Universiteit Leiden
The Netherlands

Giant galactic outflows and shocks in the cosmic web

Oei, M.S.S.L.

Citation

Oei, M. S. S. L. (2023, December 12). *Giant galactic outflows and shocks in the cosmic web*. Retrieved from https://hdl.handle.net/1887/3666253

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/3666253

Note: To cite this publication please use the final published version (if applicable).

Giant galactic outflows and shocks in the Cosmic Web

Giant galactic outflows and **shocks** in the **Cosmic Web**

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr.ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op dinsdag 12 december 2023
klokke 15.00 uur

door

Martijn Simon Soen Liong Oei

geboren te Leiden
in 1993

Promotor:	prof. dr. H. J. A. Röttgering	
Copromotor:	dr. R. J. van Weeren	

Promotiecommissie

Voorzitter:	prof. dr. I. A. G. Snellen	
Secretaris:	prof. dr. J. Schaye	
Overige leden:	prof. dr. F. Combes	Collège de France & Observatoire de Paris
	prof. dr. K. M. Blundell	University of Oxford
	prof. dr. M. J. Hardcastle	University of Hertfordshire

De *promotiecommissie* is de door het college voor promoties ingestelde commissie die beslist of de promovendus een dusdanige proeve van bekwaamheid heeft afgelegd dat deze tot de promotie kan worden toegelaten. (De *oppositiecommissie* is de commissie ten overstaan waarvan het proefschrift publiekelijk wordt verdedigd; deze omvat onder andere leden van de promotiecommissie.)

Universiteit Leiden
The Netherlands

The size and age of the Cosmos are beyond ordinary human understanding. Lost somewhere between immensity and eternity is our tiny planetary home. In a cosmic perspective, most human concerns seem insignificant, even petty. And yet our species is young and curious and brave and shows much promise. In the last few millennia we have made the most astonishing and unexpected discoveries about the Cosmos and our place within it, explorations that are exhilarating to consider. They remind us that humans have evolved to wonder, that understanding is a joy, that knowledge is prerequisite to survival. I believe our future depends powerfully on how well we understand this Cosmos in which we float like a mote of dust in the morning sky.

Carl E. Sagan, American astronomer, in *Cosmos (1980)*

To my parents, for their love and support.

© Martijn S. S. L. Oei, 2023 All rights reserved.
ISBN: 978-94-6419-999-4

Printed by Gildeprint – Enschede
A digital copy of this thesis is available at https://openaccess.leidenuniv.nl.

Design by Martijn S. S. L. Oei
The front cover shows *Alcyoneus*, a giant galactic outflow. The spine features a double dactyl composed by the author. The back cover shows the magnetic field strength within a patch of cosmic web (not necessarily *the* Cosmic Web) simulated by the TNG Collaboration. The defence invitation shows the synchrotron Milky Way.
This thesis has been typeset using LaTeX, originally developed by Leslie Lamport and based on Donald Knuth's TeX. The body text is set in Egenolff–Berner Garamond, a revival of Claude Garamont's humanist typeface. The formatting is based on the template 'Dissertate' from Jordan Suchow, which has been released under the permissive MIT (X11) license.

The work described in this thesis has been carried out at Leiden Observatory, which is part of the Faculty of Science at Leiden University. The research has been supported by the Vidi research programme with project number 639.042.729, which is financed by the Dutch Research Council (NWO).

Advisors: R. J. van Weeren & H. J. A. Röttgering **Author:** M. S. S. L. Oei

Giant galactic outflows and shocks in the Cosmic Web

Abstract

The radio sky harbours both galactic and extragalactic sources of arcminute- to degree-scale emission of various physical origins. To discover extragalactic diffuse emission in the Cosmic Web beyond galaxy clusters, one must image low–surface brightness structures amidst a sea of brighter compact fore- and background sources. Angularly separating the faint from the bright radio sky requires high-quality ionospheric calibration. This thesis introduces new advances in and investigations into ionospheric calibration, the degree-scale Milky Way foreground, and two sources of megaparsec-scale emission in cosmic filaments: giant galactic outflows and cosmological structure formation shocks. Giant galactic outflows (or 'giant radio galaxies') are generated by the jets of active supermassive black holes, and transport relativistic leptons, entrained atomic nuclei, heat, and magnetic fields from the centres of galaxies to their outskirts and beyond. These outflows embody the most energetic pathway by which galaxies respond to the Cosmic Web around them. Structure formation shocks around filaments are a generic, but still elusive, prediction of cosmological simulations and trace the gravitational flow of matter from proto-voids to filaments. Both phenomena inform on the strength, topology, and origin of magnetic fields in the Cosmic Web.

Contents

1 **Introduction** 15

1.1 Magnetism and the rise of modern physics . . . 16
1.2 Magnetism throughout the Universe . . . 18
1.3 Origin of magnetism . . . 26
1.4 Astronomical observations of magnetic phenomena . . . 31
1.5 The LOFAR . . . 31
1.6 Astronomical interferometry of electromagnetic waves: the van Cittert–Zernike theorem . . . 34
1.7 The ionosphere . . . 46
1.8 Late-time radio probes of cosmological magnetism . . . 49
1.9 This thesis . . . 61
1.10 Future directions . . . 66

2 **A probabilistic approach to direction-dependent calibration** 67

2.1 Introduction . . . 68
2.2 Ionospheric effects on interferometric visibilities . . . 70
2.3 Probabilistic relation between FED and ΔTEC: Gaussian process layer model . . . 72
2.4 Method . . . 76
2.5 Results . . . 83
2.6 Discussion . . . 86
2.7 Conclusion . . . 91
Appendix 2.A1 Derivation of tomographic equivalence . . . 93
Appendix 2.A2 ΔTEC variance function . . . 94

Appendix 2.A3 Covariance function with frozen flow . . . 97
Appendix 2.A4 Supporting figures . . . 98

3 *Filamentary Baryons and Where to Find Them:* A forecast of synchrotron radiation from merger and accretion shocks in the local Cosmic Web 101
3.1 Introduction . . . 103
3.2 Methods . . . 109
3.3 Redshift predictions from geometric cosmic web model . . . 121
3.4 Results . . . 130
3.5 Discussion . . . 140
3.6 Conclusions . . . 147
Appendix 3.A1 Additional figures . . . 148
Appendix 3.A2 Single-shock synchrotron MEC–total matter density scaling relation . . . 149
Appendix 3.A3 Ray tracing in the cosmological setting . . . 156
Appendix 3.A4 Observer's specific intensity . . . 156
Appendix 3.A5 Volume-filling fractions . . . 158
Appendix 3.A6 Notation . . . 160

4 The discovery of a radio galaxy of at least 5 Mpc 163
4.1 Introduction . . . 165
4.2 Data and methods . . . 168
4.3 Results and discussion . . . 169
4.4 Conclusion . . . 187
Appendix 4.A1 J1420-0545 comparison . . . 189
Appendix 4.A2 Inclination angle comparison . . . 190
Appendix 4.A3 Lobe volumes with truncated double cone model . . . 193

5 An intergalactic medium temperature from a giant radio galaxy 203
5.1 Introduction . . . 204
5.2 Data . . . 205
5.3 Methods and results . . . 214
5.4 Discussion . . . 228
5.5 Conclusion . . . 234
Appendix 5.A1 Average density ratio for WHIM with isothermal β-profile . . . 237
Appendix 5.A2 Large-to-small-scale density conversion: the low-density regime . . . 238
Appendix 5.A3 Lower WHIM temperature constraints from current-day radio data . . . 241

6 **Measuring the giant radio galaxy length distribution with the LoTSS** 243
6.1 Introduction . . . 245
6.2 Theory . . . 247
6.3 Sample compilation and properties . . . 268
6.4 Results . . . 281
6.5 Discussion . . . 296
6.6 Conclusions . . . 302
Appendix 6.A1 Framework derivations and details . . . 307
Appendix 6.A2 Additional images . . . 334
Appendix 6.A3 Stellar and supermassive black hole masses . . . 334
Appendix 6.A4 Surface brightness prior . . . 334
Appendix 6.A5 Likelihood function . . . 339
Appendix 6.A6 Properties of newly discovered giants . . . 341

7 **Do luminous giants populate special large-scale environments? Or: the radio luminosity–Cosmic Web density relation for radio galaxies** 343
7.1 Introduction . . . 345
7.2 Data . . . 349
7.3 Methods . . . 352
7.4 Results . . . 357
7.5 Discussion . . . 379
7.6 Conclusion . . . 391
Appendix 7.A1 Cosmic Web localisation accuracy with spectroscopic and photometric redshifts . . . 394
Appendix 7.A2 Cosmic Web density distribution: the gamma Ansatz . . . 395
Appendix 7.A3 Modelling relative baryon density measurement heteroskedasticity . . . 397
Appendix 7.A4 Radio luminosity–Cosmic Web density relation: fixed voxel method . . . 398
Appendix 7.A5 Relative number density derivation . . . 398
Appendix 7.A6 Disparity between mean squared density and squared mean density . . . 402

8 **Constraining the giant radio galaxy population with machine learning–accelerated detection and Bayesian inference** 405
8.1 Introduction . . . 407
8.2 Theory . . . 408
8.3 Data . . . 419

8.4 Methods . . . 420
8.5 Results . . . 439
8.6 Discussion . . . 445
8.7 Conclusions . . . 448
Appendix 8.A1 Curved power law PDF for L . . . 451
Appendix 8.A2 Likelihood trick . . . 452
Appendix 8.A3 PyBDSF parameters . . . 453
Appendix 8.A4 Adaptations of the radio ridgeline based host galaxy identification . . . 453
Appendix 8.A5 Sky coverages . . . 455

Appendix A **Gaussian random field ionosphere model extension: the curved Earth** 457

Appendix B **Flux scale–induced spectral index uncertainties for high-SNR sources** 463
B.1 Introduction . . . 464
B.2 Analytical results . . . 464
B.3 Numerical results . . . 468
B.4 Conclusion . . . 468

9 **Populairwetenschappelijke samenvatting** 471
9.1 Licht en elektromagnetische golven . . . 471
9.2 Radiogolven . . . 472
9.3 Het Kosmische Web . . . 474
9.4 Schokgolven in het Kosmische Web en ionosferische kalibratie . . . 476
9.5 Schokgolfvoorspellingen voor het Kosmische Web aan de noordelijke hemel . . . 478
9.6 Gevoelige LOFAR-beelden en een onverwachtse ontdekking . . . 480
9.7 Een gigantisch uitvloeisel uit een spiraalsterrenstelsel . . . 482
9.8 Een census van giganten . . . 483
9.9 De omgevingen van gigantische galactische uitvloeisels in het Kosmische Web . . . 485
9.10 Giganten, machinaal leren en magnetogenese . . . 488

10 **List of publications** 489

11 **Curriculum vitae** 491

12 **Acknowledgments** 497

Bibliography 499

Do not try to keep up with the times, try to keep up with poetry! Keep up with everything that transcends time.

Sam Dillemans, Belgian painter, in *Sam Dillemans, de Waanzin van het Detail (2007; translation from Flemish)*

1 Introduction

On our home planet, of all widespread minerals of natural origin, magnetite is most strongly magnetic. Life on Earth long ago acknowledged this power in its struggle for survival: the fossil record of Southern England shows that, when the dinosaurs of the Cretaceous roamed our world, bacteria synthesised magnetite crystals (e.g. Siponen et al., 2013) to align themselves with Earth's magnetic field and navigate[1] to zones of optimal oxygen concentration.[2] Birds — the last surviving clade of dinosaurs — have developed at least two magnetoreception systems, including a magnetite-based receptor system active within the beaks of several species (e.g. Wiltschko & Wiltschko, 2013). Thus, when China's ancient Han dynasty discovered that lodestones — magnetite fragments turned permanent magnets[3] — could be used for navigation, humans reinvented a technology already mastered by lifeforms of long-gone eras, if not eons. In the ancient Mediterranean, lodestones ('lead-stones' or 'way-stones' in Old

[1] By committing to movement along a field line, a three-dimensional search problem simplifies to a one-dimensional one.

[2] Synthesis by such magnetotactic bacteria is not the only way in which Earth's magnetite has formed. For instance, experiments demonstrate that magnetite could have precipitated directly from seawater in the deep oceans of the early Archean (Li et al., 2017).

[3] Notably, magnetite can only become permanently magnetised when its microstructure contains *maghemite* impurities (Warner, 2012). A leading theory posits that lightning, and the strong magnetic fields of lightning bolts in particular, are responsible for charging lodestones (Wasilewski & Kletetschka, 1999).

English, for their ability to guide mariners) were called Μαγνήτιδες λίθοι: Magnesian stones, presumably named after the lodestones found near the Lydian city of Magnesia ad Sipylum in Asia Minor.[4] The philosopher Thales, from the flourishing commercial city of Miletus in Ionia that led trade with Babylon and Egypt, suggested — in line with the prevailing animistic world-view of his time — that lodestones had souls that made them attract iron. Thales, now sometimes called the *Father of Science*, was rightfully intrigued, as scientific inquiry in the following millennia would reveal that the seemingly innocent stones of Magnesia held deep secrets.

1.1 Magnetism and the rise of modern physics

For example, in the early twentieth century — and within a few years of eachother — the young physicists Niels Bohr and Hendrika van Leeuwen independently showed that a consistent application of classical and statistical mechanics rules out macroscopic magnetisation for non-rotating, isolated systems in thermal equilibrium.[5] Equivalently, this result, now known as the Bohr–van Leeuwen theorem, states that the magnetic behaviour of solids must arise purely as a quantum mechanical effect. Lodestones are thus *quantum* stones. Thales's 'souls' were whispers of the subatomic world, whose description has become one of the pillars of modern physics.

Magnetism also played a key role in establishing another pillar of modern physics. In the summer of 1905, during his *annus mirabilis*, Einstein published his groundbreaking *Zur Elektrodynamik bewegter Körper*, in *Annalen der Physik*. Herein, he laid out the groundwork for his special theory of relativity, a daring reconsideration of the laws of kinematics and electrodynamics in their full generality, albeit in a setting without a gravitational field — that is, in Minkowski spacetime.[6] To do so, Einstein

[4] Scholarly debate, both in antiquity and in modern times, has been unable to establish with certainty whether Magnesia ad Sipylum in Lydia, or the district of Magnesia in Thessaly where the founders of the Lydian city originated from, lies at the root of the lodestone's Greek name (Oxford English Dictionary, 2000).

[5] Van Leeuwen wrote her doctoral thesis at Leiden University under supervision of the Nobel Prize–winning theorist Hendrik Lorentz. In particular, van Leeuwen's thesis proposed to resolve a dispute between Langevin and Kroo on the status of paramagnetism within the then-novel theory of electrons by resorting to predictions of classical and statistical mechanics only, the validity of which would be considered beyond dispute. Van Leeuwen wrote (van Leeuwen, 1919): *'Voor die methode kunnen we gebruiken die van het H-theorema van Boltzmann, maar moeten ons dan beperken tot gassen, waarin de moleculen alleen in wisselwerking zijn gedurende den korten duur van hun botsingen en tot een temperatuurgebied, waarin de wetten van de klassieke mechanica kunnen worden toegepast, zoo hoog dus, dat de quanta nog geen rol spelen.'*

[6] It is sometimes wrongfully claimed that the special theory of relativity does not describe accelerating bodies. On the contrary, directly in *Zur Elektrodynamik bewegter Körper*, Einstein considers the *Dynamik des langsam beschleunigten Elektrons* — the dynamics of the slowly accelerated electron.

required two postulates, the first of which he called the *principle of relativity*.[7] The opening sentences of *Zur Elektrodynamik bewegter Körper* indicate that Einstein's thinking on the principle had been influenced by the conceptually seemingly distinct, but equally adequate explanations offered by Maxwell's electrodynamics for the same phenomenon — a magnet and a conductor in relative motion — when view from the two bodies' respective rest frames:

> It is known that Maxwell's electrodynamics — as usually understood at the present time — when applied to moving bodies, leads to asymmetries which do not appear to be inherent in the phenomena. Take, for example, the reciprocal electrodynamic action of a magnet and a conductor. The observable phenomenon here depends only on the relative motion of the conductor and the magnet, whereas the customary view draws a sharp distinction between the two cases in which either the one or the other of these bodies is in motion. For if the magnet is in motion and the conductor at rest, there arises in the neighbourhood of the magnet an electric field with a certain definite energy, producing a current at the places where parts of the conductor are situated. But if the magnet is stationary and the conductor in motion, no electric field arises in the neighbourhood of the magnet. In the conductor, however, we find an electromotive force, to which in itself there is no corresponding energy, but which gives rise — assuming equality of relative motion in the two cases discussed — to electric currents of the same path and intensity as those produced by the electric forces in the former case.
>
> **Examples of this sort,** together with the unsuccessful attempts to discover any motion of the Earth relatively to the 'light medium', **suggest that the phenomena of electrodynamics as well as of mechanics possess no properties corresponding to the idea of absolute rest.**

The second half of *Zur Elektrodynamik bewegter Körper* introduces the idea that what manifests as an electric field to one observer, an observer in another reference frame would identify as a magnetic field — and vice versa. Thus, just as Minkowski introduced the idea of *spacetime*, a single, unified entity whose separation into space and time is reference frame–dependent, Einstein introduced the idea of *electromagnetism*.

The *special case* that 'special' refers to, is the case of negligible gravity.

[7] The unusual feature of German orthography that requires all nouns to be capitalised, prevents us from knowing whether Einstein's *Prinzip der Relativität* should be translated to English without capitals, as I do here, or *with* to emphasise its deemed importance — as is done in some well-known English translations, such as in the 1923 book *The Principle of Relativity*.

Relativity revealed that electric and magnetic phenomena are intimately linked, as the two faces of *Ianus Bifrons*.

1.2 Magnetism throughout the Universe

1.2.1 The Earth — and the dynamo mechanism

The Earth's magnetic field is not just helpful to lifeforms seeking to navigate the planet's oceans and skies, but more fundamentally supports — if not outright makes possible in the first place — the existence of life as we know it. By deflecting the Solar wind, the *geomagnetic* field prevents the Earth's atmosphere from eroding away. Only where geomagnetic field lines leave or enter the Earth, Solar wind protons and electrons ionise atmospheric nitrogen molecules and oxygen atoms, giving rise to the mythical phenomena of *aurora borealis* and *aurora australis*. Figure 1.1 shows a view of the former, as documented by the author of this thesis during a cold Icelandic night in December 2019.

The geomagnetic field can be traced back to the core, where it is amplified through the dynamo mechanism. This mechanism is driven by the flow of heat from the inner core, whose temperature of 6000 K exceeds that of the surface[8] of the Sun, to the outer core–mantle boundary, where the temperature is 3800 K. This temperature gradient, alongside material inhomogeneity, causes buoyancy and so ultimately convection currents (of current density $\mathbf{J}$) in the outer core's iron alloy fluid. By Ampère's circuital law, these convection currents generate a magnetic field $\mathbf{B}$:[9]

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t}. \tag{1.1}$$

Faraday's law, in turn, dictates that a changing magnetic field induces an electric field $\mathbf{E}$:

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}. \tag{1.2}$$

Finally, these magnetic and electric fields exert Lorentz force on the charged particles

[8] We take the 'surface' of the Sun to mean its *visual* surface: i.e. its photosphere, from which light escapes directly into space.

[9] The emergence of magnetic fields due to currents as presently discussed only requires Ampère's circuital law as Maxwell originally derived it — i.e. *without* his later correction. However, as the terminology suggests, Maxwell's equations are not complete without this additional term — which, for instance, is crucial for predicting electromagnetic wave propagation. For this reason, we provide Ampère's circuital law *with* Maxwell's correction here.

Figure 1.1: Magnetic fields around planets do not only generate dazzling auroral displays, such as those above Iceland *(top)*, but also allow for the existence of rather sophisticated lifeforms, such as astronomers *(bottom left)* and Labradors *(bottom right)*.

that make up the currents. The currents are shaped by the Lorentz force, in accordance with Ohm's law:

$$\mathbf{J} = \sigma \left(\mathbf{E} + \mathbf{u} \times \mathbf{B} \right) , \tag{1.3}$$

where σ is the electrical conductivity and $\mathbf{u}$ is the fluid velocity field. To see how the interplay between currents, magnetic fields, and electric fields give rise to a magnetic field that grows over time, we start with Faraday's law, and substitute in Ohm's law

after solving the latter for $\mathbf{E}$. This results in

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times \left(\mathbf{u} \times \mathbf{B} - \sigma^{-1}\mathbf{J}\right). \tag{1.4}$$

Now, by using Ampère's circuital law with Maxwell's correction to take out $\mathbf{J}$, and after introducing the magnetic diffusivity $\eta = (\mu_0\sigma)^{-1}$, we find

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times \left(\mathbf{u} \times \mathbf{B} - \eta\nabla \times \mathbf{B} + \frac{\eta}{c^2}\frac{\partial \mathbf{E}}{\partial t}\right). \tag{1.5}$$

From Eq. 1.5, it is clear that, if Maxwell's correction (i.e. the second term on the RHS of Eq. 1.1) could be neglected, we obtain a partial differential equation for the evolution of $\mathbf{B}$ only. When is neglecting Maxwell's correction warranted? Clearly, this is warranted when the first two terms between parentheses are much larger than the third, yielding the condition

$$||\mathbf{u} \times \mathbf{B} - \eta\nabla \times \mathbf{B}||_2 \gg ||\frac{\eta}{c^2}\frac{\partial \mathbf{E}}{\partial t}||_2, \tag{1.6}$$

where $||\mathbf{x}||_2$ is the usual Euclidean norm of vector $\mathbf{x}$. By comparing Eq. 1.2 to Eq. 1.5, we see that $\mathbf{u} \times \mathbf{B} - \eta\nabla \times \mathbf{B} = -\mathbf{E} - \frac{\eta}{c^2}\frac{\partial \mathbf{E}}{\partial t}$. We now plug this into the inequality, and use the fact that vector lengths do not change when vectors are reflected in the origin. As a result,

$$||\mathbf{E} + \frac{\eta}{c^2}\frac{\partial \mathbf{E}}{\partial t}||_2 \gg ||\frac{\eta}{c^2}\frac{\partial \mathbf{E}}{\partial t}||_2. \tag{1.7}$$

This is an inequality of the form $||\mathbf{x}+\mathbf{y}||_2 \gg ||\mathbf{y}||_2$. Loosely speaking, $\mathbf{x}+\mathbf{y}$ is at risk of being *shorter* than $\mathbf{y}$ if $\mathbf{x}$ and $\mathbf{y}$ are about as long and roughly antiparallel, so that they can cancel eachother. Inversely, when $\mathbf{x}$ is much longer than $\mathbf{y}$, then there is no such risk: $\mathbf{y}$ now takes on the role of a mere 'perturbation', and $\mathbf{x} + \mathbf{y}$ will consequently remain much longer than $\mathbf{y}$, irrespective of the orientation of $\mathbf{y}$ with respect to $\mathbf{x}$. Therefore, the inequality of Eq. 1.7 is satisfied whenever

$$||\mathbf{E}||_2 \gg ||\frac{\eta}{c^2}\frac{\partial \mathbf{E}}{\partial t}||_2, \tag{1.8}$$

or

$$\frac{||\mathbf{E}||_2}{||\frac{\partial \mathbf{E}}{\partial t}||_2} \gg \frac{\eta}{c^2} =: \tau_{\mathrm{F}}, \tag{1.9}$$

where τ_F is the Faraday time. We remark that the Faraday time depends only on the electrical conductivity σ and the fundamental constants of Nature μ_0 and c. By calculating σ for various physical systems, one can show (e.g. Brandenburg & Subramanian, 2005) that the inequality is satisfied in a wide range of astrophysical contexts, including that of Earth's outer core. We thus arrive at the *induction equation* of magnetohydrodynamics:

$$\frac{\partial \mathbf{B}}{\partial t} \approx \nabla \times (\mathbf{u} \times \mathbf{B} - \eta \nabla \times \mathbf{B}) \,. \tag{1.10}$$

The two terms on the RHS of Eq. 1.10 are called the *magnetic induction term* and the *magnetic diffusion term,* respectively. The *magnetic Reynolds number* R_m expresses the relative importance of both terms, and is defined simply as the ratio of the typical magnitudes of the induction and diffusion terms. Because the diffusion term $\eta \nabla \times \mathbf{B} \propto \eta$, it follows that $R_m \propto \eta^{-1} \propto \sigma$: the higher the electrical conductivity, the higher the magnetic Reynolds number. What are typical magnetic Reynolds numbers in astrophysical scenarios? In the accretion discs of cataclysmic variable stars, stellar mass black holes, and neutron stars, $R_m \sim 10^4$, while in the accretion discs of supermassive black holes, $R_m \sim 10^{11}$. High values of R_m are also found in the convection zones of Sun-like stars, where $R_m \sim 10^6$–10^9. Finally, in galaxies at large, and in galaxy clusters, R_m is orders of magnitude higher still. These numbers suggests that the *perfectly conducting fluid* approximation has widespread validity in astrophysics. Such fluids, as their name implies, are characterised by $\sigma \to \infty$, or $\eta \to 0$. As a result, the induction equation for perfectly conducting fluids is

$$\frac{\partial \mathbf{B}}{\partial t} \approx \nabla \times (\mathbf{u} \times \mathbf{B}) \,. \tag{1.11}$$

Both this simple induction equation as well as Eq. 1.10's more general form demonstrate that $\mathbf{B}(\mathbf{x}, t) = 0$ is a perfectly valid dynamo solution. Thus, when magnetic fields are initially absent, the dynamo mechanism does *not* generate them. This makes explicit the important fact that the dynamo mechanism can amplify magnetic fields, but always requires a seed magnetic field to get started.

For the Earth, the seed magnetic field possibly came from a time when the Sun was very young. In a putative T Tauri phase — which would have lasted for only ten millions years or so — the contracting Sun launched a stellar wind that was much more strongly magnetised than the one observed today. It is hypothesised that this early, magnetised Solar wind transferred both angular momentum and magnetic fields to the protoplanetary disc from which the Solar System planets, including the Earth,

eventually arose.

1.2.2 Other planets, in the Solar System and beyond

In the Solar System, Mars has suffered the fate of losing its atmosphere to the interplanetary medium, changing its sprawling deltas, fans, and channels of liquid water into cold and dry deserts. This, of course, has led to a strong decrease in its potential to host life as we know it. Although the early Martian magnetic field was never able to shield the planet's atmospheric hydrogen from the young Sun's relentless winds, the stripping of atmospheric carbon and oxygen only began in earnest when the Martian magnetic field disappeared 4.1 billions years ago. With the death of the planet's dynamo, the Martian atmosphere became vulnerable to slow but steady atmospheric erosion. Inadequate replenishment by processes such as cometary bombardment and vulcanism eventually led to its feeble current-day atmosphere, whose surface pressure is less than 1% of that on Earth.

The cautious tale of Mars[10] has motivated astronomers in recent years to study more closely the interplay between planetary atmospheres, planetary magnetic fields, and stellar winds (e.g. Rodríguez-Mozos & Moya, 2019). It is now believed that the habitability of exoplanets should not be judged merely on their ability to allow liquid water on their surface, but also on their ability to sustain a magnetic field.[11]

1.2.3 Stars, and islands of stars

Magnetism appears to be of existential importance not only to habitable planets, but also to stars. Magnetic fields in collapsing protostellar clouds are crucial to transferring away angular momentum. Without extraction of angular momentum, such clouds cannot give birth to stars.

The influence of magnetism on the level of stellar populations is an area of active research (for a review, see e.g. Krumholz & Federrath, 2019). The two key quantities via which astronomers quantify the outcome of the star formation process, both in

[10] This sentiment might not be echoed by antinatalists — and especially not by those with views as sweeping as David Benatar's (e.g. Benatar, 2006).

[11] If the Solar System provides some indication, exomoons may outnumber exoplanets. Thus, if the habitable fraction of exomoons is comparable to that of exoplanets, exomoons would form the Universe's most common rostrum for the emergence of life. Uniquely, tidal heating could drive the habitability of exomoons. This mechanism cannot be important for exoplanets, as the tidal heating power P scales with semi-major axis a as $P \propto a^{-6}$ or steeper (e.g. Makarov & Efroimsky, 2014) — and planets are, of course, much further away from their stars than moons are from their planets. Recent work has explored the possibility that tidal heating creates subsurface oceans on exomoons (Tjoa et al., 2020), in which life could emerge even when magnetically protected atmospheres are absent.

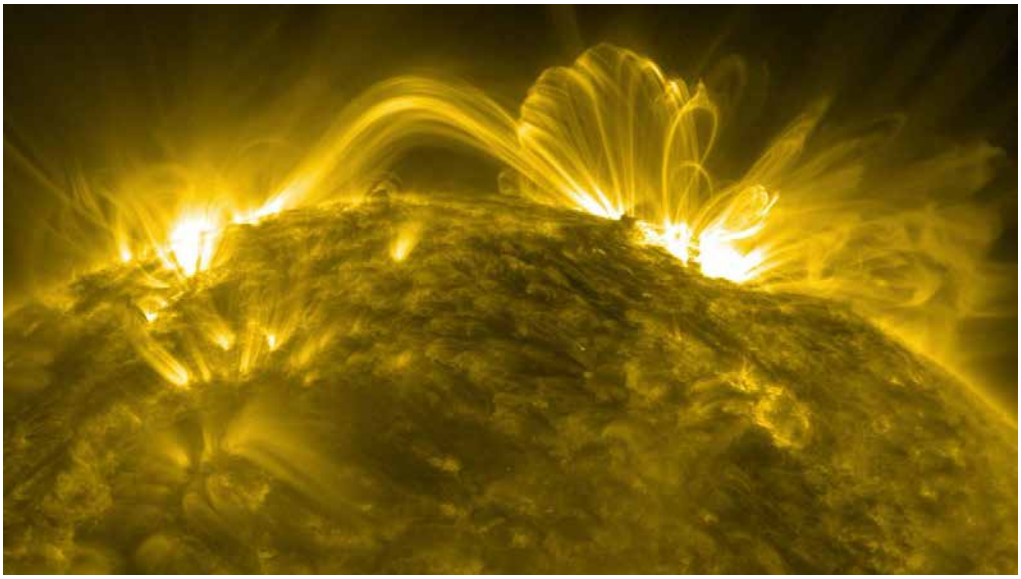


Figure 1.2: The Sun is the sizzling heart of the Solar System, and the only star that we can study up close. A magnetically active star, the Sun's atmosphere features coronal loops: radiating magnetic flux tubes that begin and end in the photosphere and that are up to a million kilometres long. These magnetic flux tubes can have complex shapes, and often appear thinner than they truly are (Malanushenko et al., 2022). *From: Solar Dynamics Observatory, NASA*

individual molecular clouds and in entire galaxies, are the star formation rate (SFR) and the initial mass function (IMF) of stars.[12] Remarkably, following the pioneering work of Salpeter (1955), careful observations of nascent stars in different regions of the Milky Way have suggested a nearly universal IMF. For more than half a century after Salpeter (1955), no consensus appeared as to whether the IMFs in other galaxies, both past and present, deviates from that of the Milky Way, with a common modelling assumption being that they do not. Historically, magnetic fields have been proposed as regulators of the fragmentation of collapsing gas clouds, with the goal of explaining the possible near-universality of the IMF. While work of the last two decades has shown that magnetic fields in star-forming regions are too weak to fulfill this role directly, magnetism may crucially impact star formation *indirectly* — for example, by allowing the formation of stellar jets and other outflows. In addition, magnetic fields may dramatically reduce the exchange of heat and particles between

[12] Regardless of its value, a star's initial mass is the prime driver of its physical properties, such as its bolometric luminosity, surface temperature, and radius, at any given instant of its ensuing life. Moreover, the initial mass determines the star's eventual fate, the type of remnant left behind, and the elemental abundances and energy returned back to the interstellar medium (ISM). For this reason, the IMF is essential to understanding stellar populations.

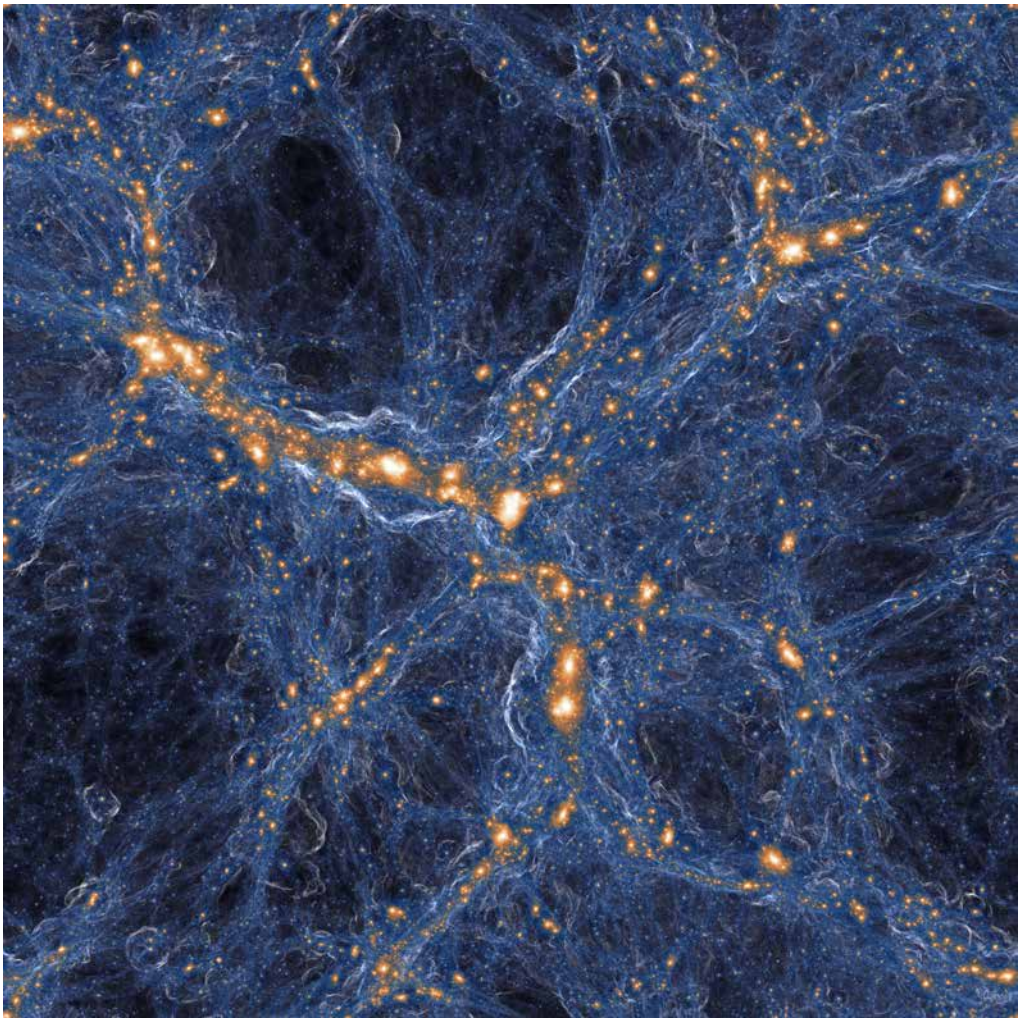


Figure 1.3: Computer simulation of the Cosmic Web — the grand, all-encompassing, network-like structure of the Universe that has been fourteen billion years in the making. With dynamics dominated by gravity, individual galaxies, galaxy groups, and galaxy clusters *(orange–white)* form in the depths of the Cosmic Web's potential wells. Supersonic shocks crash onto the filaments of the Cosmic Web *(blue–white)*, highlighting their ongoing formation. *From: TNG100 simulation, TNG Collaboration*

hot gas from supernova-driven winds and the cold ISM. If so, this would drastically change our understanding of the effectiveness of supernova feedback on the SFR.

1.2.4 The Cosmic Web

In the last decades, observations have shown that the Universe is magnetised even on its largest scales. It might not be a priori clear that the Universe even possesses a *largest*

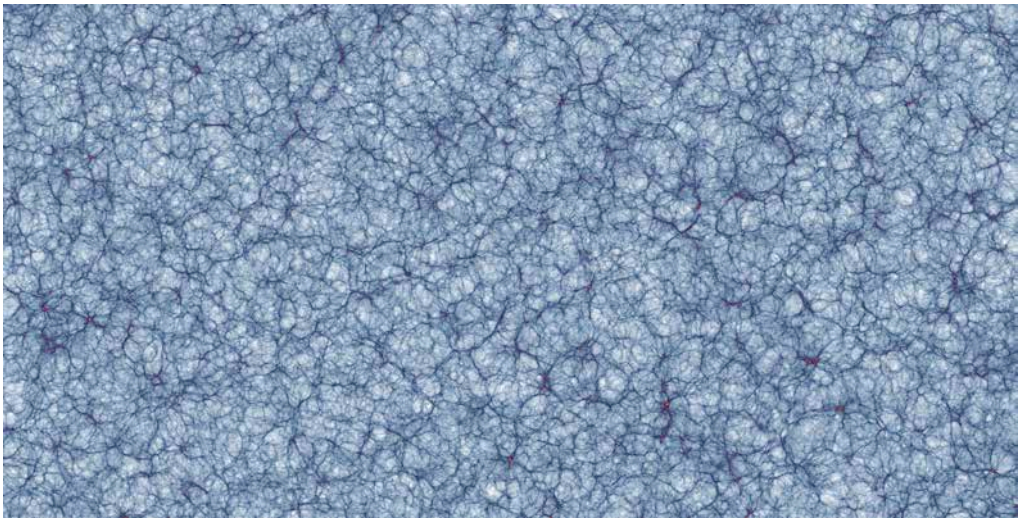


Figure 1.4: Computer simulation of the Cosmic Web at scales exceeding the End of Greatness, demonstrating the *Cosmological Principle*: at sufficiently large scales, the Universe becomes statistically homogeneous and isotropic. *From: Millennium XXL simulation, Angulo et al. (2012)*

scale, though — which in this context would refer to a spatial scale beyond which the Universe does not show additional organisation. To more clearly illustrate the meaning of such a largest scale, let us consider a Gedankenexperiment in which Hinduism's Vishnu, its primary creator, orders Brahma to build *more* of the Universe. Brahma is asked, however, to do so in a way that stays faithful to Vishnu's existing design. As to avoid repetitiveness, he need not create exact copies of the structures already out there (which would lead to a crystalline universe), but is allowed to introduce some stochasticity. One could now ask how large a volume Brahma should survey in order to learn all he needs to extend the Universe in a statistically sound manner — that is, without *ignoring* any patterns present in the current Universe, and without *introducing* any patterns that are not yet present. The answer is that this volume should measure, along each spatial dimension, roughly $1.5 \cdot 10^2$ Mpc (e.g. Cole et al., 2005; Eisenstein et al., 2005) — that is, hundreds of millions of lightyears. These are the values where, in this context, the *largest scale* refers to. The Universe's final formations on the spatial hierarchy are clusters, filaments, sheets, and voids, which together form a majestic, perennially evolving network called the Cosmic Web — sometimes fittingly dubbed the *End of Greatness*.

The Cosmic Web was first discovered by mapping out the locations of galaxies in three dimensions using their right ascensions, declinations, redshifts, and the Hubble–Lemaître law. This revealed that galaxies are not scattered uniformly throughout space, but that they organise along filament-like formations. Where filaments meet,

hundreds — and sometimes thousands — of galaxies lie clustered together. As Fig. 1.3 illustrates, modern computer simulations reproduce this spatial distribution of galaxies. They show that the Web's striking morphology is determined principally by the Gaussian random field initial conditions of the matter density fields, the collisionless gravitational dynamics of dark matter, and the expansion of the Universe, which could loosely be regarded as a force counteracting gravity. The numerical value of the End of Greatness is set by the finite age of the Universe, the finite speed of light, the strength of gravity, and the speed of expansion — that is, by the constants c, G, and H_0. Computer simulations also reproduce the End of Greatness, as shown in Fig. 1.4.

The detection of synchrotron radiation from halos and merger shocks in galaxy clusters, the nodes of the Cosmic Web, have shown that the densest regions of the Cosmic Web are magnetised at $B \sim 1\,\mu$G levels. At the moment of writing, the first synchrotron detections of the IGM within filaments have been claimed, showing that observations of large-scale structure beyond galaxy clusters are within the reach of modern low-frequency radio telescopes. Most notably, Botteon et al. (2018), Govoni et al. (2019), and Botteon et al. (2020b) report the discovery of two *radio bridges*: highly compressed filaments between clusters bound to merge in the near cosmological future. Very recently, Vernstrom et al. (2023) have presented statistical evidence for polarised radiation from accretion shocks in filaments. Finally, observations of distant, gamma ray–emitting blazars have revealed that even cosmic voids are magnetised (e.g. Neronov & Vovk, 2010; Acciari et al., 2023).

1.3 Origin of magnetism

The ubiquity of magnetism, both spatially and through time, is interpreted by many as a hint that magnetism has a unified origin — and additionally, that this origin must lie in the Early Universe. Scenarios in which cosmic magnetism already arose in the Early Universe are called *primordial magnetogenesis* scenarios. These scenarios vary wildly in their proposed creation mechanisms, amplification mechanisms, and therefore in the epochs during which they are supposed to operate. Some scenarios remain within the confines of Standard Model physics, while some venture outside (e.g. Kandus et al., 2011). I introduce primordial magnetogenesis further in Sect. 1.3.1.

Alternatively, the Universe's magnetism might stem from the post-recombination epoch. In these cases, the Biermann battery is often proposed as responsible for generating a weak magnetic field ($B \sim 10^{-21}$ G) from an initially absent one (e.g. Attia et al., 2021). In the laboratory, modern high-power laser systems have verified the reality of the Biermann battery mechanism (Gregori et al., 2012). In brief, the Biermann battery is a process in which microscopic currents arise from the difference in mass be-

tween electrons and positively charged ions on the one hand, and plasma pressure gradients on the other. Let us consider, for example, the simplest case of a pure hydrogen plasma, which are ubiquitous in astrophysics. As electrons have a mass that is three orders of magnitude lower than that of protons, any pressure gradient in the plasma will make electrons accelerate more by the same three orders of magnitude. This creates charge separation, and thus an electric field. A magnetic field is only created *ex nihilo*, however, when the electron density gradient and the plasma temperature gradient are non-parallel (for more details, see e.g. Brandenburg & Subramanian, 2005). The seed fields generated by the Biermann battery are in post-recombination magnetogenesis scenarios amplified by the dynamo mechanism, which I introduced in Sect. 1.2.1. In astrophysical magnetogenesis scenarios, dynamo-amplified magnetism in stars and discs of supermassive black holes is spread back into intergalactic space by exploding stars or jet streams. These ejectic fields can subsequently form the seeds for magnetism in other stars and galaxies. I introduce astrophysical magnetogenesis further in Sect. 1.3.2.

1.3.1 Primordial magnetogenesis

Turner & Widrow (1988) were the first to explore the possibility that primordial magnetic fields formed during cosmic inflation. This allows for a microphysical origin for magnetic fields that have a megaparsec-scale coherence length in the more recent Universe. The characteristic feature of inflation is its 'de Sitter phase', named so because the Universe's expansion during this phase resembles that of a de Sitter universe. A de Sitter universe is an idealised, matter-free universe first studied by Leiden Observatory director Willem de Sitter, whose expansion is determined by a positive cosmological constant. During the de Sitter phase of inflation, as in the de Sitter phase of the Universe expected to occur in the far future, the Universe's scale factor increases exponentially with time. The total increase in scale factor caused by inflation is $\sim 10^{30}$. Classical electromagnetism in spatially flat Friedmann–Lemaître–Robinson–Walker (FLRW) universes obeys magnetic flux conservation.[13] Concerningly, magnetic flux conservation dictates that magnetic field strengths B decay with the scale factor a over time as $B(t) \propto a^{-2}(t)$.[14] Under magnetic flux conservation, initial magnetic fields are thus reduced by a formidable factor $\sim 10^{-60}$, making these primordial magnetic fields cosmologically irrelevant. As a result, any cosmologically *relevant* primordial

[13] Together, the conformal invariance of classical electromagnetism and the conformal flatness of *flat* FLRW universes ensure that the magnetic flux conservation result of Minkowski space carries over. Non-flat FLRW universes need not obey magnetic flux conservation, as first pointed out in Tsagas (2007) and worked out further in Barrow et al. (2012).

[14] For a brief derivation, we refer the reader to Appendix 3.A2 of Chapter 3.

magnetic fields that originate from before inflation either necessitate non-zero curvature, or extensions to classical electromagnetism that break its conformal invariance. Constraints on primordial magnetism could thus offer a window into the possibly exotic physics of the inflation era.

Ideas fully in the realm of the Standard Model have been proposed that generate seed magnetic fields of significant strengths in the post-inflation, pre-recombination era. These ideas typically revolve around the generation of magnetic fields during the electroweak or quantum chromodynamical phase transitions that happened in the first fraction of a second of the Universe's existence. A problem is, however, that such post-inflation seed fields have coherence lengths that are too small to give rise to magnetism observed at the present day. For example, magnetic fields generated during the electroweak phase transition have astronomical unit–scale coherence lengths (Kandus et al., 2011).

Intruigingly, primordial magnetism strong enough to produce the magnetic fields of modern galaxy clusters *without* dynamo amplification causes baryon inhomogeneities in the Early Universe that resolve the hotly debated Hubble tension — without the need to extend the concordance ΛCDM cosmological model (Jedamzik & Pogosian, 2020).

1.3.2 Astrophysical magnetogenesis

In most post-recombination magnetogenesis scenarios, magnetic fields are created by the Biermann battery and subsequently amplified by gravitational compression and dynamos. *Astrophysical* magnetogenesis scenarios concern the subsequent seeding of the Cosmic Web by violent forms of release of the astrophysically amplified magnetic fields. Two typical carriers of the released fields are supernova shocks and jets launched from the accretion discs of supermassive black holes. The latter carrier will be studied in detail in this thesis.

Supernovae

The first generation of stars to have formed throughout the Universe, the so-called *Population III stars,* could have seeded the IGM with magnetic fields upon their cataclysmic demise as supernovae. The strength of the seed magnetic fields present when stars form, determines the ability of their protostellar clouds to lose angular momentum, and thus affects the stellar initial mass distribution (e.g. Xu et al., 2008). We note that the seed fields of Population III stars could be the result of the Biermann battery and gravitational compression only, whilst the seed fields of later generations of stars could be much higher as a result of dynamos. For this reason, Population III stars

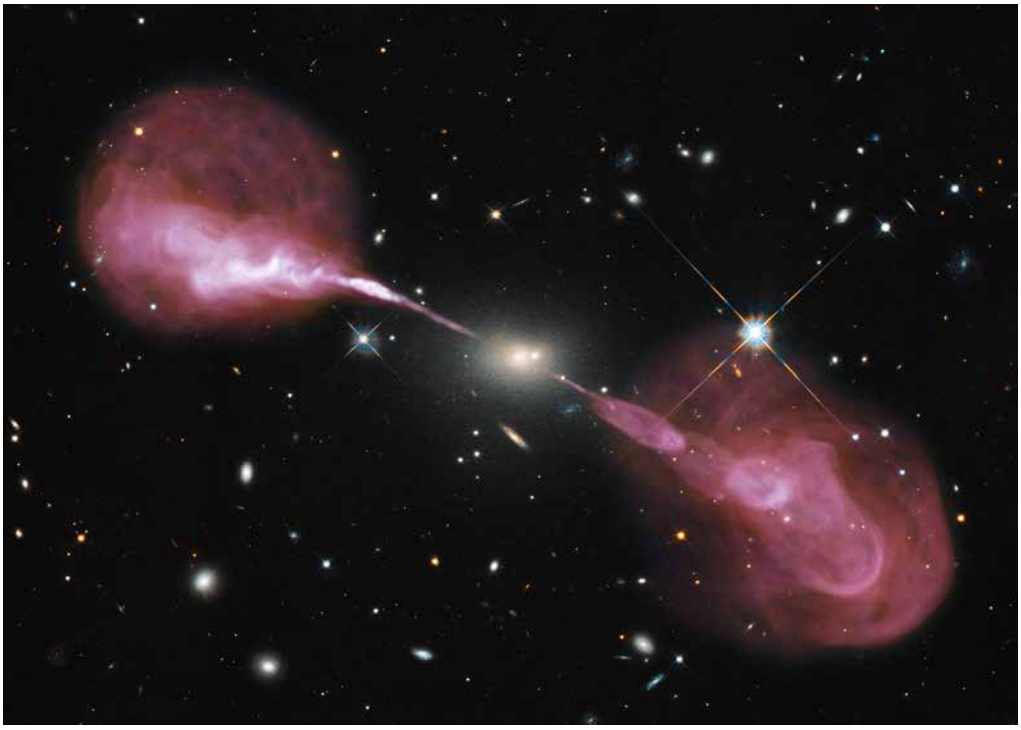


Figure 1.5: The supergiant elliptical galaxy 3C 348, whose *Hubble Space Telescope*–collected stellar light fills the image's centre, hosts a supermassive black hole that launches a pair of collimated jets. These jets carry plasma and magnetic fields into the intergalactic medium. The resulting plumes span more than a million light-years, and bask in synchrotron radiation detected by the Very Large Array. The radio-emitting structure is known as Hercules A. *From: NASA, ESA, S. Baum and C. O'Dea (RIT), the Hubble Heritage Team (STScI / AURA); R. Perley and W. Cotton (NRAO / AUI / NSF)*

were more massive than later generations of stars.[15] Currently, no reliable numerical predictions exist of the magnetogenesis potential of supernovae, as the processes that generate and amplify supernova shock magnetic fields are uncertain and require temporal and spatial resolutions far beyond those achievable in today's cosmological simulations (e.g. Garaldi et al., 2021).

Jets from supermassive black holes

Jet-driven galactic outflows — also known as radio galaxies (RGs) — are brought forth by supermassive black holes (SMBHs), and embody another possible pathway

[15] However, the most important reason that Population III stars were more massive than stars forming today is that the protostellar clouds from which they arose lacked elements heavier than lithium — especially carbon, oxygen, and silicon — which otherwise contribute to cloud collapse via radiative cooling.

for the post-recombination magnetisation of the intergalactic medium (IGM). Figure 1.5 shows a famous example, *Hercules A,* discovered in the early days of radio astronomy (Bolton, 1948). A favourable cocktail of 10^1 Jy–scale radio flux densities, a nearly two million light-year extent, and a comparatively low distance of some two billion light-years allowed for the creation of highly resolved radio maps (Dreher & Feigelson, 1984) when few others were available.[16] As a result, Hercules A was influential in the development of ideas on the evolution of radio galaxies (e.g. Mason et al., 1988) — and on the role played by the central engines that power them.

Our understanding of the importance and ubiquity of black holes has changed markedly in the 107 years since Schwarzschild found the solution to the Einstein field equations that now bears his name. Notwithstanding Einstein's own scepticism towards the astrophysical reality of Schwarzschild black holes (Einstein, 1939), Hawking and Penrose proved that black hole formation is a natural terminus of the lives of massive stars. Soon after, Penrose, Blandford, Znajek, and Rees made plausible that Kerr black holes are also the central agents in active galactic nuclei (e.g. Blandford & Znajek, 1977; Rees, 1984). From observations of radio galaxies, quasars, stellar orbits in the Galactic Centre, maser discs, reverberation mapping time delays, pre-merger gravitational waves, and most recently, the M87* and Sgr A* event horizon shadows in direct VLBI imagery (e.g. Event Horizon Telescope Collaboration et al., 2019a, 2022), it is now clear that black holes are fundamental building blocks of our Universe — on stellar, galactic, and cosmological scales alike. SMBHs occur in nearly every galaxy, and their influence over their galactic and extragalactic environments can be profound. Intriguingly, this influence reaches the gargantuan scale of the Cosmic Web — with its sprawling clusters, filaments, sheets, and voids.

The largest RGs, giants (or GRGs), attain megaparsec (Mpc) lengths (for a review, see Dabhade et al., 2023). Giants embody the most extreme known mechanism by which galaxies impact the Cosmic Web around them. Giants affect the thermodynamics of the intergalactic medium, of which they are simultaneously a probe. In this thesis, we show for the first time that giants can be used to estimate the IGM temperature in filaments. Giants in filaments may also be responsible for augmenting weak primordial magnetic fields. The extent of this giant-induced magnetism is of major interest to the study of magnetogenesis, as magnetic fields in filaments today tightly trace Early Universe magnetism if the GRG contribution is small. In this thesis, we pioneer the first steps of measuring giant-induced magnetism.

[16] For a presentation and analysis of the most recent radio imagery of Hercules A, see Timmerman et al. (2022).

1.4 Astronomical observations of magnetic phenomena

With the case made that cosmic magnetism represents a compelling astronomical frontier, what methods would be best suited to study it? Of course, cosmic magnetism can be studied in the laboratory (e.g. Gregori et al., 2012) or in simulations (e.g. Vazza et al., 2015, 2017). For observations, it appears reasonable to choose a messenger that bears signatures of magnetism in the most direct way possible. Within the electromagnetic spectrum, radio waves are perhaps the most promising magnetic messengers. They are generated by radiation mechanisms that operate only in the presence of magnetic fields, such as the cyclotron, electron–cyclotron maser instability (for a review, see Treumann, 2006), and (gyro-)synchrotron radiation mechanisms. In addition, by the Faraday effect, the polarisation direction of radio waves is rotated by the presence of magnetic fields along their journey from source to observer.

In this thesis, we will study magnetism through the radio window, and in particular through synchrotron radiation. Synchrotron radiation emerges when charged particles with relativistic speeds spiral along magnetic field lines. As predicted by the relativistic Larmor formula, the bolometric power emitted by a synchrotron-radiating particle $P \propto m^{-2}$, where m is the particle's mass (for derivations, see e.g. Rybicki & Lightman, 1986). For this reason, astronomical synchrotron radiation is dominated by that of lightest charged leptons — electrons and positrons — rather than that of, for example, muons and antimuons, or of charged baryons such as protons and antiprotons. When, say, *electrons* emit synchrotron radiation, they pay for the released energy with their kinetic energy. As the electrons slow down, their synchrotron spectrum shifts to lower frequencies. This, in brief, motivates why low-frequency ($\nu_{\rm obs} \sim 10^1$–10^2 MHz) radio telescopes are best suited to observe electron populations in magnetic fields long after they have been accelerated to relativistic speeds by supernovae or SMBH jets.

1.5 The LOFAR

The world's premier low-frequency observatory in operation today is the Low-Frequency Array (LOFAR; van Haarlem et al., 2013), initially envisioned in the summer of 1997 by George Miley, then the director of Leiden Observatory (for a personal history of the early years, see Miley, 2010). This telescope observes the Northern Sky at metre and decametre wavelengths. Its heart lies near the village of Exloo (Old Drents for *'forest of the oak'*, although the forest has since disappeared) in the rural Dutch province of Drenthe. Here, in a quiet swathe of wetland, plants and animals — such as orchids, marsh lousewort, the black-tailed godwit, and land and water rails

Figure 1.6: The LOFAR's core lies in the moors of Drenthe, one of the Netherlands' northern provinces. The *Superterp,* shown here, is a circular, elevated island with a diameter of 320 metres that contains 6 of the 24 core stations. Each core station comprises 96 low-band antennae and two sets of 24 high-band antenna tiles. The data underlying this thesis were collected using the high-band antennae only. In the bottom image, north is at the top. *From: ASTRON drone footage, 2018 (top); Kadaster aerial imagery, 2022 (bottom)*

— thrive alongside science. The area hosts the so-called *Superterp,* shown in Fig. 1.6. When the LOFAR was built, its design, consisting of $\sim 10^4$ low-cost dipole antennae linked by analog electronics (in a first processing step) and then digitally (in a subsequent processing step) to form a single telescope, was considered radical. It is organised in 24 *core* stations around Exloo, 14 *remote* stations spread throughout the Netherlands, and 14 *international* stations spread throughout Europe. When com-

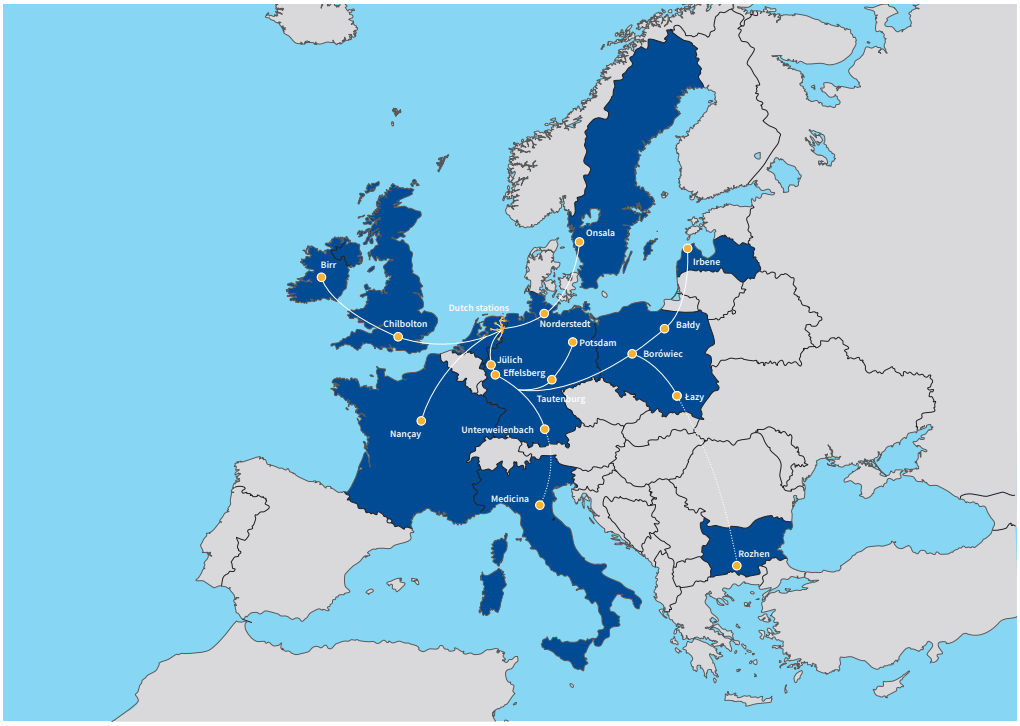


Figure 1.7: Stations of the International LOFAR Telescope, the world's premier low-frequency radio observatory, as of 2022. The pan-European network, soon to operate in ten countries, enables observations of the Universe at metre wavelengths and at resolutions comparable to those achieved by the *James Webb Space Telescope* (i.e. $\theta_{\mathrm{FWHM}} \sim 10^{-1}$ arcsec). *From: ASTRON*

bining data from the Dutch stations as well as from the non-Dutch stations, the telescope is often referred to as the International LOFAR Telescope (ILT). In Fig. 1.7, we show the stations that together form the ILT. At metre wavelengths, the ILT is able to achieve sub-arcsecond resolution — a specification unsurpassed by any other existing or planned low-frequency observatory (Morabito et al., 2022).[17] The telescope achieves this feat through aperture synthesis interferometry, a technique for which Martin Ryle of the University of Cambridge was awarded the 1974 Nobel Prize in Physics. Aperture synthesis interferometry combines the insight that pairs of recordings of the electric field at distinct spatial locations encode the sky's intensity function — a profound result known as the van Cittert–Zernike theorem — with the fact that the Earth rotation timescale is typically much shorter than the timescale over which the radio sky changes. This enables the creation of an effective telescope whose diameter is equal to the largest distance amongst pairs of stations — for a source at the

[17] The ILT is therefore not only a *pathfinder* for the Square Kilometre Array (SKA; Dewdney et al., 2009), but will remain complementary to it.

zenith, at least. As the van Cittert–Zernike theorem lies at the basis of this thesis' ionospheric calibration efforts and radio imagery, I now present an original (though somewhat simplified) derivation that is hopefully instructive to some.

1.6 Astronomical interferometry of electromagnetic waves: the van Cittert–Zernike theorem

Treating the wavefronts emitted by astronomical sources as spheres amounts to a particularly effective idealisation. Very far from the source, these wavefronts appear flat — at least locally. More quantitatively, the distance one must travel from one location to another to perceive appreciable wavefront curvature is proportional to the distance to the source. For this reason, we can model incoming electromagnetic radiation *from astronomical sources* with plane waves.

A general description of the (real) electric field displacement E_{real} induced by a plane wave at location $\vec{r} \in \mathbb{R}^3$ at time $t \in \mathbb{R}$ travelling in direction $\hat{k} \in \mathbb{S}^2$ is

$$E_{\text{real}}(\vec{r}, t) = A \cos(2\pi \frac{(\vec{r} \cdot \hat{k} + ct)}{\lambda} + \varphi). \tag{1.12}$$

Figure 1.8 shows a corresponding sequence of planes with maximum electric field displacement A. We emphasize that, for full generality, Eq. 1.12 requires a phase φ in the argument of the cosine. Without φ, Eq. 1.12 predicts that E_{real} always has maximum displacement (i.e. a displacement equal to the amplitude) in the spatiotemporal origin, where $\vec{r} = 0$ and $t = 0$. This is clearly not the most general plane wave: a general plane wave should be able to have an arbitrary displacement between (and including) $-A$ and A in the origin.

It is instructive to note that all points in a plane perpendicular to $\hat{k}$ have the same electric field displacement. To see why this is true, let $\vec{r}$ denote some point in an arbitrary plane perpendicular to $\hat{k}$. *Any* point in this plane is given by $\vec{r} + \vec{r}_\perp$, where $\vec{r}_\perp$ is an arbitrary vector perpendicular to $\hat{k}$. Because the dot product of $\hat{k}$ with any vector perpendicular to it, vanishes, the locations $\vec{r}$ and $\vec{r} + \vec{r}_\perp$ (at any t) give rise to the same

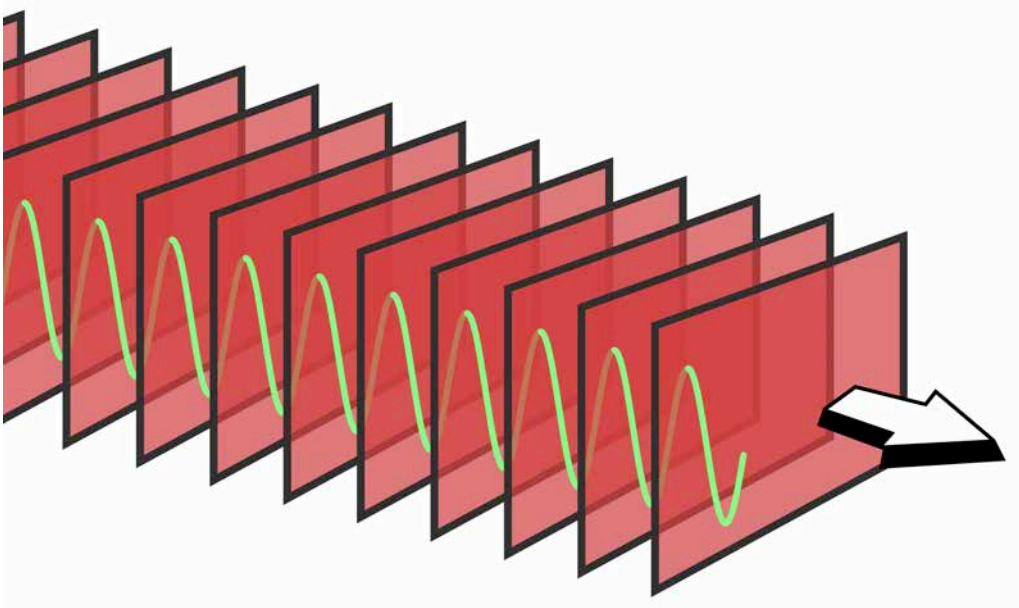


Figure 1.8: Schematic of a linearly polarised monochromatic plane wave travelling through three-dimensional space in the direction of the arrow, $\hat{k}$. The parallel planes perpendicular to $\hat{k}$ appear 'truncated', but in fact extend to infinity in this physical idealisation. This sequence of planes comprises the set of points $\vec{r}$ where, at the depicted instant in time t, E_{real} equals its amplitude A — or, in other words, the set of points that solve $E_{\text{real}}(\vec{r}, t) = A$. Subsequent planes are a distance λ apart. The green curve indicates the magnitude and orientation of the electric field displacement vectors along an arbitrary line parallel to $\hat{k}$. These vectors are oriented either 'up' or 'down' — i.e. towards the top or bottom of this page. This 'up'-direction is the wave's *polarisation direction* $\hat{p}$. *From: Frédéric Perez*

cosine argument:

$$
\begin{aligned}
2\pi\frac{(\vec{r}+\vec{r}_\perp)\cdot\hat{k}+ct}{\lambda}+\varphi &= 2\pi\frac{\vec{r}\cdot\hat{k}+\vec{r}_\perp\cdot\hat{k}+ct}{\lambda}+\varphi \\
&= 2\pi\frac{\vec{r}\cdot\hat{k}+0+ct}{\lambda}+\varphi \\
&= 2\pi\frac{\vec{r}\cdot\hat{k}+ct}{\lambda}+\varphi. \qquad (1.13)
\end{aligned}
$$

Hence, $\vec{r}$ and $\vec{r}+\vec{r}_\perp$ have the same electric field displacement. Because $\vec{r}$ and $\vec{r}_\perp$ were arbitrary, *all* points in the same plane perpendicular to $\hat{k}$ have equal E_{real}.

We shall denote the polarisation direction $\hat{p}$, so that $\vec{E}_{\text{real}} = E_{\text{real}}\,\hat{p}$. Because the polarisation direction $\hat{p}$ and the wave propagation direction $\hat{k}$ must be perpendicular, we find that they obey $\hat{p}\cdot\hat{k} = 0$. For a linearly polarised plane wave, $\hat{p}$ is constant in space and time. Under natural conditions, a superposition of two linearly polarised

plane waves of equal wavelength that travel in the same direction ($\lambda_1 = \lambda_2$ and $\hat{k}_1 = \hat{k}_2$) is almost never another linearly polarised plane wave.[18] Instead, an *elliptically* polarised plane wave arises. At any particular instant of time, all points that lie in the *same* plane perpendicular to $\hat{k}$ have the same polarisation direction. However, two *different* planes perpendicular to $\hat{k}$ generally have different polarisation directions. At any particular point, the polarisation direction rotates as time advances; the associated angular velocity is generally not constant in time.[19]

1.6.1 Radiation from all directions

We can think of the night sky as a 'celestial' *sphere,* mathematically represented by the unit 2-sphere $\mathbb{S}^2 := \{\vec{r} \in \mathbb{R}^3 : ||\vec{r}||_2 = 1\}$, with $||\vec{r}||_2$ being the usual Euclidean norm of $\vec{r}$. From each direction, at least *some* radiation from astronomical sources will be travelling towards us. Therefore, to find the total electric field at some position $\vec{r}$ on Earth, we should — thanks to the principle of superposition — *sum up* the contributions from all directions; that, of course, actually means *integration*.

For each point on the sphere, we consider an infinitesimal patch of solid angle $\mathrm{d}\Omega$ around it, and the outwards-directed unit vector $\hat{k}$ perpendicular to the sphere at that point. We assume the sky to emit monochromatically at wavelength λ. The complex electric field at position $\vec{r}$ and time t is

$$E(\vec{r}, t) = \int_{\mathbb{S}^2} A\left(\hat{k}\right) e^{i\left(\frac{2\pi}{\lambda}\left(\hat{k}\cdot\vec{r}+ct\right)+\varphi\left(\hat{k},t\right)\right)} \, \mathrm{d}\Omega. \tag{1.14}$$

Note that we have given φ not only a direction dependence, but a time dependence as well. If $\varphi(\hat{k}, t)$ is fully constant with time, we call the light from direction $\hat{k}$ to be *perfectly coherent*. If $\varphi(\hat{k}, t)$ varies a lot with time, we call the light from direction $\hat{k}$ *incoherent*. See Fig. 1.9.

Astronomers perceive Eq. 1.14's complex electric field — to which, quite astoundingly, sources at gigalightyear distances measurably contribute — through currents that start running in antennae. Here we consider a cylindrical metal rod, oriented along axis $\hat{a}$, as an idealised antenna. If the polarisation direction $\hat{p}$ of an incoming

[18] Here, 'almost never' is used as the opposite of 'almost surely'. A linearly polarised plane wave emerges only in edge cases, such as when the constituent waves are exactly in phase, or when their polarisation directions are either equal or opposite ($\hat{p}_1 = \pm\hat{p}_2$).

[19] This angular velocity *is* constant in time only for the special case of a circularly polarised plane wave, which emerges as the superposition of two linearly polarised plane waves with orthogonal polarisation directions ($\hat{p}_1 \perp \hat{p}_2$), equal amplitudes ($A_1 = A_2$), and a phase difference of 90° ($\varphi_1 - \varphi_2 = \pm\frac{\pi}{2}$). A unique feature of a circularly polarised plane wave is that the *magnitude* of its electric field displacement is constant in space and time ($E_{\text{real}}(\vec{r}, t) = E_{\text{real}}$).

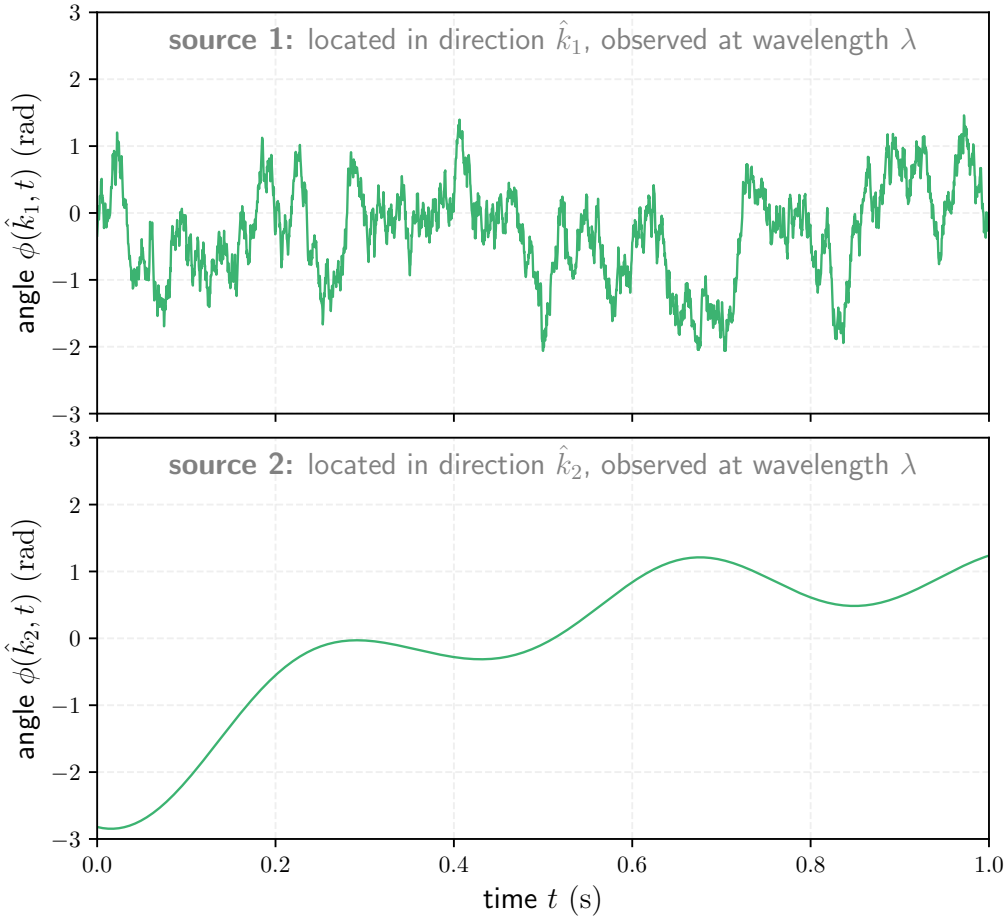


Figure 1.9: For each direction $\hat{k} \in \mathbb{S}^2$, $\varphi(\hat{k}, t)$ is an independent stochastic process with distinct statistical properties. Here we show φ for two directions: $\hat{k}_1$ and $\hat{k}_2$. From both directions, we receive light from a source — an AGN, say. Because the two AGN are far away from eachother, they cannot affect eachother, and as a result the two time series $\varphi(\hat{k}_1, t)$ and $\varphi(\hat{k}_2, t)$ exhibit no cross-correlation. However, there *is* correlation *within* the time series of each direction. The time scale over which φ remains roughly the same, is called the *coherence time* τ. The emission from AGN 1 has a *short* coherence time, whilst the emission from AGN 2 has a *long* coherence time.

plane wave is perpendicular to the cylinder axis $\hat{a}$, then the free electrons in the metal cannot commence an oscillatory motion. Such motion is only possible if the free electrons feel an electric force *along* the direction $\hat{a}$. Thus, the *measurable* fraction of the electric field amplitude $A(\hat{k})$ is reduced by a factor $\hat{a} \cdot \hat{p}(\hat{k}, t)$, where we have made the directional and temporal dependencies of the polarisation direction explicit.

1.6.2 Introducing visibility

Interferometers attempt to measure the time-averaged product of the complex electric field at *some* location with the complex conjugate of this field at *another*. Why is exactly *this* the quantity of interest?

To find out, we first establish an expression for the complex conjugate of the complex electric field. Using rules from complex analysis,[20] we find that complex conjugation of Eq. 1.14 yields

$$E^*\left(\vec{r},t\right) = \int_{\mathbb{S}^2} A\left(\hat{k}\right) e^{-i\left(\frac{2\pi}{\lambda}\left(\hat{k}\cdot\vec{r}+ct\right)+\varphi\left(\hat{k},t\right)\right)}\,\mathrm{d}\Omega. \tag{1.16}$$

Next, consider two points of measurement $\vec{r}_1, \vec{r}_2 \in \mathbb{R}^3$. These points represent the locations of two astronomical dishes or antennae (or, sometimes, 'stations') — on Earth or elsewhere in space. Calling the difference vector $\vec{r}_2 - \vec{r}_1$ the *baseline vector* $\vec{b}$, we can equivalently write these locations as $\vec{r}_1$ and $\vec{r}_2 = \vec{r}_1 + \vec{b}$. Furthermore, let $\langle f(t)\rangle$ denote the time-average of some function $f(t)$:

$$\langle f(t)\rangle := \lim_{\Delta t\to\infty}\frac{1}{\Delta t}\int_{-\frac{\Delta t}{2}}^{\frac{\Delta t}{2}} f(t)\;\mathrm{d}t. \tag{1.17}$$

Now,

$$\begin{aligned}\langle E\left(\vec{r}_1,t\right)E^*\left(\vec{r}_2,t\right)\rangle =& \langle E\left(\vec{r}_1,t\right)E^*\left(\vec{r}_1+\vec{b},t\right)\rangle\\ =& \lim_{\Delta t\to\infty}\frac{1}{\Delta t}\int_{-\frac{\Delta t}{2}}^{\frac{\Delta t}{2}}\int_{\mathbb{S}^2}\int_{\mathbb{S}^2} A\left(\hat{k}\right)A\left(\hat{k}'\right)e^{i\left(\frac{2\pi}{\lambda}\left(\hat{k}\cdot\vec{r}_1+ct\right)+\varphi\left(\hat{k},t\right)\right)}\\ &\cdot e^{-i\left(\frac{2\pi}{\lambda}\left(\hat{k}'\cdot\left(\vec{r}_1+\vec{b}\right)+ct\right)+\varphi\left(\hat{k}',t\right)\right)}\mathrm{d}\Omega\mathrm{d}\Omega'\mathrm{d}t,\end{aligned} \tag{1.18}$$

where we formed a double integral from the product of the two integrals over $\mathbb{S}^2$.

[20] Let $(z_1, z_2, \ldots, z_N) \in \mathbb{C}^N$ be a tuple of N complex numbers. Because $z^* := \mathrm{Re}\,(z) - i\,\mathrm{Im}\,(z)$,

$$\begin{aligned}\left(\sum_{j=1}^N z_j\right)^* &= \left(\sum_{j=1}^N \mathrm{Re}\left(z_j\right) + i\sum_{j=1}^N \mathrm{Im}\left(z_j\right)\right)^*\\ &= \sum_{j=1}^N \mathrm{Re}\left(z_j\right) - i\sum_{j=1}^N \mathrm{Im}\left(z_j\right) = \sum_{j=1}^N z_j^*.\end{aligned} \tag{1.15}$$

Thus, complex conjugation may be taken 'inside' sums. Because the same property holds for integrals, we obtain the required complex conjugation rule.

Next, thanks to cancelling terms, we can simplify the phasor exponent. Finally, we can take the integral over time inside, yielding

$$\begin{aligned}\langle E(\vec{r}_1, t)\, E^*(\vec{r}_2, t)\rangle = \int_{\mathbb{S}^2}\int_{\mathbb{S}^2} A\left(\hat{k}\right) A\left(\hat{k}'\right) e^{i\frac{2\pi}{\lambda}\left(\hat{k}-\hat{k}'\right)\cdot\vec{r}_1} e^{-i\frac{2\pi}{\lambda}\hat{k}'\cdot\vec{b}} \\ \cdot \lim_{\Delta t\to\infty} \frac{1}{\Delta t}\int_{-\frac{\Delta t}{2}}^{\frac{\Delta t}{2}} e^{i\left(\varphi\left(\hat{k},t\right)-\varphi\left(\hat{k}',t\right)\right)}\, \mathrm{d}t\, \mathrm{d}\Omega \mathrm{d}\Omega'. \end{aligned} \tag{1.19}$$

The proof of the van Cittert–Zernike theorem relies on the assumption that the 'phase jitter' time series $\varphi(\hat{k}, t)$ and $\varphi(\hat{k}', t)$ are independent stochastic processes when $\hat{k} \neq \hat{k}'$ (where $\hat{k}, \hat{k}' \in \mathbb{S}^2$). For some fixed sky direction $\hat{k}$, $\{\varphi(\hat{k}, t) \mid t \in \mathbb{R}\}$ is an uncountable set of *dependent* random variables (RVs). To many, this notion may seem rather abstract and daunting. We now provide a concrete example of how the time series $\varphi(\hat{k}, t)$ could be simulated numerically — which requires us to discretise the time axis — and what the corresponding formula for the coherence timescale would be.

Consider an $(N+1)$-tuple of times $(t_0, t_1, t_2, \ldots, t_N)$, with subsequent elements separated by time interval Δt_{step}. We define

$$\varphi(\hat{k}, t_j) := \varphi(\hat{k}, t_{j-1}) + \Delta\varphi_j \tag{1.20}$$

for all $j \in J := \{1, 2, \ldots, N\}$, with the $\{\Delta\varphi_j\}_{j\in J}$ being independent and identically distributed (IID) RVs with $\mathbb{E}[\Delta\varphi_j] = 0$ and $\mathbb{V}[\Delta\varphi_j] =: \sigma^2$. (Note that these assumptions still leave the distribution of $\Delta\varphi_j$ somewhat unconstrained: we have not specified whether the distribution is e.g. normal, uniform, or something else.) Figure 1.10 shows two realisations of phase offset time series simulated in the described way. Next, let $\Delta\varphi$ (*without* subscript) be the total phase change over the course of N time steps:

$$\Delta\varphi := \sum_{j=1}^{N} \Delta\varphi_j. \tag{1.21}$$

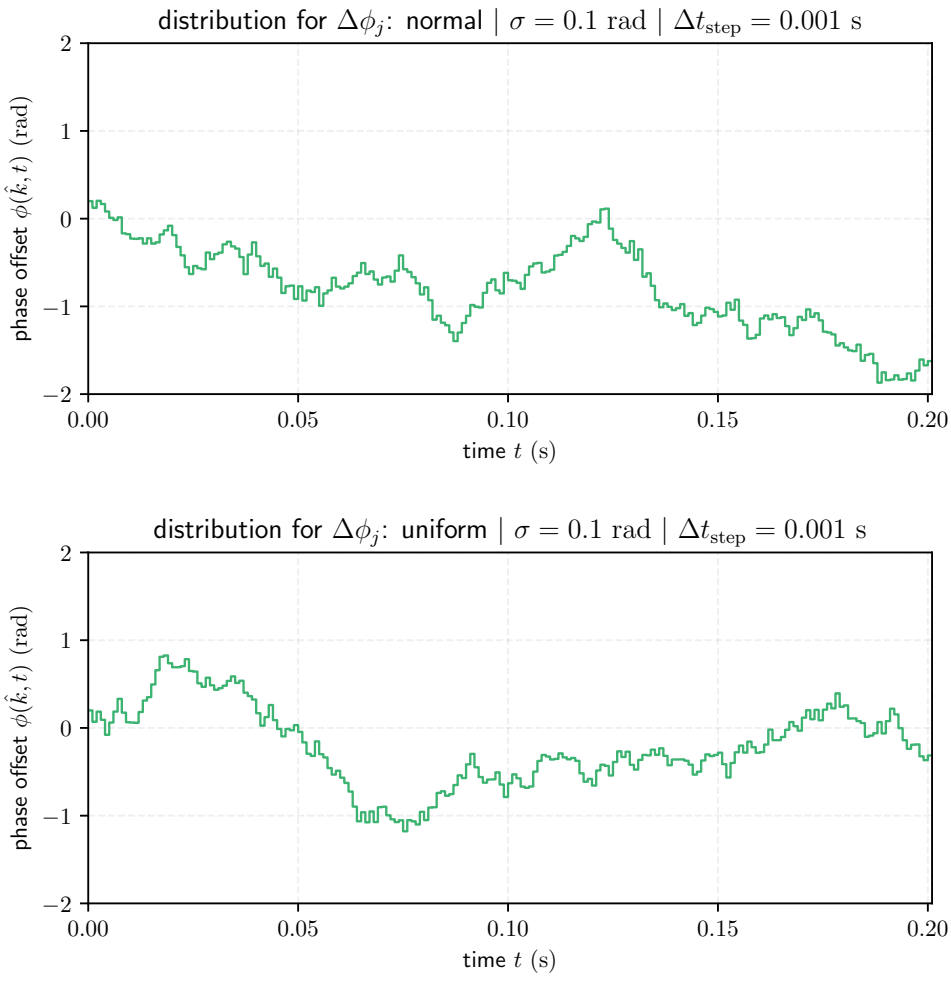


Figure 1.10: Two simulated $\varphi(\hat{k}, t)$ time series, using different distributions for the RVs $\{\Delta\varphi_j\}_{j\in J}$.

What are the mean and variance of this RV?

$$\mathbb{E}\left[\Delta\varphi\right] = \mathbb{E}\left[\sum_{j=1}^{N} \Delta\varphi_j\right] = \sum_{j=1}^{N} \mathbb{E}\left[\Delta\varphi_j\right] = N \cdot 0 = 0; \tag{1.22}$$

$$\mathbb{V}\left[\Delta\varphi\right] = \mathbb{V}\left[\sum_{j=1}^{N} \Delta\varphi_j\right] = \sum_{j=1}^{N} \mathbb{V}\left[\Delta\varphi_j\right] = N \cdot \sigma^2. \tag{1.23}$$

In the second line, to take the sum out of the variance operator, we use that the RVs

are independent.

We define N^* as the numbers of steps necessary for the standard deviation of $\Delta\varphi$ to equal 1 rad. To find an expression for N^* in terms of σ, we remark that the standard deviation of $\Delta\varphi$ equals $\sqrt{\mathbb{V}\left[\Delta\varphi\right]} = \sqrt{\mathbb{V}\left[\Delta\varphi\right](N)}$. Then, by the definition of N^*, we have

$$\sqrt{\mathbb{V}\left[\Delta\varphi\right](N^*)} = 1\,\text{rad}. \tag{1.24}$$

Thus, we find

$$\sqrt{N^*} \cdot \sigma = 1, \;\text{ or }\; N^* = \frac{1}{\sigma^2}. \tag{1.25}$$

It would make sense to define the coherence timescale of this stochastic process as the time passed during N^* steps:

$$\tau := \Delta t_{\text{step}} \cdot N^* = \frac{\Delta t_{\text{step}}}{\sigma^2}. \tag{1.26}$$

After all, this would correspond to the typical amount of time necessary for the phase offset to deviate a significant amount — such as 1 rad. Both simulated time series of Fig. 1.10 have a coherence timescale $\tau = \frac{0.001\text{ s}}{(0.1\text{ rad})^2} = 0.1\,\text{s}$. For the van Cittert–Zernike theorem to work, we much integrate for much longer than the coherence timescale: $\Delta t \gg \tau$. So, say that we are intending to perform interferometric observations, and that the time series shown in Fig. 1.10 are typical realisations for the celestial sky at the observing wavelength. If one had to choose between integration times of $\Delta t = 0.001$ s, $\Delta t = 0.1$ s, or $\Delta t = 10$ s, it would be best to choose $\Delta t = 10$ s. It is not wise to integrate for longer, as this will cause artefacts (*time smearing*) due to the rotation of the Earth, which continuously changes the (u, v)-coordinates of baselines.

Consider two different directions in the sky, $\hat{k}, \hat{k}' \in \mathbb{S}^2$, and their corresponding phase jitter time series, $\varphi(\hat{k}, t)$ and $\varphi(\hat{k}', t)$. In Fig. 1.11, we visualise the complex exponential of their difference,

$$\Upsilon(t) = e^{i\left(\varphi\left(\hat{k},t\right) - \varphi\left(\hat{k}',t\right)\right)}, \tag{1.27}$$

where we have suppressed Υ's dependence on both $\hat{k}$ and $\hat{k}'$ for brevity. This quantity evolves over time in a fickle motion along the Argand diagram's unit circle. The van Cittert–Zernike theorem works because the mean value of the complex numbers that emerge from this random 'dance' around the origin of the Argand diagram is zero. If the integration time Δt is well-chosen, we expect the mean of the complex numbers

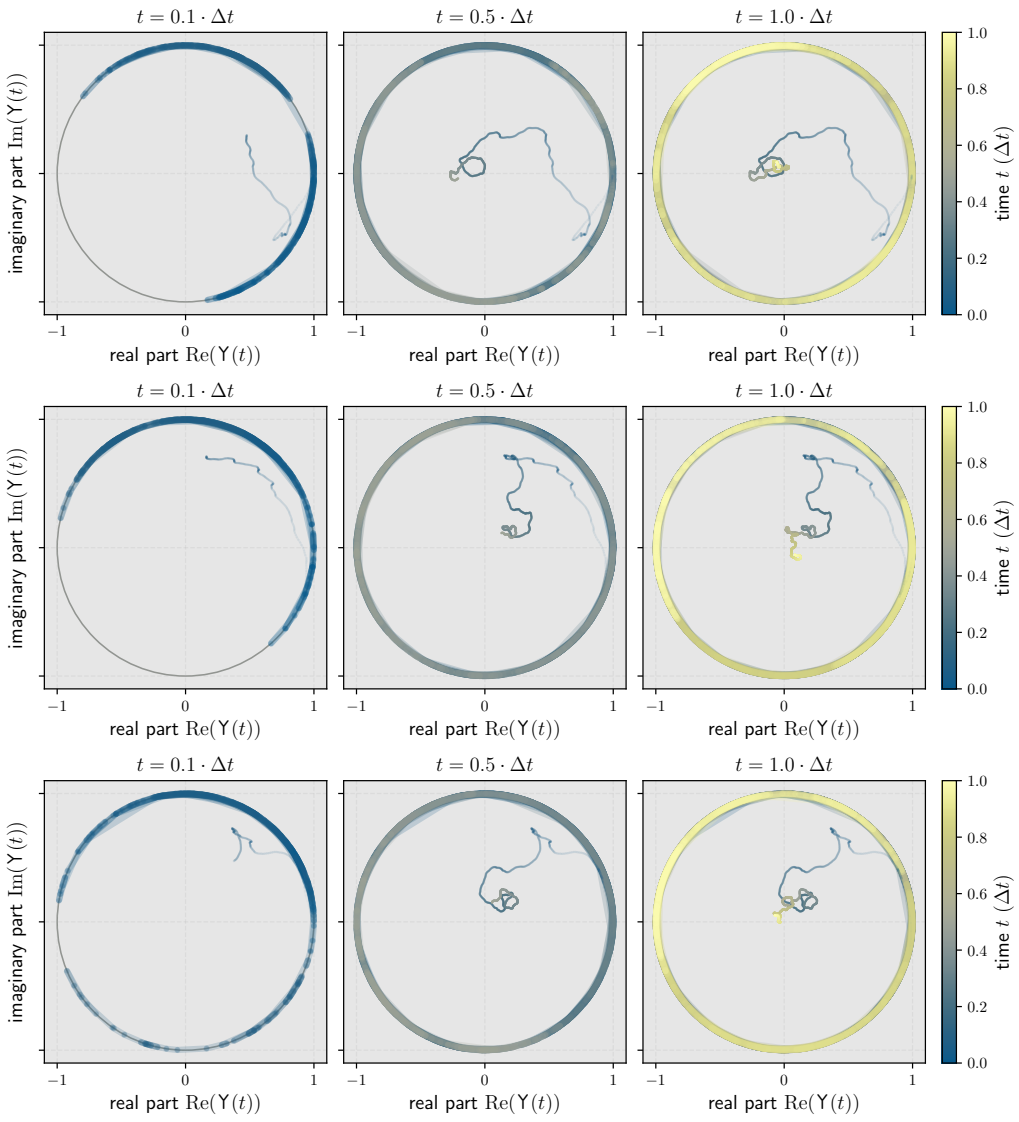


Figure 1.11: Three examples (rows) of Eq. 1.27's $\Upsilon(t)$ (thick circular arcs), the complex exponential of the difference between the phase jitter time series of two distinct sky directions $\hat{k}$ and $\hat{k}'$. Time progresses from left to right, leading up to a full integration time Δt. We also show how $\Upsilon(t)$'s time average evolves (thin meandering curves). The tendency of this time average to approach the origin of the complex plane forms the essence of the van Cittert–Zernike theorem.

to come closer to the origin as the time passed approaches Δt.

Now we use the critical assumption that the stochastic processes $\varphi(\hat{k}, t)$ and $\varphi(\hat{k}', t)$ are independent if $\hat{k} \neq \hat{k}'$. Physically, one says that the sources corresponding to these directions are *mutually incoherent*. Because $\varphi(\hat{k}, t) - \varphi(\hat{k}', t)$ thus is a random angle,

$\Upsilon(t)$ will be a random point on the unit circle in the complex plane. Taking the time average of many such random points on the unit circle will eventually ($\Delta t \to \infty$) make the result vanish:

$$\langle e^{i\varphi(\hat{k},t)} \rangle = \mathbb{E}\left[e^{i\varphi(\hat{k},t)}\right] = 0. \tag{1.28}$$

As a result,

$$\begin{aligned}
\lim_{\Delta t \to \infty} \frac{1}{\Delta t} \int_{-\frac{\Delta t}{2}}^{\frac{\Delta t}{2}} \Upsilon(t)\, \mathrm{d}t &= \mathbb{E}\left[e^{i\varphi\left(\hat{k},t\right)} \left(e^{i\varphi\left(\hat{k}',t\right)}\right)^*\right] \\
&= \mathbb{E}\left[\left(e^{i\varphi\left(\hat{k},t\right)} - \mathbb{E}\left[e^{i\varphi\left(\hat{k},t\right)}\right]\right)\left(e^{i\varphi\left(\hat{k}',t\right)} - \mathbb{E}\left[e^{i\varphi\left(\hat{k}',t\right)}\right]\right)^*\right] \\
&=: \mathrm{Cov}\left[e^{i\varphi\left(\hat{k},t\right)}, e^{i\varphi\left(\hat{k}',t\right)}\right] \\
&= \delta(\hat{k} - \hat{k}'),
\end{aligned} \tag{1.29}$$

where $\delta(\hat{k} - \hat{k}')$ is the Dirac delta distribution. Looking back at Eq. 1.19, we see that we have now calculated its time integral, which we can replace by this Dirac delta distribution:

$$\begin{aligned}
\langle E(\vec{r}_1, t)\, E^*(\vec{r}_2, t) \rangle = \int_{\mathbb{S}^2} \int_{\mathbb{S}^2} A\left(\hat{k}\right) A\left(\hat{k}'\right) \\
\cdot\, e^{i\frac{2\pi}{\lambda}\left(\hat{k}-\hat{k}'\right)\cdot\vec{r}_1}\, e^{-i\frac{2\pi}{\lambda}\hat{k}'\cdot\vec{b}}\, \delta(\hat{k} - \hat{k}')\, \mathrm{d}\Omega \mathrm{d}\Omega'.
\end{aligned} \tag{1.30}$$

In practice, we cannot integrate for an infinite amount of time to obtain just *one* visibility: we would never get anywhere with interferometry in the Universe's lifetime! Instead, we choose a finite integration time. This integration time Δt should be *longer* than the coherence time of celestial sources. In this way, we make sure that the time integral still roughly vanishes for different directions $\hat{k}$ and $\hat{k}'$.[21]

Next, we evaluate one of the integrals over the 2-sphere $\mathbb{S}^2$. For example, one could evaluate the integral with infinitesimal solid angle $\mathrm{d}\Omega'$ and direction vector $\hat{k}'$. The integral vanishes for all values of $\hat{k}'$, except for when $\hat{k}' = \hat{k}$. Note that precisely when that happens, $A(\hat{k}') = A(\hat{k})$, so that we obtain a factor $A^2(\hat{k})$. Thus,

$$\langle E(\vec{r}_1, t)\, E^*(\vec{r}_2, t) \rangle = \int_{\mathbb{S}^2} A^2\left(\hat{k}\right) e^{-i\frac{2\pi}{\lambda}\hat{k}\cdot\vec{b}}\, \mathrm{d}\Omega. \tag{1.31}$$

[21] Usually, the coherence time of astronomical sources is much less than a second. With the LOFAR, integration times of roughly a second therefore suffice.

Note that the result depends on the baseline vector $\vec{b}$, but is independent of $\vec{r}_1$: the absolute locations of the antennae do not matter! Now treated as a *function* of $\vec{b}$, this expression is called the *visibility function* at wavelength λ:

$$V_\lambda(\vec{b}) := \int_{\mathbb{S}^2} A^2\left(\hat{k}\right) e^{-i\frac{2\pi}{\lambda}\hat{k}\cdot\vec{b}}\,\mathrm{d}\Omega. \tag{1.32}$$

The astute reader may see the contours of a Fourier transform appearing.

1.6.3 Visibility and specific intensity

Let us say we are interested in reconstructing the sky's specific intensity function at wavelength λ around some central direction $\hat{k}_\mathrm{c} \in \mathbb{S}^2$. Without loss of generality, we now choose a Cartesian coordinate system in which $\hat{k}_\mathrm{c}$ is the positive z-axis. Any direction vector $\hat{k}_+$ in the same hemisphere as $\hat{k}_\mathrm{c}$ can be written $\hat{k}_+ = [k_x, k_y, \sqrt{1-k_x^2-k_y^2}]$. Likewise, any direction vector $\hat{k}_-$ in the *opposing* hemisphere can be written $\hat{k}_- = [k_x, k_y, -\sqrt{1-k_x^2-k_y^2}]$. We can therefore write — without doing any approximations — that

$$\begin{aligned} V_\lambda(\vec{b}) = \int_{\mathbb{R}^2} \mathbb{I}\left(k_x^2+k_y^2 \le 1\right) \Big(A^2\left(\hat{k}_+\right) e^{-i\frac{2\pi}{\lambda}\hat{k}_+\cdot\vec{b}} + \\ A^2\left(\hat{k}_-\right) e^{-i\frac{2\pi}{\lambda}\hat{k}_-\cdot\vec{b}}\Big)\,\mathrm{d}k_x\mathrm{d}k_y. \end{aligned} \tag{1.33}$$

Here, $\mathbb{I}$ is the *indicator function*, which equals 1 if the condition in parentheses is true, and 0 if this condition is false. Typically, $A(\hat{k})$ falls off quickly away from $\hat{k}_\mathrm{c}$ because telescopes have a limited field of view. Therefore, it is *very* safe to assume that all directions *in the hemisphere opposing* $\hat{k}_\mathrm{c}$ have a vanishing contribution to the electric field as measured by the stations: $A(\hat{k}_-) = 0$. We obtain

$$V_\lambda(\vec{b}) = \int_{\mathbb{R}^2} \mathbb{I}\left(k_x^2+k_y^2 \le 1\right) A^2\left(\hat{k}_+\right) e^{-i\frac{2\pi}{\lambda}\left(b_x k_x + b_y k_y + b_z\sqrt{1-k_x^2-k_y^2}\right)}\,\mathrm{d}k_x\mathrm{d}k_y. \tag{1.34}$$

Let us interpret the product of the first two factors in the integral: this represents the *observed* (rather than *ground-truth*) specific intensity at wavelength λ. (One reason is that A in direction $\hat{k}$ at time t is decreased by a factor that depends on the orientation of our dipole $\hat{a}$, and the polarisation direction $\hat{p}(\hat{k}, t)$.) *In concreto,* $I_{\lambda,\mathrm{obs}} : \mathbb{R}^2 \to \mathbb{R}_{\ge 0}$

is

$$I_{\lambda,\text{obs}}\left(k_x, k_y\right) := \begin{cases} A^2\left(\hat{k}_+\right) & \text{if } k_x^2 + k_y^2 \leq 1; \\ 0 & \text{otherwise.} \end{cases} \tag{1.35}$$

Let us now consider an interesting limiting case: stations with a small field of view. Quantitatively, this means that $I_{\lambda,\text{obs}}\left(k_x, k_y\right)$ decays quickly as $k_x^2 + k_y^2$ increases, so that the only relevant contributions to the integral are for $k_x^2 + k_y^2 \ll 1$. Clearly, if $k_x^2 + k_y^2 \ll 1$, then $\sqrt{1 - k_x^2 - k_y^2} \approx 1$. This means that we can take a phasor $e^{-i\frac{2\pi}{\lambda}b_z}$ out of the integral; after all, this factor does not depend on integration variables k_x and k_y anymore. As a result,

$$V_\lambda(\vec{b}) \approx e^{-i\frac{2\pi}{\lambda}b_z} \int_{\mathbb{R}^2} I_{\lambda,\text{obs}}\left(k_x, k_y\right) e^{-i\frac{2\pi}{\lambda}\left(b_x k_x + b_y k_y\right)} \, \mathrm{d}k_x \mathrm{d}k_y. \tag{1.36}$$

For baseline vectors $\vec{b}$ in the plane $b_z = 0$ (so that b_x and b_y are the only two remaining coordinates), the complex exponential in front of the integral reduces to 1. Let us define the dimensionless coordinates $u_x := b_x\lambda^{-1}$, $u_y := b_y\lambda^{-1}$ and $u_z := b_z\lambda^{-1}$ $(= 0)$. We also introduce new notation for the visibility function at wavelength λ restricted to the plane $b_z = 0$. *In concreto,* $V_\lambda^0 : \mathbb{R}^2 \to \mathbb{C}$ is

$$V_\lambda^0(u_x, u_y) := V_\lambda\left([b_x, b_y, 0]^T\right). \tag{1.37}$$

Thus,

$$V_\lambda^0(u_x, u_y) = \int_{\mathbb{R}^2} I_{\lambda,\text{obs}}\left(k_x, k_y\right) e^{-2\pi i\left(u_x k_x + u_y k_y\right)} \, \mathrm{d}k_x \mathrm{d}k_y. \tag{1.38}$$

To cast this result in the most common notation, we must relabel: $u_x \to u$, $u_y \to v$, $u_z \to w$, $k_x \to l$, $k_y \to m$, $[u_x, u_y]^T \to \vec{u}$ and $[k_x, k_y]^T \to \vec{l}$. Then

$$V_\lambda^0\left(\vec{u}\right) = \int_{\mathbb{R}^2} I_{\lambda,\text{obs}}(\vec{l}) \, e^{-2\pi i \vec{u}\cdot\vec{l}} \, \mathrm{d}\vec{l}, \tag{1.39}$$

which says that $V_\lambda^0(\vec{u})$ is the Fourier transform of $I_{\lambda,\text{obs}}(\vec{l})$. This result is the famous van Cittert–Zernike theorem, established by Dutch physicists Pieter Hendrik van Cittert and Frits Zernike. Van Cittert was the first to derive the result (van Cittert, 1934), whilst Zernike found a simpler proof (Zernike, 1938). Equation 1.39 represents one of the most important equations in astronomy, elucidating why telescopes

are Fourier transform machines.[22]

If we measure V_λ for a set of baseline vectors $\{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_N\}$, we can trivially find V^0_λ for these vectors:

$$V^0_\lambda(\vec{u}_i) = e^{2\pi i w_i} V_\lambda(\vec{b}_i). \tag{1.40}$$

(Here, $w_i = u_{z,i} = b_{z,i}\lambda^{-1}$.) With some measured values $V^0_\lambda(\vec{u}_i)$ in hand, we can make an estimate of the full function $V^0_\lambda(\vec{u})$ via an interpolation method of choice. Next, we simply use the inverse Fourier transform to estimate $I_{\lambda,\mathrm{obs}}(\vec{l})$, the sky's specific intensity function at wavelength λ.[23]

1.7 The ionosphere

Finally, a major antagonist in our quest to measure signals from cosmic magnetism through radio interferometry is the Earth's own ionosphere. Here we briefly introduce the theory that underpins the probabilistic calibration of ionospheric distortions as presented in this thesis.

In the 1920s and 1930s, Hans Lassen, Edward Appleton, Sydney Goldstein, and Douglas Hartree (in that order) each independently published versions of the theory of EM wave propagation through a plasma of ions and electrons (e.g. Anduaga, 2021).[24] These efforts led to what is now known as the Appleton–Lassen equation,

[22] The astute reader may remark that in this derivation, we have assumed that the amplitude A is different for EM waves of different directions, but the same for all measurement locations $\vec{r}$ on Earth and in its vicinity. In reality, the flux from a point source obeys the inverse square law, and A is inversely proportional to distance. Fortunately, the assumption of a position-independent amplitude function $A(\vec{r}, \hat{k}) \approx A(\hat{k})$ is accurate as long as the sources are astronomically far away. To see why, assume that, at $\vec{r}_1$, some source is a distance R away. Then, at $\vec{r}_2 = \vec{r}_1 + \vec{b}$, the same source is at most a distance $R + b$ away, where $b = ||\vec{b}||_2$. The relative amplitude at $\vec{r}_2$ with respect to $\vec{r}_1$ is $\frac{R}{R+b} = \frac{1}{1+\frac{b}{R}}$. The longest baseline ever used in astronomy is of the order of one Earth diameter; moreover, choose a relatively nearby astronomical object at 1 lightyear. Then $\frac{1}{1+\frac{b}{R}} = 0.999999999$. For astronomical objects that are further, the approximation of an position-independent amplitude function is even better. For VLBI observations of Solar System planet Mars, using the same longest baseline, we have $\frac{1}{1+\frac{b}{R}} = 0.99995$.

[23] One popular approach to interpolating the visibility function and estimating the sky's specific intensity function, is CLEAN (e.g. Högbom, 1974).

[24] As the historical reconstruction of Gillmor (1982) details, Appleton, who won the 1947 Nobel Prize in Physics for his *'investigations of the physics of the upper atmosphere'*, received substantial theoretical help from the young Austrian Wilhelm Altar in the years 1925–1926. Strikingly, during this time, Altar wrote a draft manuscript *'Wellenausbreitung in ionisierten Gasen unter dem Einfluss eines Magnetfelds'*, in which he derived the Appleton–Lassen equation *years before* Appleton would publish it — without mentioning Altar.

which can quantify radio wave phase changes due to refraction in atmospheric media. From here onwards, for historical accuracy, I shall call this equation the *Altar–Appleton–Lassen equation* instead. The full Altar–Appleton–Lassen equation gives the (possibly complex) indices of refraction n_- and n_+ at some point in spacetime $x^\mu = (ct, \vec{r})$ as

$$n_\pm^2(x^\mu) = 1 - \frac{X}{1 - iZ - \frac{Y^2 \sin^2\theta}{2(1-X-iZ)} \pm \sqrt{\frac{Y^4 \sin^4\theta}{4(1-X-iZ)^2} + Y^2 \cos^2\theta}}, \tag{1.41}$$

where $X = X(x^\mu) := \left(\frac{\nu_\mathrm{p}}{\nu}\right)^2$, $Y = Y(x^\mu) := \frac{\nu_\mathrm{g}}{\nu}$, and $Z = Z(x^\mu) := \frac{1}{2\pi}\frac{\nu_\mathrm{c}}{\nu}$ (e.g. Altman & Suchy, 2011). ν_p is the plasma frequency[25], ν_g is the electron gyro frequency[26], and ν_c is the electron collision frequency, whilst θ is the angle between the local magnetic field and the wave vector. Note that each term that depends on θ also depends on Y, and therefore on B: $Y \propto \nu_\mathrm{g}$ and $\nu_\mathrm{g} \propto B$. Therefore, all terms containing θ vanish when $B = 0$. This *should* be the case, as θ becomes a meaningless quantity in the absence of magnetic fields.

To obtain a rough estimate of the severity of phase distortions generated in the ionosphere, we idealise the layer's plasma as cold, and as unaffected by external magnetic fields $\vec{B}$.[27] Thus, we set $Y = 0$ and $Z = 0$. The ordinary and extraordinary modes now coalesce (so we use n instead of $n_\pm$), and the Altar–Appleton–Lassen equation reduces to the form

$$n_\pm(x^\mu) \approx n_\pm(X) = \sqrt{1 - X} = \sqrt{1 - \left(\frac{\nu_\mathrm{p}}{\nu}\right)^2} = n\left(\frac{\nu_\mathrm{p}}{\nu}\right) = n(x^\mu). \tag{1.42}$$

In this cold-plasma limit, where electronic thermal motion is ignored, ν_p depends on the ionosphere's free electron density n_e as

$$\nu_\mathrm{p}(n_\mathrm{e}) = \frac{e}{2\pi\sqrt{\varepsilon_0 m_\mathrm{e}^*}} \cdot \sqrt{n_\mathrm{e}}. \tag{1.43}$$

[25] The plasma frequency is a cut-off frequency below which electromagnetic wave propagation does not occur. EM waves with $\nu < \nu_\mathrm{p}$ are reflected by the plasma.

[26] The gyro frequency of a particle of charge q and mass m in a locally uniform magnetic field of strength B is $\nu_\mathrm{g} = \frac{|q|B}{2\pi\ \mathrm{rad}\cdot m}$.

[27] If electronic thermal motion is negligible, the collision rate $\nu_\mathrm{c} \ll \nu$. In such cases, the plasma is said to be *collisionless* (at least on timescales $\sim\nu^{-1}$). In practice, taking the plasma temperature to be close to 0 might not always yield valid results. The approximation under consideration also neglects the influence of Earth's magnetic field on radio wave propagation. Numerical calculations may adopt more sophisticated forms of the Altar–Appleton–Lassen equation that correspond to ionospheres in which the electrons are not necessarily cold and a magnetic field is present.

Here, e is the elementary charge, ε_0 is the vacuum permittivity, and m_e^* is the effective electron mass. The free electron density is not homogeneous throughout the plasma — rather, $n_e = n_e(x^\mu)$. As a result, the plasma frequency and so also the refractive index will vary along the optical path (OP). This explains why the total phase change induced by refraction during a radio wave's journey through the ionosphere φ_i is found by summing up many small contributions along the OP:

$$\varphi_\frown = \varphi_\frown(\nu) = 2\pi \text{ rad} \cdot \frac{\int_{\text{OP}}(1-n)\mathrm{d}l}{\lambda} = 2\pi \text{ rad} \cdot \frac{\nu}{c}\int_{\text{OP}}(1-n)\mathrm{d}l, \tag{1.44}$$

where c is the speed of light in vacuum and $\varphi_\frown$ is assumed to be measured in radians. It is readily seen that in the vacuum limit ($n \to 1$), the phase delay vanishes.

Together, Eqs. 1.42, 1.43, and 1.44 provide a recipe to calculate the ionospheric phase disturbance for a radio wave of an arbitrary frequency along an arbitrary line of sight, given a known electron density distribution $n_e(\vec{r}, t)$.[28]

As an illustration, let us consider the limit in which the frequency of the radio wave is consistently much higher than the plasma frequency (*id est* $\nu \gg \nu_p(\vec{r}, t)$, or equivalently, $\frac{\nu_p(\vec{r},t)}{\nu} \ll 1$, at all spacetime loci of the wave's trajectory).[29] In this case, we can approximate $n = n\left(\frac{\nu_p}{\nu}\right)$ with a second-degree Taylor polynomial around $\frac{\nu_p}{\nu} = 0$. Substituting that result in Eq. 1.44 yields

$$\varphi_\frown \approx \frac{\pi}{c}\int_{\text{OP}} \nu_p^2 \,\mathrm{d}l \cdot \frac{1}{\nu} = \frac{e^2}{4\pi c \varepsilon_0 m_e^*}\int_{\text{OP}} n_e \,\mathrm{d}l \cdot \frac{1}{\nu} = \frac{\beth}{\nu}. \tag{1.46}$$

In this limit, it is apparent that $\varphi_\frown \propto \nu^{-1}$. We thus see that the frequency-dependency of ionospheric phase distortions over the observed radio bandwidth should not be ignored, given a large enough proportionality factor $\beth$. Also, the $\varphi_\frown \propto \nu^{-1}$ behaviour makes clear that the disturbances are more severe at low frequencies than at higher frequencies. In particular, observations at 30 MHz are plagued by ionospheric phase effects that are 10 times more severe than those at 300 MHz, and even 100 times more

[28] More explicitly,

$$\varphi_\frown(\nu, \text{OP}, n_e) = 2\pi \text{ rad} \cdot \frac{\nu}{c}\int_{\text{OP}} \left(1 - \sqrt{1 - \frac{e^2}{4\pi^2 \varepsilon_0 m_e^*}\frac{n_e}{\nu^2}}\right)\mathrm{d}l. \tag{1.45}$$

[29] This assumption is reasonable in the description of Earth's ionosphere for the purposes of low-frequency radio observations (understood to be $\nu \sim 10^2$ MHz). Considering that the average free electron density is $\sim 10^{10}$ m^{-3}, Eq. 1.43 yields $\nu_p \sim 1$ MHz, so that indeed $\frac{\nu_p}{\nu} \sim 10^{-2} \ll 1$.

severe than those at 3 GHz.[30]

1.8 Late-time radio probes of cosmological magnetism

In the preceding sections, we have reviewed the presence of magnetic fields at planetary, stellar, and cosmological scales. We have highlighted their role in the shielding of planetary atmospheres against stellar winds, through which they plausibly contribute to maintaining habitable conditions, and in the formation of stars. We have also seen that the ultimate origin of magnetism is hitherto uncertain. In the last few sections, we have argued that interferometric observations in the low-frequency radio window appear particularly promising to study energetic magnetic phenomena in the Universe — provided that ionospheric distortions can be effectively calibrated out.

In this section, I introduce the radio probes of cosmological magnetism at late times that we have studied in this thesis.

1.8.1 Structure formation shocks in the Cosmic Web

Large-scale structure formation is a process that started in the Early Universe and continues to this day. In the prevailing *cold* dark matter cosmological model, small structures collapse first and superclusters collapse last.[31] At present, massive galaxy clusters are arising, which grow in mass by merging with other clusters, by the episodic infall of galaxy groups and individual galaxies from connecting Cosmic Web filaments, and by continuous dark matter and IGM accretion streams — again from filaments. With cluster growth being fuelled by filaments, the filament population must be evolving at present, too. Although galaxy clusters are much easier to observe than filaments throughout the electromagnetic spectrum, they may not be the most promising theatres to study cosmic magnetogenesis, as their eventful histories make it hard to tie a detected magnetic field to a specific physical origin. By contrast, filaments have had more quiet pasts, and their magnetic fields therefore possibly resemble *cosmic fossils* of primordial magnetogenesis: i.e. witnesses of a bygone age that have been largely unaffected by the passing of time.

[30] At multi-metre and decametre wavelengths, ionospheric effects with even stronger inverse frequency dependencies (i.e. at least $\propto \nu^{-2}$; see e.g. de Gasperin et al., 2018) become important.

[31] In (disfavoured) scenarios in which *hot* dark matter dominates the composition of dark matter, this order is reversed: hot dark matter particles — such as massive neutrinos (e.g. Gershtein & Zel'dovich, 1966) — move so fast that they can escape from small structures, dissolving them in the process. The remaining large structures are only weakly overdense, and thus collapse slowly. Galaxies eventually form by fragmentation of collapsed galaxy clusters.

One promising way to detect magnetic fields in Cosmic Web filaments is to search for synchrotron radiation generated in the downstream regions of structure formation shocks. The emergence of such shock waves is a generic prediction from theory and simulations of cosmological structure formation. To appreciate the basic properties of structure formation shocks in the Cosmic Web, we consider a simple analytic model.

Gauss's law in integral form for Newtonian gravity is

$$\oiint \vec{g} \cdot \hat{n} \, \mathrm{d}S = 4\pi G M_{\mathrm{enc}}, \tag{1.47}$$

where $\vec{g}$ denotes the gravitational field and $\hat{n}$ is an outward-oriented unit vector normal to the surface enclosing a volume with total mass M_{enc}. Let us now consider a Cosmic Web filament, idealised as an infinitely long cylinder of proper radius R and proper density ρ. The gravitational field for this configuration will depend only on the proper distance r to the filament spine: $\vec{g} = g(r)\hat{r}$, where $\hat{r}$ is the unit vector field pointing away radially from the filament spine. To invoke Gauss's law, we imagine a cylinder of proper length L and proper radius r concentric with the filament. For $r \geq R$, Eq. 1.47 then becomes

$$g(r) \cdot 2\pi r L = 4\pi G \cdot \pi R^2 L \rho, \tag{1.48}$$

so that

$$\vec{g}(\vec{r}) = -\frac{2G\,\lambda_{\mathrm{m}}}{r}\,\hat{r}, \tag{1.49}$$

where $\lambda_{\mathrm{m}} := \pi R^2 \rho$ is the filament's proper *linear mass density*: its mass per unit of proper length. Let us now consider, within an expanding universe, the free-fall dynamics of a gas pocket that starts out in a void and that is gravitationally attracted to the filament. After how much time, and at what speed relative to the filament, will the gas pocket crash into it? We shall consider the proper speed v of the gas pocket — at every instant measured relative to the local CMB rest frame.[32] It is therefore apt to interpret v as a *peculiar* speed. If the initial peculiar velocity vector of the pocket lies in the plane containing both the filament spine and the pocket, the problem becomes

[32] In an expanding universe, only the notion of the speed of one object relative to another at the *same* spacetime location carries meaning. Mathematically speaking, distinct points of non-Euclidean manifolds have distinct tangent spaces, and the difference between vectors from distinct vector spaces is undefined. (For Euclidean manifolds, the tangent spaces of distinct points are also distinct *in principle*, but they can all be identified with one another.)

one-dimensional.[33] In this case, we only need to track the pocket's proper coordinate x. Placing the spatial origin at the filament spine, we find from Eq. 1.49 for $|x| \geq R$ that

$$g(x) = -\frac{2G\,\lambda_{\mathrm{m}}}{x}. \tag{1.50}$$

In a simple cosmological picture with a comoving galaxy cluster number density that remains the same over time, the *comoving* inter-cluster distances remain roughly constant, while the *proper* inter-cluster distances are inversely proportional to $1 + z$. (Here, z denotes cosmological redshift.) A typical filament, spanned between two neighbouring non-merging clusters, is thus stretched as this universe expands. If the matter streams entering the filament balance out the matter streams leaving it, the filament's mass is conserved — or, equivalently, its *comoving* linear mass density is constant: $\lambda_{\mathrm{m,c}}(z) = \lambda_{\mathrm{m}}(0)$. As a first approximation, it is therefore reasonable to assume that the *proper* linear mass density evolves as $\lambda_{\mathrm{m}}(z) = (1+z)\lambda_{\mathrm{m}}(0)$. Under this assumption, g is not only a function of x but also of t, as $z = z(t)$.

In a static (rather than expanding) universe, within an infinitesimal time $\mathrm{d}t$, the peculiar speed would change simply by $\mathrm{d}v = g\,\mathrm{d}t$. By contrast, in an expanding universe, peculiar speeds die out in the absence of gravity or forces.[34] To see why, we consider a family of observers along the pocket's trajectory, each with vanishing peculiar motion. Within $\mathrm{d}t$, the pocket will travel between family members a proper distance $\mathrm{d}x = v\,\mathrm{d}t$ apart. However, by the expansion of the universe, these relatives move away from eachother with proper speed $H\,\mathrm{d}x$, where H is the Hubble parameter. Once it arrives, the second family member will measure the pocket's peculiar speed to be $v - H\,\mathrm{d}x$, as described by a simple Galilean transformation. As a result, the pocket's peculiar speed evolves as

$$\frac{\mathrm{d}v}{\mathrm{d}t} = g(x(t), t) - H(t)v(t). \tag{1.51}$$

The proper position of the pocket (and any other object) evolves as a result of both peculiar motion and Hubble expansion, leading to

$$\frac{\mathrm{d}x}{\mathrm{d}t} = v(t) + H(t)x(t). \tag{1.52}$$

[33] The astute reader will remark that the formulation of this sentence's conditional implicitly assumes a *Euclidean* geometry. This assumption is not material to its conclusion.

[34] This fact is the key to understanding why matter cools as universes expand.

These coupled differential equations can be solved as a function of time t. Alternatively, we can recast them using z, using the facts that the scale factor $a = \frac{1}{1+z}$ and $H := \frac{\dot{a}}{a}$, where $\dot{a} := \frac{\mathrm{d}a}{\mathrm{d}t}$. The chain rule implies that the differentials relate as

$$\mathrm{d}z = -H(z)(1+z)\,\mathrm{d}t. \tag{1.53}$$

Furthermore, it is convenient to use the pocket's comoving position $x_\mathrm{c}(z) = x(z)(1+z)$ rather than its proper position. Starting off from Eqs. 1.51 and 1.52, these changes of variable yield

$$\frac{\mathrm{d}v}{\mathrm{d}z} = \frac{2G(1+z)\lambda_\mathrm{m}(0)}{H(z)\,x_\mathrm{c}(z)} + \frac{v(z)}{1+z}; \tag{1.54}$$

$$\frac{\mathrm{d}x_\mathrm{c}}{\mathrm{d}z} = -\frac{v(z)}{H(z)}. \tag{1.55}$$

Adopting $H(z)$ from the concordance ΛCDM cosmology, we solve these differential equations numerically for $v(z)$ and $x_\mathrm{c}(z)$. Of course, we need to assume a value for $\lambda_\mathrm{m}(0)$ and choose the initial conditions $v(z_\mathrm{i})$ and $x_\mathrm{c}(z_\mathrm{i})$ at some initial redshift z_i. By definition of the *overdensity* δ, one can dissect $\rho = \rho_\mathrm{c}(0)\Omega_\mathrm{M}(0)(1+\delta(0))$, where $\rho_\mathrm{c}(0)$ is the critical density at $z = 0$, $\Omega_\mathrm{M}(0)$ is the matter density parameter at $z = 0$, and $\delta(0)$ is the overdensity at $z = 0$. We set $R = 1$ Mpc and vary $\delta(0)$: $\delta(0) \in \{10, 15, 20, 25, 30\}$. As initial conditions, we choose $z_\mathrm{i} = 6$, $v(z_\mathrm{i}) = 0$, and $x_\mathrm{c}(z_\mathrm{i}) = 5$ Mpc. As such, the initial conditions define the pocket's state when the Universe was about a billion years old, and the Epoch of Reionisation came to an end. We simulate the pocket's dynamics until it crashes onto the filament — that is to say, until its proper distance to the filament spine is less than R — or, alternatively, until $z = 0$.

In Fig. 1.12, we show the corresponding numerical results. The gas pockets infalling on Cosmic Web filaments with current-day overdensities $\delta(0) \in \{20, 25, 30\}$ collide before $z = 0$. These violent events do not occur simultaneously, but are temporally scattered across billions of years of cosmic history, and make an end to journeys that can last more than ten billion years. The more massive the filament, the earlier the collision takes place. The top panel shows that, when this happens, the peculiar speeds $v \sim 10^2$ km s^{-1}. As the sound speed in the warm–hot intergalactic medium (warm–hot IGM, or WHIM) $c_\mathrm{s} \sim 10^1$ km s^{-1} (e.g. Ryu et al., 2003), we draw the key conclusion that these infalling gas pockets are *supersonic*, and thus generate shock waves.[35] The shocks of Fig. 1.12 have Mach numbers $\mathcal{M} := \frac{v}{c_\mathrm{s}} \sim 10$. Full-

[35] A particularly easy way to remember the correct order of magnitude for the speed of sound in the

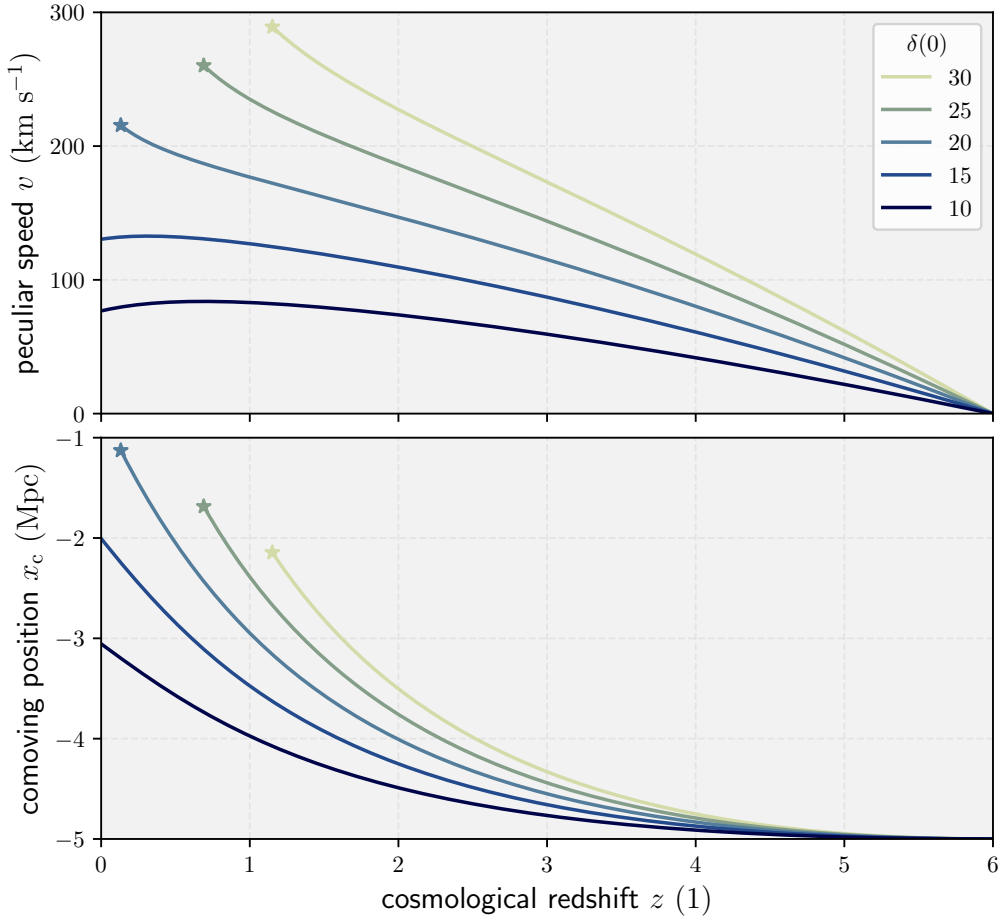


Figure 1.12: Dynamics of void gas pockets falling towards Cosmic Web filaments through cosmic time, simulated from $t = 1$ Ga onwards. Time proceeds from right to left. Over the course of billions of years, such gas pockets can traverse megaparsec-scale distances, and crash onto filaments with peculiar speeds $v \sim 10^2$ km s^{-1}. The accreting matter is supersonic, because sound speeds in filaments are $c_s \sim 10^1$ km s^{-1}. From dark to light, the colours indicate filaments with current-day overdensities $\delta(0) \in \{10, 15, 20, 25, 30\}$. The stars mark pockets that have crashed onto filaments before the present day.

fledged cosmological simulations yield probability distributions over Mach numbers (e.g. Fig. 17 of Vazza et al., 2011), showing $\mathcal{M} \sim 10^0$–10^2 to be the typical range;

WHIM is as follows. A well-known fact from secondary school is that the speed of sound in Earth's atmosphere $c_s \sim 10^{-1}$ km s^{-1}. In an ideal gas, $c_s = \sqrt{\gamma\, m^{-1}\, k_B T}$, where γ is the adiabatic index, m is the particle mass, and k_B is Boltzmann's constant. As $\sqrt{\gamma\, m^{-1}}$ varies between Earth's atmosphere and the WHIM by a factor of order unity at most, sound speed variations are almost exclusively due to temperature variations: $c_s \propto \sqrt{T}$. Using this proportionality along with the facts that $T \sim 10^2$ K for Earth's atmosphere and $T \sim 10^6$ K for the WHIM, we find $c_s \sim 10^1$ km s^{-1} for the WHIM.

however, shocks with $\mathcal{M} \sim 10^4$ do sometimes occur.

The seminal work of Ryu et al. (2008) predicts that these structure formation shocks cause vortical motions in the IGM, which then cause the IGM to develop turbulence. The so-called *turbulence dynamo* could then amplify weak seed magnetic fields to $B \sim 1$ μG in clusters and $B \sim 10$ nG in filaments. However, structure formation shocks do not only presumably amplify the magnetic fields of the Cosmic Web, but also serve as radio beacons for their detection. This is because the downstream regions of structure formation shocks glow in synchrotron light, as the shocks accelerate the IGM's high-energy electrons to ultrarelativistic velocities. Shocks accomplish this task through *diffusive shock acceleration* (DSA; for a review, see e.g. Malkov & Drury, 2001). DSA is a process in which charges gain energy by diffusing back and forth across the magnetised shock front, as if they were trapped between two mirrors. It is an open question where the high-energy electrons that enter the DSA process come from to begin with. A popular proposal (e.g. Hoeft & Brüggen, 2007) is that they simply stem from the Maxwell–Boltzmann velocity distribution tail of the IGM's thermal electron pool. Recently, Brunetti & Vazza (2020) have proposed that acceleration mechanisms other than DSA, such as second-order Fermi reacceleration in super-Alfvénic turbulence, can also lead to synchrotron emitting particles that illuminate the Cosmic Web's magnetic fields.

All in all, detecting synchrotron radiation from the intergalactic medium in filaments of the Cosmic Web constitutes an upcoming frontier to test models of astrophysical shocks and their radiation mechanisms, trace the missing baryons (e.g. Driver, 2021), and constrain magnetogenesis. Simulations have predicted that the LOFAR might just be sensitive enough to directly image the very brightest shocks in filaments of the Cosmic Web. Such an achievement would usher in a new era of opportunity to study extragalactic magnetism.

1.8.2 Active galactic nuclei and their jet-mediated outflows

Active galactic nuclei (AGN), together with the outflows their jets carry into the IGM, form the second radio probe of magnetism in the Cosmic Web that we have studied in this thesis.

A small-scale look at AGN

Figure 1.13 shows an artist's impression of an AGN and one of its two jets, depicted under such an angle that the accretion disc is directly visible to the observer. Precisely how SMBHs generate jets remains one of astrophysics' major unsolved problems. However, recent VLBI observations of the SMBH accretion disc in Messier

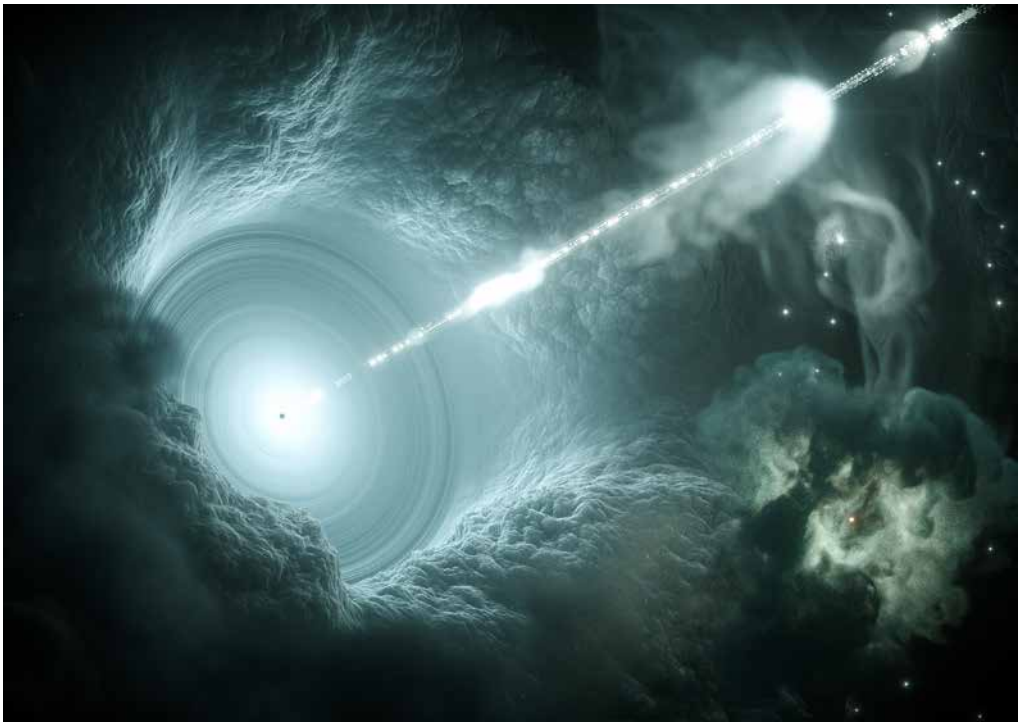


Figure 1.13: Active galactic nuclei launch jets perpendicular to their luminous accretion discs, which can be obscured to observers because of the tori of dust that tend to surround them. A supermassive black hole resides in the middle of each disc. Artist's impression. *From: DESY, Science Communication Lab*

87 (M87) — such as those of Event Horizon Telescope Collaboration et al. (2021a) and Lu et al. (2023) — have led to significant progress in our understanding of jet launching. Initially described as *'a curious straight ray ... apparently connected with the nucleus'* by Lick Observatory's Heber Curtis (Curtis, 1918), the Earth-facing jet of M87 was the first of its kind to be found. The middle-left panel of Fig. 1.14 shows a modern optical view, by the *Hubble Space Telescope*. Nearly sixty years would pass after Curtis' finding before a convincing case was made that jets such as M87's are brought forth by SMBHs. This feat was eventually accomplished by Roger Blandford and Roman Znajek, encouraged by Christodoulou (1970)'s demonstration that the mass of a Kerr black hole consists of a *reducible* and an *irreducible* component — with the reducible component being in principle extractable. As young researchers at the University of Cambridge of the 1970s, they appreciated that the accretion of baryonic matter onto a spinning black hole would generate viscous stress and viscous heat, leading to the dissolution of atoms and molecules into a magnetised plasma. In a groundbreaking publication, Blandford & Znajek (1977) showed that, as a result

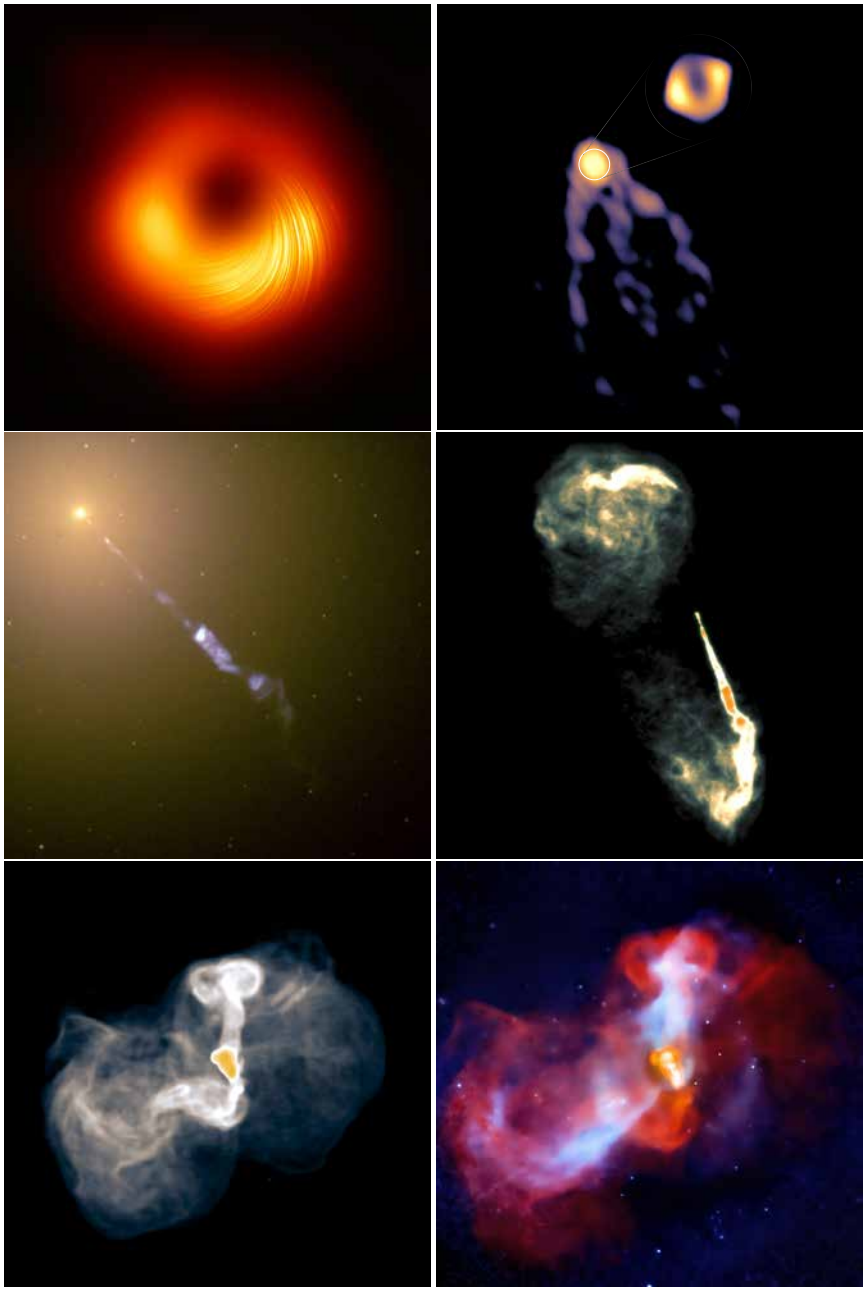


Figure 1.14: Multi-scale, multi-wavelength observations of the AGN and jet-driven outflows of Messier 87, a giant elliptical galaxy in the Virgo Cluster. From top to bottom, and from left to right, the widening views show: the polarised accretion disc at 1.3 mm, the connection between the accretion disc and one of the nascent jets at 3.5 mm, this jet at kiloparsec scale in the optical, the jet and its most recent plasma deposits at 2 cm, the jet's earlier deposits within the Virgo Cluster at large at 90 cm, and the same scene in a blue–red X-ray–radio overlay. The images are not fully rotationally aligned. *From: Biretta et al. (1995), Owen et al. (2000), Werner et al. (2010), Event Horizon Telescope Collaboration et al. (2021a), Lu et al. (2023), NASA, and The Hubble Heritage Team (STScI / AURA)*

of ergospheric frame dragging, magnetic field lines infalling on the black hole twist into a helix aligned with the hole's spin axis. These magnetic fields in motion generate an electric field that accelerates charges. The accelerated charges subsequently inverse Compton scatter with background photons, such as those from the accretion disc, creating gamma rays. In what is called the Breit–Wheeler process, collisions of gamma rays with background photons produce electron–positron pairs (for a review, see Ruffini et al., 2010). These fresh leptons are likewise accelerated by the electric field, inverse Compton scatter with background photons, *et caetera*; in other words, the mechanism just described starts anew. As a result, a cascade of electron–positron pair production arises (e.g. Ford et al., 2018). Finally, the leptons flow away through the helix in both directions — thus forming relativistic jets — and extract angular momentum from the black hole in the process.

Jets do not remain purely leptonic (and may not even start out as such) with evidence existing that powerful jets can emerge from the centres of galaxies energetically dominated by protons (e.g. Blandford et al., 2019). While piercing through the ISM, CGM, and IGM, the jets entrain additional protons and other atomic nuclei. The exact composition of the jets, and the lobes in which they terminate, is a matter of active research (e.g. Croston et al., 2018; Hardcastle & Croston, 2020). One approach is, of course, to turn to simulations; those solving the equations of magnetohydrodynamics have traditionally been most popular. Nowadays, state-of-the-art computer simulations instead use plasma kinetics to capture, for the first time, the electron–positron pair production process central to nascent Blandford–Znajek jets (e.g. Parfrey et al., 2019).[36] Additionally, the low densities in jets imply that jet particles have large mean free paths before colliding with eachother, making the plasma essentially collisionless. Whereas collisionless plasma can be described well with plasma kinetics, magnetohydrodynamics — in which plasmata are modelled as fluids — is only reliable for strongly collisional plasmata.

The first computer simulations that demonstrated the Blandford–Znajek mechanism assumed *weakly* magnetised accretion discs with randomly oriented field lines. This was eventually called the SANE scenario — a questionable acronym standing for *standard and normal evolution*. By contrast, the influential work of Narayan et al. (2003) argued that the accretion flow drags in a poloidal magnetic field that accumulates near the centre, creating a magnetic 'wall' around the black hole that is oriented perpendicular to the disc.[37] Incidentally, blobs of plasma find their way through the barrier, but are repeatedly frustrated: their inward speed is much less than the free-

[36] Strikingly, although these simulations start off without particles, they eventually approach an approximate steady state boasting continuously generated Blandford–Znajek jets.

[37] This barrier forms a sheath around the hole's nascent jets.

fall speed. As the barrier largely blocks the disc's plasma from venturing closer to the SMBH, the accretion disc is said to be locked in place. Meanwhile, the disc's magnetic field is *strong* and coherent. Narayan et al. (2003) called this the *magnetically arrested disc* (MAD) scenario.

The top-left panel of Fig. 1.14 shows the first polarised millimetre-wave observations of M87's SMBH accretion disc (Event Horizon Telescope Collaboration et al., 2021a). In particular, it visualises the vector field of linear polarisation position angles, revealing a manifestly azimuthal pattern. Such an azimuthal pattern arises for regions of the disc where the magnetic field lines locally exhibit a radial or polar orientation (i.e. pointing towards or away from the hole, or along the hole's rotation axis). Consider, for example, a disc region with a radially oriented magnetic field, observed along the hole's polar axis. Leptons spiralling along these field lines generate synchrotron radiation; the part that eventually arrives at our polar observer is linearly polarised. Standard synchrotron radiation theory (e.g. Rybicki & Lightman, 1986) predicts that the polarisation axis is perpendicular to the magnetic field — that is, *azimuthally* oriented. A similar pattern arises for disc regions with a polar magnetic field (Event Horizon Telescope Collaboration et al., 2021b), although our slight inclination relative to the hole's polar axis ($i \sim 17°$), light bending, and relativistic aberration make it harder to immediately see why. Note that it is precisely a large-scale *poloidal* magnetic field that features disc regions where the field is locally radial or polar. Indeed, a detailed comparison between these observations and a suite of general relativistic magnetohydrodynamic simulations (in particular those of Event Horizon Telescope Collaboration et al., 2019b) has singled out the MAD scenario as the only one consistent with the data (Event Horizon Telescope Collaboration et al., 2021b).

The top-right panel of Fig. 1.14 shows a larger-scale view, revealing the connection between the accretion disc and the nascent Earth-facing jet over tens of Schwarzschild radii (Lu et al., 2023). The overall jet shape is parabolic, as predicted by the Blandford–Znajek mechanism (Nakamura et al., 2018). However, close to the black hole — within ten Schwarzschild radii or so — the imaged structure is wider than expected for a parabolic jet. Lu et al. (2023) interpret this as evidence for the presence of another luminous physical component, in addition to the accretion disc and the jet: a non-relativistic, gravitationally unbound wind arising from the disc that plays a key role in collimating the jet into its parabolic shape.

A large-scale look at AGN

AGN control the evolution of their host galaxies by launching winds and jets that warm and rarefy the ISM (e.g. King & Pounds, 2015). Generally, this prevents the

formation of new stars, especially in galactic centres (e.g. Di Matteo et al., 2005); however, star formation might *increase* locally within expanding kiloparsec-radius rings (e.g. Dugan et al., 2017). This feedback mechanism is always on in galaxies with stellar masses $M_\star > 10^{11}\,M_\odot$ (Sabater et al., 2019). Given the central role of AGN in galaxy evolution and cosmology, no cosmological simulation can be trusted without a proper implementation of their physics. For example, cosmological simulations without radio galaxy feedback erroneously predict an abundance of massive starburst galaxies in the centres of galaxy clusters. In reality, heat from RG plasma stops bremsstrahlung-mediated cooling flows that otherwise cause rapid baryonic collapse (e.g. Croton et al., 2006). As an example, the bottom row of Fig. 1.14 shows the plasma deposited by M87 into the Virgo Cluster. However, because SMBHs are astronomical unit–sized, galaxies are kiloparsec-sized (ratio 10^8–10^9), and the Cosmic Web is megaparsec-sized (ratio 10^{11}–10^{12}), it is not possible to build simulations in which a realistic interplay between SMBHs, their host galaxies, and the enveloping Cosmic Web naturally arises. Finding appropriate sub-grid formulations of SMBH activity is therefore a topic of major current interest within the simulation community (e.g. Ward et al., 2022).

Queerly, multi-wavelength observations have revealed a bewildering phenomenological breadth of active galactic nuclei. This breadth is remarkable in light of the no-hair theorem: if the entire formation history of a Kerr black hole lies encoded in just two simple properties — its mass $M_\bullet$ and angular momentum $\vec{J}$ — how can AGN, of which SMBHs are the essence, appear so wildly different? Only part of the solution can be sought in differences in orientation with respect to the observer: although relativistic beaming can explain the exceptional properties of blazars, observations have ruled out simple orientation-based unification schemes for the AGN population as a whole.

The Cosmic Web could be the missing link for a standard model of AGN. Understanding the vast diversity of stars in terms of a standard model of stellar evolution has been a triumph of 20th-century astronomy; similarly, understanding the diversity of active black holes in terms of a 'standard model' could well be within astronomy's reach in the current century. For stars, the initial mass turned out to be the key parameter that determines evolutionary trajectories. However, the initial metallicity — as set by the star's large-scale *environment* — proved to be an important additional parameter guiding stellar evolution. Likewise, evidence is mounting that the large-scale environments of galaxies determine the incidence and strength of SMBH activity, and consequently the properties of AGN and RGs. Striking evidence includes the fact that the comoving quasar number density has decreased after peaking 10 Gyr ago, when both the Universe's mean matter density and the proper galaxy number density

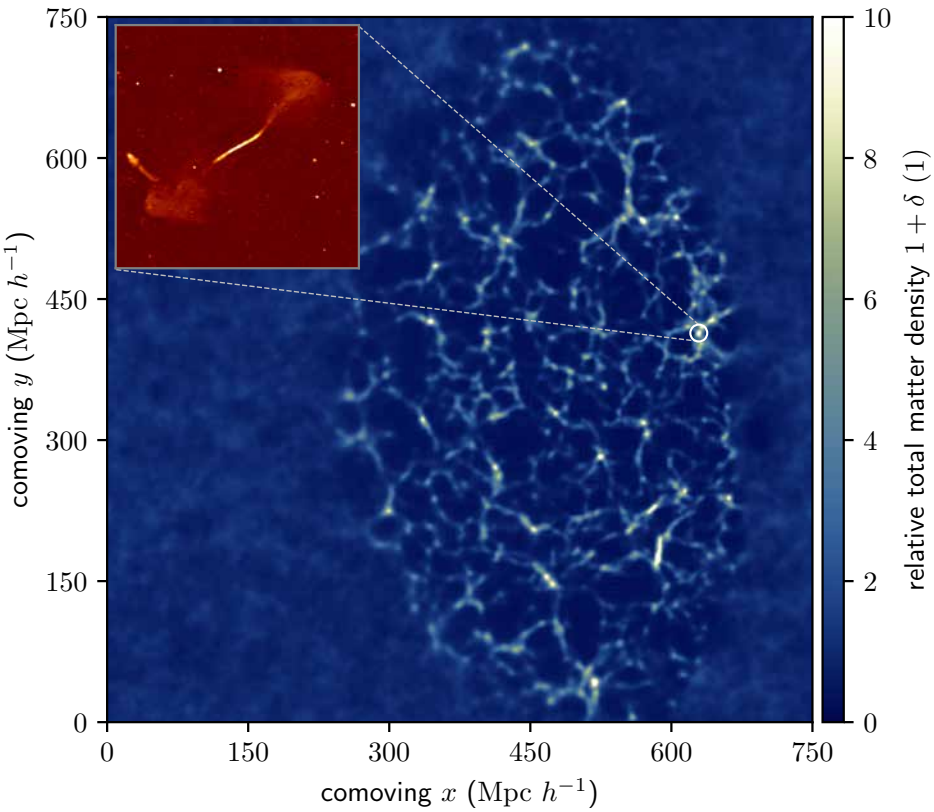


Figure 1.15: Thanks to a Cosmic Web reconstruction breakthrough with Hamiltonian Monte Carlo–guided forward modelling of a dark matter fluid from the Early Universe to the present day, it is now possible to measure the density and gravitational stability of the Cosmic Web around individual active SMBHs and their RGs (Chapter 7). As an example, the GRG shown in the inset, discovered in Chapter 5 and hosted by SDSS J100451.83+543404.2, resides in the galaxy cluster indicated by the white circle. The main panel shows a slice through the BORG SDSS posterior mean (Jasche et al., 2015).

(Conselice et al., 2016) were an order of magnitude higher than they are today. The fact that the population properties of AGN change with cosmic time, perhaps chiefly because of changing densities, suggests that AGN properties also change with large-scale environmental density at a *fixed* epoch. Elegantly, the distribution of galaxies throughout the Cosmic Web, whose Mpc-scale IGM density varies by two orders of magnitude, could therefore already partly explain the enigmatic diversity of AGN.

The influence of the Cosmic Web on supermassive black holes and their activity is likely significant, but remains poorly explored — especially outside of galaxy clusters. To measure this connection, observations of SMBH activity have to be combined with reconstructions of the *actual* Cosmic Web around us. Previous research

has been severely limited by incomplete and unreliable reconstructions that required ad-hoc assumptions. Enticingly, over the last two decades, a breakthrough in large-scale structure reconstruction from spectroscopic galaxy surveys has led to a family of highly principled Bayesian inference techniques that unveil the content, gravitational stability, and formation history of the nearby Cosmic Web (e.g. Kitaura & Enßlin, 2008; Jasche et al., 2010a; Jasche & Kitaura, 2010a; Jasche & Wandelt, 2013; Jasche et al., 2015; Jasche & Lavaux, 2019; Kitaura et al., 2021). Such reconstructions allow one to measure the density, including uncertainty, of the Mpc-scale environment of any galaxy. This is a radical departure from the simple galaxy counting methods that have been used to probe Cosmic Web density before.

Simultaneously, advanced radio interferometers such as the LOFAR and MeerKAT are now operational; the DSA-2000, ngVLA, and SKA will follow suit within this decade. Clearly, bringing these developments together opens an exciting branch of research. In particular, by combining state-of-the-art large-scale structure reconstructions with LOFAR observations of AGN and their RGs, this thesis strives to force a leap in our understanding of the interplay between the Cosmic Web, active black holes, and magnetogenesis. We carry out the first steps towards a measurement of magnetogenesis by radio galaxies — filling a gap in knowledge of how the Universe's largest magnetic fields came to be.

1.9 This thesis

In this thesis, we set out on a quest to measure — through low-frequency radio interferometry — the magnetic and thermodynamic state of filaments of the Cosmic Web. In addition, we propose and investigate the possibility that giant radio galaxies have played a major role in magnetising the Cosmic Web.

1. Calibrating out distortions caused by the ionosphere is the radio equivalent of using adaptive optics in visible-light astronomy. Overcoming the ionospheric challenge is a *condicio sine qua non* to achieve highly sensitive, sharp images of the metre- and decametre-wavelength sky. To detect the faint, magnetised Cosmic Web, one would like to use a 'coronagraphic' masking approach analogous to the one used in the direct imaging of exoplanets. This requires us to confine, as much as possible, the radiation of foreground and background galaxies to their actual sky coordinates, as to subtract and suppress their emission in a subsequent processing step. Calibration methods in operation today do not optimally exploit the correlations between ionospheric distortions of different antennae, times, and sky directions, because they lack an underlying spatiotemporal model of the ionosphere above the interferometer. In Chapter 2,

we develop such a probabilistic spatiotemporal model. We find that a Gaussian random field description for the free electron density $n_e(\vec{r}, t)$ leads to Gaussian random field behaviour of the radio observable $\Delta\mathrm{TEC}(\vec{r}, t, \hat{k})$, whose covariance function — i.e. the infinite-dimensional generalisation of the covariance matrix — we derive explicitly. To first approximation, the ionosphere's n_e will be Gaussian; for this case, our work explicates the best unbiased ΔTEC inference method possible. Our formalism forms the basis of advanced ionospheric calibration techniques trialled for the LOFAR and the DSA-2000.

2. In Chapter 3, we derive the first probabilistic predictions of synchrotron and thermal bremsstrahlung emission from the IGM within clusters and filaments of the *actual,* nearby Cosmic Web. We show that the strongly variable emissivity of cluster merger and accretion shocks, together with their currently unknown sky coordinates, causes the 150 MHz specific intensity in a fixed direction and at degree-scale resolution to vary over two orders of magnitude among random realisations. Our statistical framework is general and thus flexible: it allows one to predict radio skies under other particle acceleration mechanisms than DSA. This research has highlighted the previously underappreciated fact that the quality of such predictions strongly affects the likelihood of success of ongoing filament detection efforts; thus, creating sophisticated predictions is necessary. Detecting radio, microwave, and X-ray emission from Cosmic Web filaments would constrain the IGM's density, chemical enrichment history, and magnetic field strength, and therefore test models of astrophysical shocks and their radiation mechanisms, probe the Universe's missing baryons, and constrain magnetogenesis. Our predictions accelerate the opening of a new, non-thermal window to study the elusive Cosmic Web beyond galaxy clusters.

3. By reprocessing and reimaging the LOFAR Two-metre Sky Survey (LoTSS) DR2 in search of the magnetised Cosmic Web, we have generated the most sensitive arcminute-resolution sky survey yet at metre wavelengths. It reveals a hitherto unknown patchwork of aligned synchrotron stripes of Milky Way origin that, surprisingly, extends to high Galactic latitudes. Figure 1.16 shows a particularly striking region. Serendipitously, as described in Chapter 4, we discover *Alcyoneus*, the largest known AGN–Cosmic Web feedback system, and the Universe's largest known structure of galactic origin. This discovery reveals that galaxies can have spheres of influence around them with diameters of 5 Mpc within which they enrich the IGM with relativistic electrons, heat, and magnetism. Intriguingly, in comparison to other currently known giants, Alcyoneus does not appear generated by a particularly massive black hole.

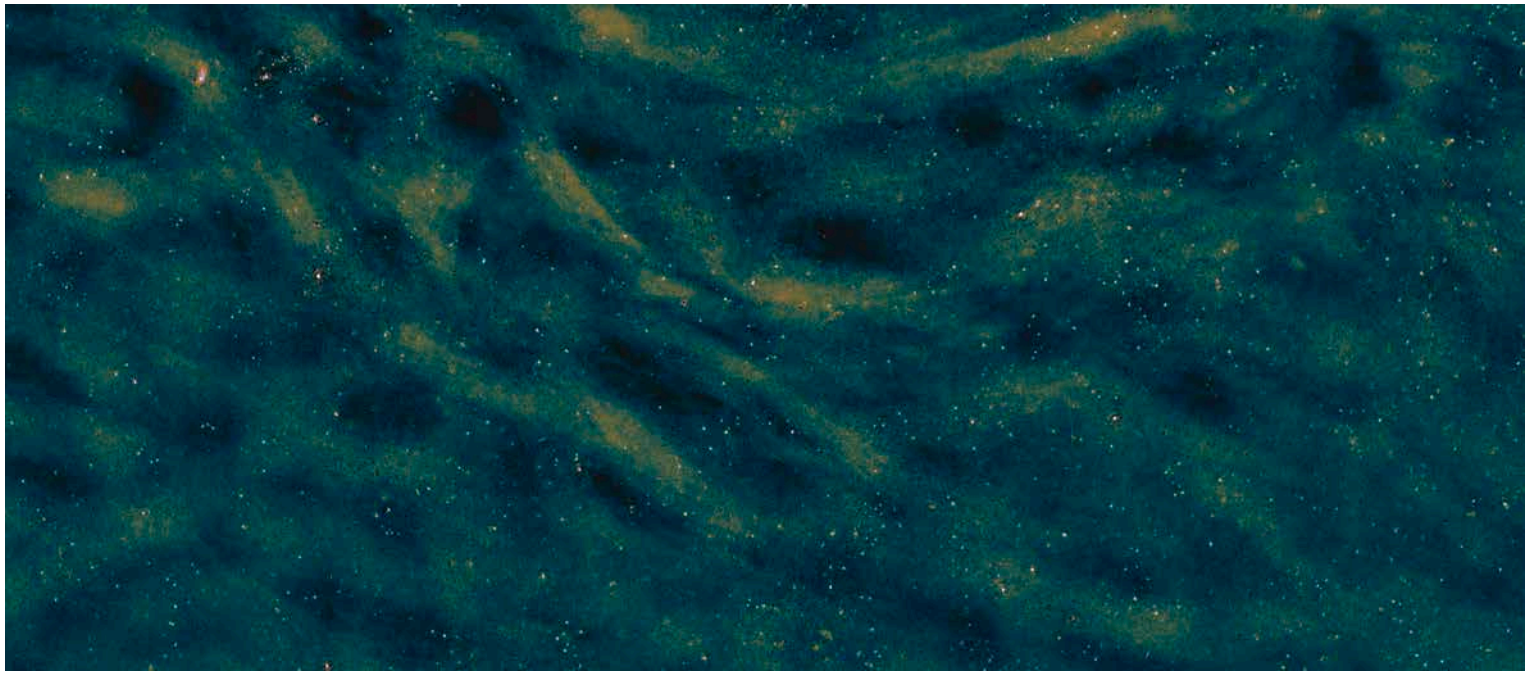


Figure 1.16: By reprocessing the LOFAR Two-metre Sky Survey DR2 (as discussed briefly in Chapters 4 and 6), we have created the most sensitive low-frequency radio survey variety hitherto made. This has led to the serendipitous discovery that our home galaxy fills the metre-wave sky with an intricate patchwork of degree-scale stripes — even up to high Galactic latitudes b. Here we show a $\sim 10^2$ deg^2 mosaic at $b \sim 45°$. Some compact-source residuals inadvertently remain.

The host galaxy's stellar mass is not too high either, and also the jet power is typical for known giants. Alcyoneus' discovery thus shows that exceptionally massive black holes and stellar populations, or exceptionally powerful jets, are not necessary for RGs to achieve multi-Mpc extents. In addition, we develop a Bayesian method to infer RG lobe properties, and apply it to Alcyoneus. This leads to lobe pressure and magnetic field strength measurements of $P = 4.8 \pm 0.3 \cdot 10^{-16}$ Pa and $B = 460 \pm 10$ nG, the lowest found yet.

4. Whereas Chapter 4 reports the discovery of the largest known giant generated by an elliptical galaxy, Chapter 5 reports the discovery of the largest known giant generated by a *spiral* galaxy. With only a handful found, luminous spiral galaxy–generated giants appear to be exceedingly rare. The current giant is generated by NGC 6185, a dynamically disturbed SAa galaxy that is the most massive member of a low-redshift ($z = 0.03$) galaxy group. Just a few methods exist to probe the thermodynamics of the dilute IGM that pervades galaxy groups and Cosmic Web filaments. In this case study, we propose and bring into practice a new IGM temperature estimation technique based on Cosmic Web density reconstructions and a radio image of the GRG's lobes. We infer the temperature at the boundary between the NGC 6185 group and the enveloping warm–hot IGM to be $T_{\rm IGM} = 11^{+12}_{-5} \cdot 10^6$ K. This work bypasses expensive X-ray observations and paves the way to mapping IGM temperatures via GRG lobe dynamics in more places throughout the nearby Cosmic Web.

5. In Chapter 6, we present the discovery of 2060 hitherto unknown giants in both preexisting and novel, low-resolution LoTSS images — more than have been described in all literature studies combined. Besides the giants described in Chapters 4 and 5, spectacular discoveries include a giant whose host has a record-*low* stellar mass $M_\star = 5 \cdot 10^{10}\, M_\odot$, giants whose hosts have record-*high* supermassive black hole masses $M_\bullet \gtrsim 5 \cdot 10^{10}\, M_\odot$, and 13 giants with an angular length larger than that of the full Moon. Among the latter giants is the angularly longest ($\phi = 2°$) known radio galaxy in the Northern Sky, which is also the angularly longest known giant overall. Using these data, we perform a precision measurement of the distribution of giant growth's central dynamical quantity: total length. To do so, we formulate a statistical geometric framework for RGs that is both rigorous and practical. We carefully forward model selection effects, and infer that giant radio galaxy lengths are well described by a Pareto distribution with tail index $\xi = 3.5 \pm 0.5$. This finding is a new observational constraint for models and simulations of RG growth. In addition, we determine the comoving number density of giants, $n_{\rm GRG} = 5 \pm 2\ (100\ {\rm Mpc})^3$,

and the volume-filling fraction of giant radio galaxy lobes in clusters and filaments, $\mathcal{V}_{\mathrm{GRG-CW}} = 5^{+8}_{-2} \cdot 10^{-6}$, both for the first time. We conclude that, at any moment in time, most clusters and filaments — the building blocks of the modern Cosmic Web — do not harbour giants.

6. Chapter 7 demonstrates how modern Cosmic Web reconstructions of the Local Universe can be used to probe the large-scale environment of individual radio galaxies. We pinpoint 260 giants (of which 80% are LoTSS DR2 discoveries) and 1443 general RGs in the BORG SDSS (Jasche et al., 2015). We present the first distributions of the large-scale density around both giant radio galaxies and their smaller kin. To explain giant growth, it has been proposed that giants live in especially low-density environments. However, our results show that currently known giants live in *denser* environments than general RGs. Currently known giants are also more radio luminous than general RG, as required to overcome the surface brightness limits of today's surveys. We propose that their higher densities stem from a positive Cosmic Web density–RG jet power relation that holds universally among RGs. To back up this hypothesis, we present the first quantitative, observational relation between Cosmic Web density and RG radio luminosity, a proxy for jet power. Our findings are consistent with the view that giants are regular, rather than mechanistically special, members of the RG population.

7. Chapter 6 presented a manual, visual search for angularly large ($\phi \gtrsim 5'$) giants in the LoTSS. In Chapter 8, we present a comprehensive search for angularly small ($\phi \lesssim 5'$) giants in the LoTSS DR2, harnessing the power of both citizen science and machine learning. In particular, through the Radio Galaxy Zoo project (Hardcastle et al., 2023) and region-based convolutional neural networks (Mostert et al., 2022), we arrive at a census of giants that now exceeds 10^4 specimen. We furthermore extend the geometrical (giant) radio galaxy population model introduced in Chapter 6, and constrain its parameters using the updated census and Bayesian inference. We obtain improved estimates of the comoving number density of giants, $n_{\mathrm{GRG}} = 13 \pm 10\ (100\ \mathrm{Mpc})^{-3}$, and the instantaneous volume-filling fraction of the lobes of giants in clusters and filaments of the Cosmic Web, $\mathcal{V}_{\mathrm{GRG-CW}} = 1.1 \pm 0.9 \cdot 10^{-5}$. Our work shows that, within the intrinsic RG population, giants are common. Recognising giants as a standard outcome of RG evolution, we reason that the estimated value of $\mathcal{V}_{\mathrm{GRG-CW}}$ is sufficient for giants to have contributed significantly to magnetising the Cosmic Web. Once thought to represent a fringe phenomenon, giant

galactic outflows are now emerging as prime suspects of astrophysical magnetogenesis.

1.10 Future directions

There exists a plethora of exciting future opportunities to extend the work developed in this thesis. One measurement that beckons to be done is a statistical search for synchrotron emission from filaments of the Cosmic Web, using the physical, probabilistic predictions of Chapter 3 — or enhanced varieties. With the newest Bayesian large-scale structure reconstructions, it has become possible to extend these predictions to the entire sky, to push to higher redshifts, to achieve higher spatial resolutions, and to reproduce more accurately the density fields around galaxy clusters. In deep, small–solid angle searches, such as one towards the Ursa Major Supercluster, the ionospheric distortion suppression method of Chapter 2 could be of use.

Furthermore, this thesis has laid the groundwork for using giants in cosmology. In particular, we have developed methodology for measuring the contribution of giant radio galaxies to astrophysical magnetogenesis. The formulae and forward modelling methods for extracting constraints on the giant radio galaxy population from strongly selected data that Chapters 6 and 8 have presented, could yield more detailed and more reliable estimates of giant-induced magnetogenesis if enriched by more data on GRG lobe volumes and magnetic field strengths. To realise this, future work could focus on extending the GRG lobe volume and magnetic field strength inference that Chapters 4 and 5 have put forth. In addition, the machine learning–accelerated discovery of giants of Chapter 8 yields strong promise, and in particular if the procedure's radio–optical association is improved.

Finally, by combining the newest Bayesian large-scale structure reconstructions with large samples of giants and other radio galaxies, we could study the growth of giants and their astrophysical magnetisation potential as a function of Cosmic Web density, expanding upon the work of Chapter 7. Simultaneously, by accurately analysing individual giants and their lobes within the Cosmic Web that envelops them, as an extension of the work of Chapter 5, we could aspire to build a temperature map of the intergalactic medium.

When human ingenuity is linked to radio interferometers and digital computing infrastructure to study magnetic fields in filaments of the Cosmic Web, five great networks conspire: the neurological, the astronomical, the Information Age's, the magnetic, and the cosmological. Together, as this thesis demonstrates, they yield new ways to interrogate Nature, new answers, and new questions too.

Every portrait that is painted with feeling is a portrait of the artist, not of the sitter.

Oscar Wilde, Irish Aestheticist and poet, *The Picture of Dorian Gray* *(1891)*

2

A probabilistic approach to direction-dependent calibration

J. G. Albert, **M. S. S. L. Oei**, R. J. van Weeren, H. T. Intema, H. J. A. Röttgering — *Astronomy & Astrophysics*, 633, 77, 2020

Abstract

Calibrating for direction-dependent ionospheric distortions in visibility data is one of the main technical challenges that must be overcome to advance low-frequency radio astronomy. In this paper, we propose a novel probabilistic, tomographic approach that utilises Gaussian processes to calibrate direction-dependent ionospheric phase distortions in low-frequency interferometric data. We suggest that the ionospheric free electron density can be modelled to good approximation by a Gaussian process restricted to a thick single layer, and show that under this assumption the differential total electron content must also be a Gaussian process. We perform a comparison with a number of other widely successful Gaussian processes on simulated differential total electron contents over a wide range of experimental conditions, and find that, in all experimental conditions, our model is better able to represent observed data and generalise to unseen data. The mean equivalent source shift imposed by our predictive errors are half as large as those of the best competitor model. We find that it is possible to partially constrain the hyperparameters of the ionosphere from sparse-and-noisy observed data. Our model provides an alternative explanation for

observed phase structure functions deviating from Kolmogorov's five-thirds turbulence, turnover at high baselines, and diffractive scale anisotropy. We show that our model performs tomography of the free electron density both implicitly and cheaply. Moreover, we find that even a fast, low-resolution approximation of our model yields better results than the best alternative Gaussian process, implying that the geometric coupling between directions and antennae is a powerful prior that should not be ignored.

Key words: techniques: interferometric – methods: analytical – methods: statistical

2.1 Introduction

Since the dawn of low-frequency radio astronomy, the ionosphere has been a confounding factor in the interpretation of radio data. This is because the ionosphere has a spatially and temporally varying refractive index, which perturbs the radio-frequency radiation that passes through it. This effect becomes more severe at lower frequencies; see (e.g. de Gasperin et al., 2018). The functional relation between the sky brightness distribution — the image — and interferometric observables — the visibilities — is given by the radio interferometry measurement equation (RIME; Hamaker et al., 1996), which models the propagation of radiation along geodesics from source to observer as an ordered set of linear transformations (Jones, 1941).

A mild ionosphere will act as a weak-scattering layer resulting in a perturbed inferred sky brightness distribution, analogous to the phenomenon of seeing in optical astronomy (Wolf, 1969). Furthermore, the perturbation of a geodesic coming from a bright source will deteriorate the image quality far more than geodesics coming from faint sources. Therefore, the image-domain effects of the ionosphere can be dependent on the distribution of bright sources on the celestial sphere, that is they can be heteroscedastic. This severely impacts experiments which require sensitivity to faint structures in radio images. Such studies include the search for the epoch of reionisation (e.g. Patil et al., 2017), probes of the morphology of extended galaxy clusters (e.g. van Weeren et al., 2019), efforts to detect the synchrotron Cosmic Web (e.g. Vernstrom et al., 2017), and analyses of weak gravitational lensing in the radio domain (e.g. Harrison et al., 2016). Importantly, these studies were among the motivations for building the next generation of low-frequency radio telescopes like the Low-Frequency Array (LOFAR), Murchison Widefield Array (MWA), and the future Square Kilometre Array (SKA). Therefore, it is of great relevance to properly calibrate the ionosphere.

Efforts to calibrate interferometric visibilities have evolved over the years from single-

direction, narrow-band, narrow-field-of-view techniques (Cohen, 1973), to more advanced multi-directional, wide-band, wide-field methods (e.g. Kazemi et al., 2011; van Weeren et al., 2016; Tasse et al., 2018). The principle underlying these calibration schemes is that if you start with a rough initial model of the true sky brightness distribution, then you can calibrate against this model and generate an improved sky brightness model. One can then repeat this process for iterative improvement. Among the direction-dependent calibration techniques the most relevant for this paper is facet-based calibration, which applies the single-direction method to piece-wise independent patches of sky called facets. This scheme is possible if there are enough compact bright sources — calibrators — and if sufficient computational resources are available. Ultimately, there are a finite number of calibrators in a field of view and additional techniques must be considered to calibrate all the geodesics involved in the RIME. We note that there are other schemes for ionosphere calibration that do not apply the facet-based approach, such as image domain warping (Hurley-Walker et al., 2017).

There are two different approaches for calibrating all geodesics involved in the RIME. The first approach is to model the interferometric visibilities from first principles and then solve the joint calibration-and-imaging inversion problem. This perspective is the most fundamental; however, applications (e.g. Bouman et al., 2016) of this type are very rare and often restricted to small data volumes due to exploding computational complexity. However, we argue that investing research capital — in small teams to minimise risk — could be fruitful and disrupt the status quo (Wu et al., 2019). The second approach is to treat the piece-wise independent calibration solutions as data and predict calibration solutions for missing geodesics (e.g. Intema et al., 2009a; van Weeren et al., 2016; Tasse et al., 2018). In this paper, we consider an inference problem of the second kind.

In order to perform inference for the calibration along missing geodesics, a prior must be placed on the model. One often-used prior is that the Jones operators are constant over some solution interval. For example, in facet-based calibration the implicit prior is that two geodesics passing through the same facet and originating from the same antenna have the same calibration — which can be thought of a nearest-neighbour interpolation. One often-neglected prior is the 3D correlation structure of the refractive index of the ionosphere. An intuitive motivation for considering this type of prior is as follows: The ionosphere has some intrinsic 3D correlation structure, and since cosmic radio emission propagates as spatially coherent waves. It follows that the correlation structure of the ionosphere should be present in ground-based measurements of the electric field correlation — the visibilities. The scope of this paper is therefore to build the mathematical prior corresponding to the above

intuition.

We arrange this paper by first reviewing some properties of the ionosphere and its relation to interferometric visibilities via differential total electron content in Section 2.2. In Section 2.3, we then introduce a flexible model for the free electron density based on a Gaussian process restricted to a layer. We derive the general relation between the probability measure for free electron density and differential total electron content, and use this to form a strong prior for differential total electron content along missing geodesics. In Section 2.4 we describe a numerical experiment wherein we test our model against other widely successful Gaussian-process models readily available in the literature. In Section 2.5 we show that our prior outperforms the other widely successful priors in all noise regimes and levels of data sparsity. Furthermore, we show that we are able to hierarchically learn the prior from data. In Section 2.6 we provide a justification for the assumptions of the model, and show the equivalence with tomographic inference.

2.2 Ionospheric effects on interferometric visibilities

The telluric ionosphere is formed by the geomagnetic field and a turbulent low-density plasma of various ion species, with bulk flows driven by extreme ultraviolet solar radiation (Kivelson & Russell, 1995). Spatial irregularities in the free electron density (FED) n_e and magnetic field $\mathbf{B}$ of the ionosphere give rise to a variable refractive index n, described by the Appleton–Hartree equation (Cargill, 2007) — here given in a Taylor series expansion to order $O(\nu^{-5})$:

$$n(\mathbf{x}) \approx 1 - \frac{\nu_p^2(\mathbf{x})}{2\nu^2} \pm \frac{\nu_H(\mathbf{x})\nu_p^2(\mathbf{x})}{2\nu^3} - \frac{\nu_p^4(\mathbf{x}) - 4\nu_H^2(\mathbf{x})\nu_p^2(\mathbf{x})}{8\nu^4}. \tag{2.1}$$

Here $\nu_p(\mathbf{x}) = \left(\frac{n_e(\mathbf{x})q^2}{4\pi^2\varepsilon_0 m}\right)^{1/2}$ is the plasma frequency, $\nu_H(\mathbf{x}) = \frac{B(\mathbf{x})q}{2\pi m}$ is the gyro frequency, ν is the frequency of radiation, q is the elementary charge, ε_0 is the vacuum permittivity, and m is the effective electron mass. This form of the Appleton–Hartree equation assumes that the ionospheric plasma is cold and collisionless, that the magnetic field is parallel to the radiation wavevector, and that $\nu \gg \max\{\nu_p, \nu_H\}$. The plus symbol corresponds to the left-handed circularly polarised mode of propagation, and the minus symbol corresponds to the right-handed equivalent. Going forward, we will only consider up to second-order effects, and therefore neglect all effects of polarisation in forthcoming analyses.

In the regime where refractive index variation over one wavelength is small, we can ignore diffraction and interference, or equivalently think about *wave* propaga-

tion as *ray* propagation (e.g. Koopmans, 2010). This approximation is known as the Jeffreys–Wentzel–Kramers–Brillouin approximation (Jeffreys, 1925), which is equivalent to treating this as a scattering problem, and assuming that the scattered wave amplitude is much smaller than the incident wave amplitude — the weak scattering limit (e.g. Yeh, 1962; Wolf, 1969). Light passing through a varying refractive index n will accumulate a wavefront phase proportional to the path length of the geodesic traversed. Let $\mathcal{R}_{\mathbf{x}}^{\hat{\mathbf{k}}}$ be a functional of n, so that the geodesic $\mathcal{R}_{\mathbf{x}}^{\hat{\mathbf{k}}}[n] : [0, \infty) \to \mathbb{R}^3$ maps from some parameter s to points along it. The geodesic connects an Earth-based spatial location $\mathbf{x}$ to a direction on the celestial sphere, indicated by unit vector $\hat{\mathbf{k}}$. The accumulated wavefront phase along the path is then given by

$$\varphi_{\mathbf{x}}^{\hat{\mathbf{k}}} = \frac{2\pi\nu}{c} \int_0^\infty n\left(\mathcal{R}_{\mathbf{x}}^{\hat{\mathbf{k}}}[n](s)\right) - 1 \, \mathrm{d}s, \tag{2.2}$$

where c is the speed of light *in vacuo*. Hamilton's principle of least-action states that geodesics are defined by paths that extremise the total variation of Eq. 2.2.

By substituting Eq. 2.1 into Eq. 2.2, and by considering terms up to second order in ν^{-1} only, we find that the phase deviation induced by the ionosphere is proportional to the integral of the FED along the geodesic, $\varphi_{\mathbf{x}}^{\hat{\mathbf{k}}} \approx \frac{-q^2}{4\pi\varepsilon_0 mc\nu} \tau_{\mathbf{x}}^{\hat{\mathbf{k}}}$, where,

$$\tau_{\mathbf{x}}^{\hat{\mathbf{k}}} \triangleq \int_0^\infty n_e\left(\mathcal{R}_{\mathbf{x}}^{\hat{\mathbf{k}}}[n](s)\right) \, \mathrm{d}s. \tag{2.3}$$

Equation 2.3 defines the total electron content (TEC).

In radio interferometry, the RIME states that the visibilities, being a measure of coherence, are insensitive to unitary transformations of the electric field associated with an electromagnetic wave. Thus, the phase deviation associated with a geodesic is a relative quantity, usually referenced to the phase deviation from another fixed parallel geodesic — the origin of which is called the reference antenna. Going forward we use Latin subscripts to specify geodesics with origins at an antenna location; for example $\mathcal{R}_i^{\hat{\mathbf{k}}}[n]$ is used as shorthand for $\mathcal{R}_{\mathbf{x}_i}^{\hat{\mathbf{k}}}[n]$. Correspondingly, we introduce the notion of differential total electron content (ΔTEC),

$$\tau_{ij}^{\hat{\mathbf{k}}} \triangleq \tau_i^{\hat{\mathbf{k}}} - \tau_j^{\hat{\mathbf{k}}}, \tag{2.4}$$

which is the TEC of $\mathcal{R}_i^{\hat{\mathbf{k}}}[n]$ relative to $\mathcal{R}_j^{\hat{\mathbf{k}}}[n]$.

2.3 Probabilistic relation between FED and ΔTEC: Gaussian process layer model

In this section we derive the probability distribution of ΔTEC given a specific probability distribution for FED. It helps to first introduce the concept of the ray integral (RI) and the corresponding differenced ray integral (DRI). The RI is defined by the linear operator $G_i^{\hat{\mathbf{k}}} : \mathcal{V} \to \mathbb{R}$ mapping from the space of all scalar-valued functions over $\mathbb{R}^3$ to a scalar value according to,

$$G_i^{\hat{\mathbf{k}}} f \triangleq \int_0^\infty f\left(\mathcal{R}_i^{\hat{\mathbf{k}}}[n](s)\right) \, \mathrm{d}s, \tag{2.5}$$

where $f \in \mathcal{V} = \left\{g \mid \int_{\mathbb{R}^3} g^2(\mathbf{x}) d\mathbf{x} < \infty\right\}$. Thus, an RI simply integrates a scalar field along a geodesic. The DRI $\Delta_{ij}^{\hat{\mathbf{k}}} : \mathcal{V} \to \mathbb{R}$ for a scalar field f is straightforwardly defined by

$$\Delta_{ij}^{\hat{\mathbf{k}}} f \triangleq \left(G_i^{\hat{\mathbf{k}}} - G_j^{\hat{\mathbf{k}}}\right) f. \tag{2.6}$$

Both the RI and DRI are linear operators in the usual sense. Using Eqs. 2.3 up to 2.6, we see that

$$\tau_{ij}^{\hat{\mathbf{k}}} = \Delta_{ij}^{\hat{\mathbf{k}}} n_e. \tag{2.7}$$

Let us now specify that the FED is a Gaussian process (GP) restricted to (and indexed by) the set of spatial locations $\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^3 \mid (\mathbf{x} - \mathbf{x}_0) \cdot \hat{z} \in [a - b/2, a + b/2]\}$. This defines a layer of thickness b at height a above some reference point $\mathbf{x}_0$ (see Fig. 2.1). Within this layer the FED is realised from,

$$n_e \sim \mathcal{N}[\mu, K], \tag{2.8}$$

where $\mu : \mathcal{X} \to \mathbb{R}_{>0}$ is the mean function, and $K : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is the covariance kernel function. In other words, the ionospheric FED is regarded to be a uncountable infinite set of random variables (RVs) indexed by spatial locations in $\mathcal{X}$, such that for any finite subset of such locations the corresponding FEDs have a multivariate normal distribution.

In order to extend the scalar field n_e to all of $\mathbb{R}^3$, so that we may apply the operator in Eq. 2.6 to FED, we impose that for all $\mathbf{x} \in \mathbb{R}^3 \setminus \mathcal{X} : n_e(\mathbf{x}) = 0$. This simply means that we take electron density to be zero outside the layer, and makes $G_i^{\hat{\mathbf{k}}}$ well-defined. To further simplify the model, we assume that the mean FED in the layer is constant;

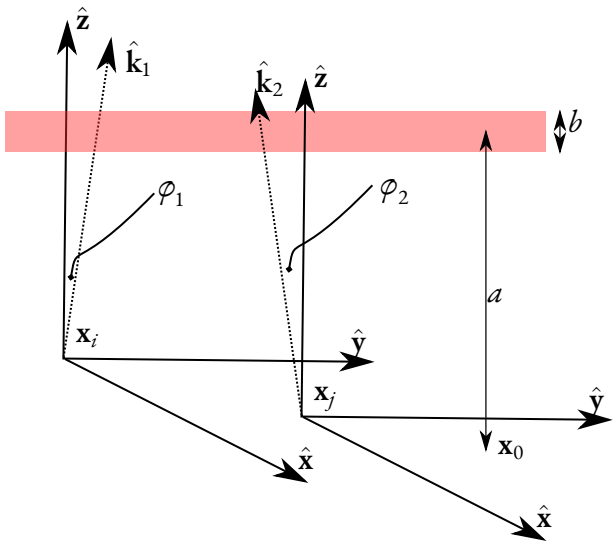


Figure 2.1: Geometry of the toy model. The ionosphere is a layer of thickness b at height a above a reference location $\mathbf{x}_0$. In general, ΔTEC is the TEC along one geodesic minus the TEC along another parallel geodesic. Usually, these geodesics are originating at antennae i and j (locations $\mathbf{x}_i$ and $\mathbf{x}_j$), and pointing in directions $\hat{\mathbf{k}}_1$ and $\hat{\mathbf{k}}_2$, respectively. One common choice is to have a fixed reference antenna for all ΔTEC measurements. The corresponding zenith angles are φ_1 and φ_2.

that is, for all $\mathbf{x} \in \mathcal{X} : \mu(\mathbf{x}) = \bar{n}_e$.

One immediate question that arises pertains to the reasoning behind using a GP to model the FED in the ionosphere. Currently, there is no adequate probabilistic description of the ionosphere that is valid for all times and at the spatial scales that we require. The state-of-the-art characterisation of the ionosphere at the latitude and scales we are concerned with are measurements of the phase structure function, a second-order statistic (Mevius et al., 2016). It is well known that second-order statistics alone do not determine a distribution. In general, all moments are required to characterise a distribution, with a determinancy criterion known as Carleman's condition. Furthermore, the ionosphere is highly dynamic and displays a multitude of behaviours. Jordan et al. (2017) observed four distinct behaviours of the ionosphere above the MWA. It is likely that there are innumerable states of the ionosphere.

Due to the above issue, it is not our intent to precisely model the ionosphere. We rather seek to describe it with a flexible and powerful probabilistic framework. Gaussian processes have several attractive properties, such as the fact that they are highly expressive, easy to interpret, and (in some cases) allow closed-form analytic integration over hypotheses (Rasmussen & Williams, 2006).

However, a Gaussian distribution assigns a non-zero probability density to negative values, which is unphysical. One might instead consider the FED to be a log-GP,

$n_e(\mathbf{x}) = \bar{n}_e \exp \rho(\mathbf{x})$, where the dimensionless quantity $\rho(\mathbf{x})$ is a Gaussian process. In the limit $\rho(\mathbf{x}) \to 0$, we recover that n_e is itself a GP. This is equivalent to saying that the $\sigma_{n_e}/\bar{n}_e \ll 1$. As explained in Section 2.4, we determine estimates of σ_{n_e} and $\bar{n}_e$ by fitting our models to actual observed calibrator data, the International Reference Ionosphere (IRI), and observations taken from Kivelson & Russell (1995). This places the ratio at $\sigma_{n_e}/\bar{n}_e \lesssim 0.06$, suggesting that if the FED can be accurately described with a log-GP, then to good approximation it can also be described with a GP.

We now impose that the geodesics are straight rays, a simplification valid in the weak-scattering limit considered here. The geodesics therefore become $\mathcal{R}_{\mathbf{x}}^{\hat{k}}[n](s) = \mathbf{x} + s\hat{k}$. In practice, strong scattering due to small-scale refractive index variations in the ionosphere is negligible at frequencies far above the plasma frequency when the ionosphere is well-behaved, which is about 90% of the time (Vedantham & Koopmans, 2015). For frequencies $\lesssim 50$ MHz however, this simplification becomes problematic. Under the straight-ray assumption, Equation 2.7 becomes

$$\tau_{ij}^{\hat{\mathbf{k}}} = \int_{s_i^{\hat{\mathbf{k}}-}}^{s_i^{\hat{\mathbf{k}}+}} n_e(\mathbf{x}_i + s\hat{k})\, ds - \int_{s_j^{\hat{\mathbf{k}}-}}^{s_j^{\hat{\mathbf{k}}+}} n_e(\mathbf{x}_j + s'\hat{k})\, ds'. \tag{2.9}$$

Here, the integration limits come from the extension of the FED to spatial locations outside the index-set $\mathcal{X}$, and are given by

$$s_i^{\hat{\mathbf{k}}\pm} = \left(a \pm \frac{b}{2} - (\mathbf{x}_i - \mathbf{x}_0) \cdot \hat{\mathbf{z}}\right) \sec \varphi, \tag{2.10}$$

where $\sec \varphi = (\hat{\mathbf{k}} \cdot \hat{z})^{-1}$ denotes the secant of the zenith angle. It is convenient to colocate the reference point $\mathbf{x}_0$ with one of the antenna locations, and then to also specify this antenna as the reference antenna, i.e. the origin of all reference geodesics. When this choice is made, ΔTEC becomes $\tau_{i0}^{\hat{\mathbf{k}}}$.

Equation 2.7 shows directly that if n_e is a GP, then so is ΔTEC. This can be understood by viewing the RI as the limit of a Riemann sum. We reiterate that every univariate marginal of a multivariate Gaussian is also Gaussian, and that every finite linear combination of Gaussian RVs is again Gaussian. Taking the Riemann sum to the infinitesimal limit preserves this property. Since the DRI is a linear combination of two RIs, the result follows (e.g. Jidling et al., 2018).

The index-set for the ΔTEC GP is the product space of all possible antenna locations and vectors on the unit 2-sphere, $\mathcal{S} = \left\{(\mathbf{x}, \hat{k}) \mid \mathbf{x} \in \mathbb{R}^3, \hat{k} \in \mathbb{S}^2\right\}$. This is analogous to saying that the coordinates of the ΔTEC GP are a tuple of antenna lo-

cation and calibration direction. Thus, given any $\mathbf{y} = (\mathbf{x}, \hat{k}) \in \mathcal{S}$, the ΔTEC is denoted by $\tau_{\mathbf{x}0}^{\hat{k}}$. Because ΔTEC is a GP, its distribution is completely specified by its first two moments.

Since we assume a flat layer geometry, the intersections of two parallel rays with the ionosphere layer have equal lengths of $b \sec \varphi$. This results in the mean TEC of two parallel rays being equal, and thus the first moment of ΔTEC is,

$$m_{\Delta\text{TEC}}(\mathbf{y}) = 0, \tag{2.11}$$

where $\mathbf{y} = (\mathbf{x}_i, \hat{k}) \in \mathcal{S}$. It is important to note that this is not a trivial result. Indeed, a more realistic but slightly more complicated ionosphere layer model would assume the layer follows the curvature of the Earth. In this case, the intersections of two parallel rays with the ionosphere layer have unequal lengths, and the first moment of ΔTEC would depend on the layer geometry and $\bar{n}_e$.

We now derive the second central moment between two ΔTEC along two different geodesics, as visualised in Fig. 2.1.

$$\begin{aligned} K_{\Delta\text{TEC}}(\mathbf{y}, \mathbf{y}') &= \mathbb{E}\left[\tau_{i0}^{\hat{\mathbf{k}}} \tau_{j0}^{\hat{\mathbf{k}}'}\right] && (2.12)\\ &= \mathbb{E}\left[(G_i^{\hat{\mathbf{k}}} n_e - G_0^{\hat{\mathbf{k}}} n_e)(G_j^{\hat{\mathbf{k}}'} n_e - G_0^{\hat{\mathbf{k}}'} n_e)\right] && (2.13)\\ &= I_{ij}^{\hat{\mathbf{k}}\hat{\mathbf{k}}'} + I_{00}^{\hat{\mathbf{k}}\hat{\mathbf{k}}'} - I_{i0}^{\hat{\mathbf{k}}\hat{\mathbf{k}}'} - I_{0j}^{\hat{\mathbf{k}}\hat{\mathbf{k}}'}, && (2.14) \end{aligned}$$

where $\mathbf{y} = (\mathbf{x}_i, \hat{k}) \in \mathcal{S}$ and $\mathbf{y}' = (\mathbf{x}_j, \hat{k}') \in \mathcal{S}$ and,

$$I_{ij}^{\hat{\mathbf{k}}\hat{\mathbf{k}}'} = \int_{s_i^{\hat{\mathbf{k}}-}}^{s_i^{\hat{\mathbf{k}}+}} \int_{s_j^{\hat{\mathbf{k}}'-}}^{s_j^{\hat{\mathbf{k}}'+}} K\left(\mathbf{x}_i + s\hat{k}, \mathbf{x}_j + s'\hat{k}'\right) \mathrm{d}s\mathrm{d}s'. \tag{2.15}$$

We now see that the GP for ΔTEC is zero-mean with a kernel that depends on the kernel of the FED and layer geometry. The layer geometry of the ionosphere enters through the integration limits of Eq. 2.15. Most notably, the physical kernel is non-stationary even if the FED kernel is. Non-stationarity means that the ΔTEC model is not statistically homogeneous, a fact that is well known since antennae near the reference antenna typically have small ionospheric phase corrections. We henceforth refer to Eq. 2.14 as the physical kernel, or our kernel.

Related work. Modelling the ionosphere with a layer has been used in the past. Yeh (1962) performed analysis of transverse spatial covariances of wavefronts (e.g.

Wilcox, 1962; Keller et al., 1964) passing through the ionosphere. Their layer model was motivated by the observation of scintillation of radio waves from satellites (Yeh & Swenson, 1959). One of their results is a simplified variance function, which can be related to the phase structure functions in Section 2.6.4. In van der Tol (2009), a theoretical treatment of ionospheric calibration using a layered ionosphere with Kolmogorov turbulence is done. More recently, Arora et al. (2016) attempted to model a variable-height ionosphere layer above the MWA using GPS measurements for the purpose of modelling a TEC gradient; however unfortunately they concluded that the GPS station array of the MWA is not dense enough to constrain their model.

2.4 Method

In order to investigate the efficacy of the physical kernel for the purpose of modelling ΔTEC we devise a simulation-based experiment. Firstly, we define several observational setups covering a range of calibration pierce-point sparsity and calibration signal-to-noise ratios. A high signal-to-noise-ratio calibration corresponds to better determination of ΔTEC from gains in a real calibration program. Secondly, we characterise two ionosphere varieties as introduced in Section 2.3. Each ionosphere variety is defined by its layer height and thickness, and GP parameters. For each pair of observational setup and ionosphere variety we realise FED along each geodesic and numerically evaluate Eq. 2.7 thereby producing ΔTEC. We then add an amount of white noise to ΔTEC which mimics the uncertainty in a real calibration program with a given calibration signal-to-noise ratio. Finally, we compare the performance of our kernel against several other common kernels used in machine learning on the problem of Gaussian process regression, known as Kriging. In order to do this, we generate ΔTEC for extra geodesics and place them in a held-out dataset. This held-out dataset is used for validation of the predictive performance to new geodesics given the observed ΔTEC. We refer to the other kernels, which we compare our kernel to, as the competitor kernels, and the models that they induce, as the competitor models.

2.4.1 Data generation

For all simulations, we have chosen the core and remote station configuration of LOFAR (van Haarlem et al., 2013), which is a state-of-the-art low-frequency radio array centred in the Netherlands and spread across Europe. The core and remote stations of LOFAR are located within the Netherlands with maximal baselines of 70 km, and we term this array the Dutch LOFAR configuration. We thinned out the array such that no antenna is within 150 m of another. We made this cutoff to reduce the data

Table 2.1: Summary of the parameters of the simulated ionospheres.

Variety	a (km)	b (km)	K_{FED}	σ_{n_e} (m^{-3})	HPD (km)
dawn	250	100	M32	$6 \cdot 10^9$	15
dusk	350	200	EQ	$3 \cdot 10^9$	15

size because nearby antennae add little new information and inevitably raise computational cost. For example, antennae like CS001HBA0 and CS001HBA1 are so close that their joint inclusion was considered redundant.

We consider several different experimental conditions, with a particular choice denoted by η, under which we compare our model to competitors. We consider five levels of pierce-point sparsity: $\{10, 20, 30, 40, 50\}$ directions per field of view (12.6 deg^2). For a given choice of pierce-point sparsity we place twice as many directions along a Fibonacci spiral — scaled to be contained within the field of view — and randomly select half of the points to be in the observed dataset and the other half to be in the held-out dataset. The Fibonacci spiral is slightly overdense in the centre of the field of view, which mimics selecting bright calibrators from a primary-beam uncorrected radio source model. We consider a range of calibration signal-to-noise ratios, which correspond to Gaussian uncertainties of ΔTEC that would be inferred from antenna-based gains in a real calibration program. We therefore consider 11 uncertainty levels on a logarithmic scale from 0.1 to 10 mTECU. A typical state-of-the-art Dutch LOFAR-HBA (high-band antennae) direction-dependent calibration is able to produce on the order of 30 calibration directions (Shimwell et al., 2019), based on the number of bright sources in the field of view, and produce ΔTEC with an uncertainty of approximately 1 mTECU; these levels of sparsity and noise probe above and below nominal LOFAR-HBA observing conditions.

We define an ionosphere variety as an ionosphere layer model with a particular choice of height a, thickness b, mean electron density $\bar{n}_e$, and FED kernel K_{FED} with associated hyperparameters, namely length-scale and variance. As mentioned in Section 2.3, due to the innumerable states of the ionosphere our intent is not to exactly simulate the ionosphere, but rather to derive a flexible model. Therefore, to illustrate the flexibility of our model, we have chosen to experiment with two very different ionosphere varieties which we designate the *dawn* and *dusk* ionosphere varieties. These ionosphere varieties are summarised in Table 2.1. In Section 2.6.4 we show that these ionosphere varieties predict phase structure functions which are indistinguishable from real observations. In order to select the layer height and thickness parameters for the dawn and dusk varieties we took height profiles from the International Reference Ionosphere (IRI; Bilitza & Reinisch, 2008) model.

In order to choose the FED GP kernels and hyperparameters we note that it has been suggested that scintillation is more pronounced during mornings, due to increased FED variation (e.g. Spoelstra, 1983); therefore we chose a rough FED kernel for our dawn simulation. Roughness corresponds to how much spectral power is placed on the shorter length-scales, and also relates to how differentiable realisations from the process are; e.g. see Fig. 2.3. For the dawn ionosphere we choose the Matérn-3/2 (M32) kernel,

$$K_{\mathrm{M32}}(\mathbf{x}, \mathbf{x}') = \sigma_{n_e}^2 \left(1 + \frac{\sqrt{3}}{l_{\mathrm{M32}}}|\mathbf{x} - \mathbf{x}'|\right) \exp\left[\frac{-\sqrt{3}}{l_{\mathrm{M32}}}|\mathbf{x} - \mathbf{x}'|\right], \tag{2.16}$$

which produces realisations that are only once differentiable and therefore rough. For the dusk ionosphere we choose the exponentiated quadratic (EQ) kernel,

$$K_{\mathrm{EQ}}(\mathbf{x}, \mathbf{x}') = \sigma_{n_e}^2 \exp\left[\frac{-|\mathbf{x} - \mathbf{x}'|^2}{2l_{\mathrm{EQ}}^2}\right], \tag{2.17}$$

which produces realisations that are infinitely differentiable and smooth.

Both kernels have two hyperparameters, variance $\sigma_{n_e}^2$ and length-scale l. In order to estimate the FED variation, σ_{n_e}, we used observations from Kivelson & Russell (1995) that TEC measurements are typically on the order of 10 TECU, with variations of about 0.1 TECU. Following the observation that the dawn typically exhibits more scintillation we choose a twice higher σ_{n_e} for our dawn simulation. In addition to the length-scale we consider the half-peak distance (HPD) h, which corresponds to the distance at which the kernel reaches half of its maximum. This parameter has a consistent meaning across all monotonically decreasing isotropic kernels, whereas the meaning of l depends on the kernel. It is related to h by $h \approx 1.177\, l_{\mathrm{EQ}}$ for the EQ and $h \approx 0.969\, l_{\mathrm{M32}}$ for the M32 kernel. The length-scales were chosen by simulating a set of ionospheres with different length-scales and choosing the length-scale that resulted in ΔTEC screens that are visually similar to typical Dutch LOFAR-HBA calibration data. For a given ionosphere variety, we numerically integrate the FED realised from the corresponding GP along the rays in order to compute TEC. From TEC we compute DTEC by taking the difference with the reference antenna TEC. We note that this requires a much higher relative precision in the absolute TEC calculations, since TEC is typically two orders of magnitude large than DTEC. Due to computational limits, we only realise one simulation per experimental condition — that is, we do not average over multiple realisations per experimental condition — however given the large number of experimental conditions there is enough variation to robustly

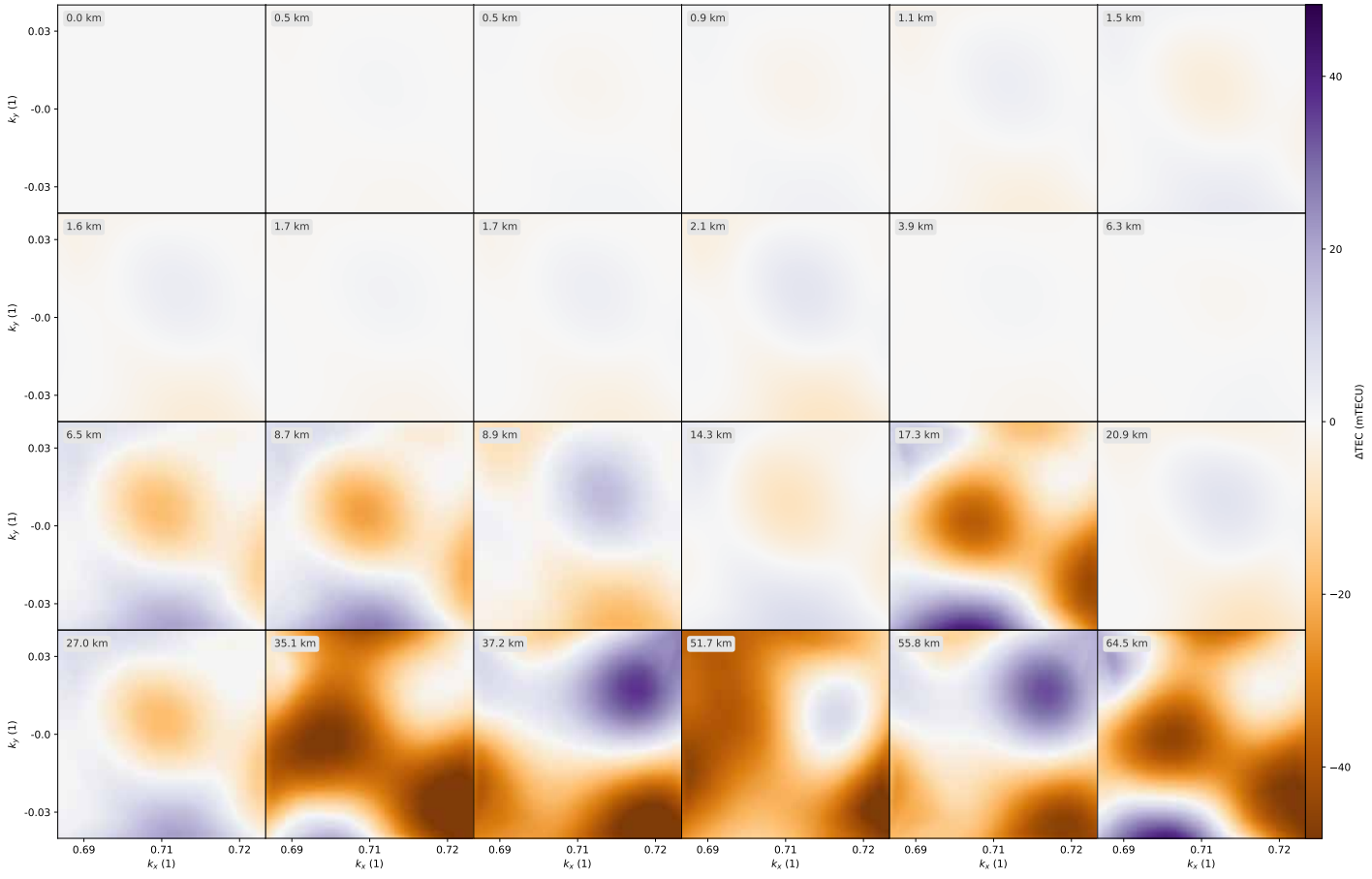


Figure 2.2: Example of antenna-based ΔTEC screens from the dusk ionosphere simulation. Each plot shows the simulated ground truth (noise-free) ΔTEC for each geodesic originating from that station with axes given in direction components k_x and k_y. The inset label gives how far the antenna is from the reference antenna. Antennae further from the reference antenna tend to have a larger magnitude ΔTEC as expected. Each plot box bounds a circular 12.6 deg^2 field of view.

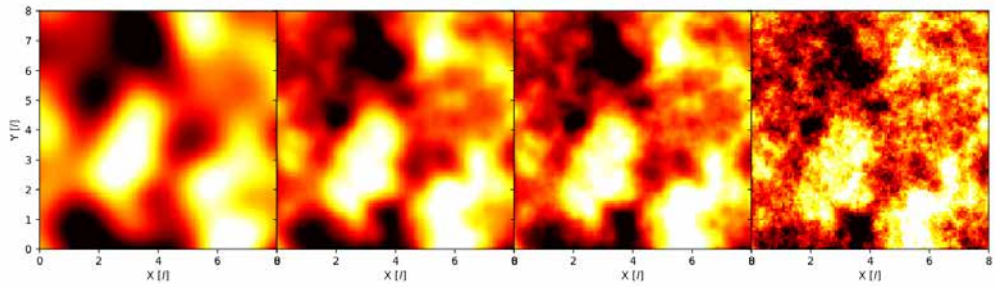


Figure 2.3: Example realisations from exponentiated quadratic, Matérn-5/2, Matérn-3/2, and Matérn-1/2 kernels. The same HPD was used in all kernels, however the smoothness of the resulting process realisation is different for each.

perform a comparative analysis.

2.4.2 Competitor models

For the comparison with competitor models, we compare the physical kernel with: exponentiated quadratic (EQ), Matérn-5/2 (M52), Matérn-3/2 (M32), and Matérn-1/2 (M12) (Rasmussen & Williams, 2006). The EQ and M32 kernels have already been introduced as FED kernels. The M52 and M12 are very similar except for having different roughness properties. Each of these kernels results in a model that spatially smooths — and therefore interpolates — the observed data, but involves a different assumption on the underlying roughness of the function. In order to use these kernels to model ΔTEC, we give each subspace of $\mathcal{S}$ its own kernel and take the product. For example, if K_C is the competitor kernel type, and $(\mathbf{x}, \hat{k}), (\mathbf{x}', \hat{k}') \in \mathcal{S}$, then we form the kernel $K_C((\mathbf{x}, \hat{k}), (\mathbf{x}', \hat{k}')) = K_C^1(\mathbf{x}, \mathbf{x}')K_C^2(\hat{k}, \hat{k}')$ thereby giving each subspace of the index set, $\mathcal{S}$, its own kernel with associated hyperparameters.

Figure 2.4 shows each kernel profile with the same HPD and Fig. 2.3 shows example realisations from the same kernels. It can be visually verified that the M32 kernel has more small-scale variation than the EQ kernel, while maintaining similar large-scale correlation features.

We note that evaluation of the physical kernel requires that a double integral be performed, which can be done in several ways (e.g. Hendriks et al., 2018). In our experiments we tried both explicit adaptive step-size Runge–Kutta quadrature, and two-dimensional trapezoid quadrature. We found via experimentation that we could simply use the trapezoid quadrature with each abscissa partitioned into four equal intervals without loss of effectiveness. However, we chose to use seven partitions. We discuss this choice in Section 2.6.5.

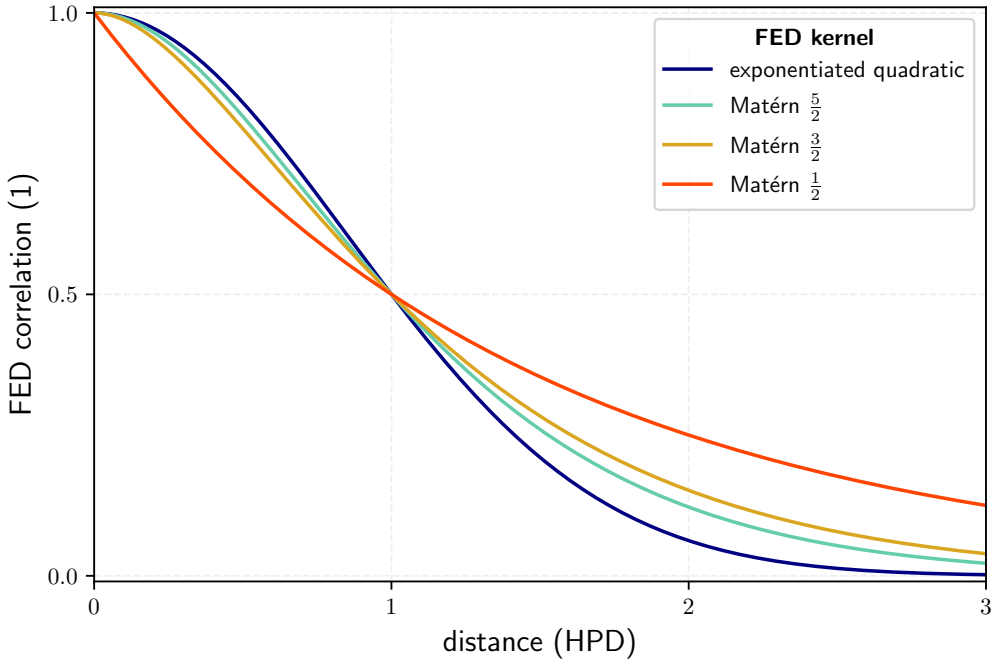


Figure 2.4: Shape of several kernels as a function of separation in units of the HPD of the kernel.

2.4.3 Model comparison

For model comparison, we investigate two key aspects of each model: the ability to accurately model observed ΔTEC, and the ability to accurately infer the held-out ΔTEC. In the language of the machine learning community these are often referred to as minimising the data loss and the generalisation error, respectively. We also investigate the ability to learn the hyperparameters of the physical kernel from sparse data. Finding that the physical model accurately models both observed and held-out ΔTEC, while also being able to learn the hyperparameters, would be a positive outcome.

To measure how well a model represents the observed data, given a particular choice of kernel K and hyperparameters, we compute the log-probability of the observed (LPO) ΔTEC data — Bayesian evidence — which gives a measure of how well a GP fits the data with intrinsically penalised model complexity. If we have data measured at $\mathbf{X} \in \mathcal{S}$ according to $\tau_{\text{obs}} = \tau(\mathbf{X}) + \varepsilon$ where $\varepsilon \sim \mathcal{N}[0, \sigma^2]$ and $\tau(\mathbf{X}) \sim \mathcal{N}[0, K(\mathbf{X}, \mathbf{X})]$, then the LPO is

$$\log P_K(\tau_{\text{obs}}) = \log \mathcal{N}[0, B], \tag{2.18}$$

where $B = K(\mathbf{X}, \mathbf{X}) + \sigma^2 I$. To measure how well a model generalises to unseen data, given a particular choice of kernel K, we compute the conditional log-probability of held-out (LPH) data given the observed data. That is, if we have a held-out dataset measured at $\mathbf{X}^* \in \mathcal{S}$ according to $\tau^*_{\text{obs}} = \tau(\mathbf{X}^*) + \varepsilon^*$ with $\varepsilon^* \sim \mathcal{N}[0, \sigma^2]$, then the LPH conditional on observed τ_{obs} is

$$\begin{aligned}\log P_K\left(\tau^*_{\text{obs}} \mid \tau_{\text{obs}}\right) = \log \mathcal{N}[&K(\mathbf{X}^*, \mathbf{X}) B^{-1} \tau_{\text{obs}}, \\ &B^* - K(\mathbf{X}^*, \mathbf{X}) B^{-1} K(\mathbf{X}, \mathbf{X}^*)]\end{aligned} \tag{2.19}$$

where $B^* = K(\mathbf{X}^*, \mathbf{X}^*) + \sigma^2 I$.

In order to make any claims of model superiority, we will define the following two figures of merit (FOMs),

$$\Delta\text{LPO}_{\text{C}}(\eta) \triangleq \frac{P_{\Delta\text{TEC}}\left(\tau_{\text{obs}} \mid \eta\right)}{P_{\text{C}}\left(\tau_{\text{obs}} \mid \eta\right)}, \tag{2.20}$$

$$\Delta\text{LPH}_{\text{C}}(\eta) \triangleq \frac{P_{\Delta\text{TEC}}\left(\tau^*_{\text{obs}} \mid \tau_{\text{obs}}, \eta\right)}{P_{\text{C}}\left(\tau^*_{\text{obs}} \mid \tau_{\text{obs}}, \eta\right)}, \tag{2.21}$$

where $P_{\Delta\text{TEC}}$ is the probability distribution using the physical kernel and P_{C} is the distribution using a competitor kernel. The variable η represents a particular choice of experimental conditions, for example pierce point sparsity and noise.

These FOMs specify how much more or less probable the physical kernel model is than a competitor for the given choice of experimental conditions, and are therefore useful interpretable numbers capable of discriminating between two models. For example, a $\Delta\text{LPO}_{\text{C}}(\eta)$ value of 1 implies that for the given experimental conditions, η, both models have an equal probability of representing the observed data, and a value of 1.5 would imply that the physical kernel representation is 50% more probable than the competitor kernel. We note that considering the ratio of marginal probabilities is the canonical way of model selection (Rasmussen & Williams, 2006). For a rule-of-thumb using these FOMs, we empirically visually find that models produce noticeably better predictions starting at around 1.10 (10%).

For each choice of experimental conditions, η, and kernel model, we first infer the maximum *a posteriori* estimate of the hyperparameters of the kernel by maximising the marginal log-likelihood of the corresponding GP (Rasmussen & Williams, 2006), which is equivalent to maximising the LPO of that model on the available observed dataset. We maximise the marginal log-likelihood using the variable metric BFGS method, which uses a low-rank approximation to the Hessian to perform gradient-based convex optimisation (Byrd et al., 1995). We use the GPFlow library (Matthews

Table 2.2: Average and standard deviation, over all experimental conditions, of the difference between the learned physical hyperparameters and the true hyperparameters.

Variety	a (km)	b (km)	HPD (km)	$b \cdot \sigma_{n_e}$ (10^{11} km $\cdot$ m^{-3})
dawn	10 ± 10	48 ± 18	4 ± 3	1.9 ± 1.2
dusk	16 ± 9	82 ± 20	1 ± 0.5	2.2 ± 0.3

et al., 2017), which simplifies the algorithmic process considerably. On top of this we perform optimisation from multiple random initialisations to avoid potential local minima. For the physical kernel this corresponds to learning the layer height a and thickness b, and FED kernel length-scale l, and variance $\sigma^2_{n_e}$, and for the competitor kernels this corresponds to learning a variance and the length-scales for each subspace.

2.5 Results

In Table 2.2 we report the average and standard deviation, over all experimental conditions, of the difference between the learned physical hyperparameters and the true hyperparameters, which we term the discrepancy. The optimisation converged in all cases. We observe that for both ionosphere varieties the discrepancy of a is on the order of a $\sim$ 10 km, or a few percent, implying that a can be learned from data. The discrepancy of HPD, is on the order of 1 km, or around 10%, implying the spectral shape information of the FED can be constrained from data. We observe that the discrepancy of layer thickness, b, is large and on the order of 50%. One reason for this is because Eq. 2.15 will scale to first order with b — which is degenerate with the function of σ_{n_e} — and the only way to break the degeneracy is to have enough variation in the secant of the zenith angle. In a sparse and noisy observation of ΔTEC, the secant variation is poor and it is expected that this degeneracy exists. Therefore we also show the product $b \cdot \sigma_{n_e}$, and we see that this compound value discrepancy is smaller by approximately 35%.

In Table 2.3 we summarise the performance of the physical kernel against each competitor kernel. We display the mean of $\Delta\mathrm{LPO}_\mathrm{C}(\eta)$, and $\Delta\mathrm{LPH}_\mathrm{C}(\eta)$ over all experimental conditions, as well as their values at the nominal experimental conditions of 30 directions per 12.6 deg^2, and ΔTEC noise of 1 mTECU, which is indicated with η_nom. We use bold font in Table 2.3 to indicate the best competitor model.

We first consider the ability of each model to represent the observed data. For the dawn ionosphere, the M52 competitor kernel has the best (lowest) $\langle \Delta\mathrm{LPO}_\mathrm{C} \rangle_\eta = 1.55$ and $\Delta\mathrm{LPO}_\mathrm{C}^{\eta_\mathrm{nom}} = 1.46$, implying that the M52 kernel model is 55% and 46% less

Table 2.3: Shows the probability ratio FOMs (see text) averaged over experimental conditions and at nominal conditions. Larger values indicate that the physical model is more probable. Bold face indicates the best performing competitor model (lower number).

	$\langle \Delta\mathrm{LPO}_\mathrm{C} \rangle_\eta$	$\Delta\mathrm{LPO}_\mathrm{C}^{\eta_\mathrm{nom}}$	$\langle \Delta\mathrm{LPH}_\mathrm{C} \rangle_\eta$	$\Delta\mathrm{LPH}_\mathrm{C}^{\eta_\mathrm{nom}}$
		dawn		
M12	1.86	1.79	1.82	1.61
M32	1.56	1.49	1.50	1.33
M52	**1.55**	**1.46**	**1.49**	**1.31**
EQ	1.63	1.48	1.84	1.35
		dusk		
M12	2.72	2.19	2.24	1.73
M32	1.96	1.69	1.50	1.29
M52	1.82	1.60	1.33	1.20
EQ	**1.73**	**1.54**	**1.16**	**1.12**

probable than the physical kernel model on average over all experimental conditions, and at nominal conditions, respectively. We note that the M32 kernel produced similar results. For the dusk ionosphere, the EQ kernel model is likewise the best among all competitors, being only 73% and 54% less probable than the physical kernel model on average over all experimental conditions, and at nominal conditions, respectively. In all experimental conditions, the physical model provides a significantly more probable explanation of the observed data.

We now consider the ability of each model to infer the held-out data. For the dawn ionosphere, the M52 competitor kernel has the best (lowest) $\langle \Delta\mathrm{LPH}_\mathrm{C} \rangle_\eta = 1.49$ and $\Delta\mathrm{LPO}_\mathrm{C}^{\eta_\mathrm{nom}} = 1.31$, implying that the M52 kernel prediction is 49% and 31% less probable than the physical kernel model on average over all experimental conditions, and at nominal conditions, respectively. We note that the M32 kernel produced similar results. For the dusk ionosphere, the EQ kernel model is likewise the best among all competitors, with predictions only 16% and 12% less probable than the physical kernel model on average over all experimental conditions, and at nominal conditions, respectively. In the case of the rougher dawn ionosphere, the physical model provides a significantly more probable prediction of the held-out data in all experimental conditions. However, for the smoother dusk ionosphere at nominal conditions, the physical model is only 12% more probable than the EQ kernel model, which is not very significant.

Figure 2.7 shows a visual comparison of the predictive distributions of the physical and best competitor kernel for the dawn ionosphere, for nominal and sparse-and-noisy conditions, for a subset of antennae over the field of view. In the first row we

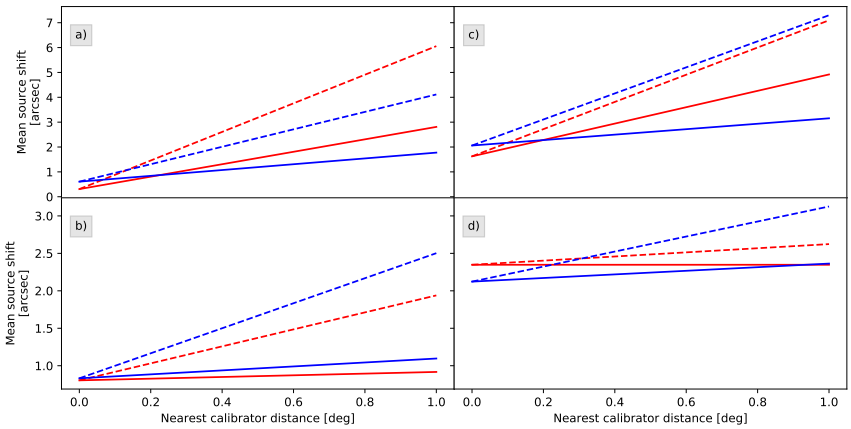


Figure 2.5: Mean equivalent source shift as a function of angular distance from the nearest calibrator caused by inference errors from the ground truth for **a)** remote stations (RS; $>$ 3 km from the reference antenna) at nominal conditions (30 calibrators for 12.6deg^2 and 1 mTECU noise), **b)** core stations (CS; $<$ 2 km) at nominal conditions, **c)** RS with sparse-and-noisy conditions (10 calibrators for 12.6deg^2 and 2.6 mTECU noise), and **d)** CS with sparse-and-noisy conditions. The *dash line styles* are the best competitor models (see text), the *solid line styles* are the physical model. The red lines are dawn ionospheres, and the blue lines are dusk ionospheres.

show the ground truth and observed data. In the second and third rows we plot the mean of the predictive distribution with uncertainty contours of the physical and best competitor models, respectively. At nominal conditions, the predictive means of the best competitor and physical models both visually appear to follow the shape of the ground truth. However, for the sparse-and-noisy condition, only the physical model predictive mean visually follows the shape of the ground truth. The uncertainty contours of the physical model vary in height slowly over the field of view, and are on the order of 0.5–1 mTECU. The uncertainty contours for the physical model indicate that we can trust the predictions near the edges of the field of view. In comparison, the uncertainty contours of the best competitor model steeply grow in regions without calibrators, and are on the order of 2–10 mTECU, indicating that only predictions in densely sampled regions should be trusted.

The last two rows show the residuals between the posterior means and the ground truth for the physical and best competitor models respectively. From this we can see that even when the best-competitor predictive mean visually appears to follow the ground truth the residuals are larger in magnitude than those of the physical models.

In order to quantify the effect of the residuals, a ΔTEC error, $\delta\tau$, can be conveniently represented by the equivalent source shift for a source at zenith on a baseline

of r,

$$\delta l \approx \frac{q^2}{\varepsilon_0 m_e \nu^2 r} \delta\tau \tag{2.22}$$

$$\approx 1.16'' \left(\frac{r}{10\ \mathrm{km}}\right)^{-1} \left(\frac{\nu}{150\ \mathrm{MHz}}\right)^{-2} \left(\frac{\delta\tau}{\mathrm{mTECU}}\right). \tag{2.23}$$

Figure 2.5 shows the mean linear regression of the absolute equivalent source shift of the residuals for each point in the held-out data set, for nominal (left) and sparse-and-noisy (right) conditions, at 150 MHz on a baseline of 10 km, as a function of the nearest calibrator. For visual clarity we have not plotted confidence intervals, however we note that for nominal conditions the 1σ confidence width is about $2''$ and for the sparse-and-noisy conditions it is about $4''$. Because there are few nearest-calibrator distances exceeding 1 degree at nominal conditions, we only perform a linear regression out to 1 degree.

The upper row shows the source shift for the remote stations (RS) residuals, which are generally much larger than the source shifts for core stations (CS) in the bottom row, since the CS antennae are much closer to the reference antenna and have smaller ΔTEC variance. We observe that the physical model (dashed line styles) generally has a shallower slope than the best competitor model (solid line styles). Indeed, for the CS antennae the physical model source shift is almost independent of distance from a calibrator. The offset from zero at 0 degrees of separation comes from the fact that the predictive variance cannot be less than the variance of the observations; see the definition of B^* in Eq. 2.19. At 1 degree of separation, the physical model mean equivalent source shift is approximately half of that of the best competitor model. At 0 degrees of separation, the mean source shift is the same for both models as expected.

2.6 Discussion

2.6.1 Model selection bias

Our derived model is a probabilistic model informed by the physics of the problem. We use the same physical model to simulate the data. Therefore it should perform better than any other general-purpose model. The fact that we simulate from the same physical model as used to derive the probabilistic model does not detract from the efficacy of the proposed model to represent the data. Indeed, it should be seen as a reason for preferring physics-based approaches when the physics are rightly known. The Gaussian random field layer model for the ionosphere has been a useful prescription for the ionosphere for a long time (e.g. Yeh & Swenson, 1959).

One type of bias that should be addressed is the fact that we assume we know the FED kernel type of the ionosphere. We do not show, for example, what happens when we assume the wrong FED kernel. However, since we are able to converge on optimal hyperparameters for a given choice of FED kernel, we can therefore imagine performing model selection based on the values of the Bayesian evidence (LPO) for different candidate FED kernels. Thus, we can assume that we could correctly select the right FED kernel in all the experimental conditions that we chose in this work.

2.6.2 Implicit tomography

The results of Section 2.5 indicate that the physical model provides a better explanation of ΔTEC data than any of the competitor models. One might ask how it performs so well. The approach we present is closely linked to tomography, where (possibly non-linear) projections of a physical field are inverted for a scalar field. In a classical tomographic approach, the posterior distribution for the FED given observed ΔTEC data would be inferred and then the predictive ΔTEC would be calculated from the FED, marginalising over all possible FEDs,

$$P(\tau \mid \tau_{\text{obs}}) = \int_{n_e} P(\tau \mid n_e)\, P(n_e \mid \tau_{\text{obs}})\, \mathrm{d}n_e, \tag{2.24}$$

where $n_e = \{n_e(\mathbf{x}) \mid \mathbf{x} \in \mathcal{X}\}$ is the set of FEDs over the entire index set $\mathcal{X}$, $\tau = \{\tau_{\mathbf{x}}^{\hat{k}} \mid (\mathbf{x}, \hat{k}) \in \mathcal{S}_* \subset \mathcal{S}\}$ is the ΔTEC over some subset $\mathcal{S}_*$ of the index set $\mathcal{S}$, $\tau_{\text{obs}} = \{\tau_{\mathbf{x}}^{\hat{k}} + \varepsilon \mid (\mathbf{x}, \hat{k}) \in \mathcal{S}_{\text{obs}} \subset \mathcal{S}\}$ is the observed ΔTEC over a different subset $\mathcal{S}_{\text{obs}}$ of $\mathcal{S}$, and $\varepsilon \sim \mathcal{N}[0, \sigma^2 I]$.

In our model, the associated equation for $P(\tau \mid \tau_{\text{obs}})$ is found by conditioning the joint distribution on the observed ΔTEC and then marginalising out FED,

$$P(\tau \mid \tau_{\text{obs}}) = \int_{n_e} P(n_e, \tau \mid \tau_{\text{obs}})\, \mathrm{d}n_e \tag{2.25}$$

$$= \int_{n_e} P(n_e \mid \tau_{\text{obs}})\, P(\tau \mid n_e, \tau_{\text{obs}})\, \mathrm{d}n_e, \tag{2.26}$$

where in the second line we used the product rule of probability distributions (Kolmogorov, 1956). By working through Eqs. 2.24 and 2.26, we discover that if $P(\tau \mid n_e) = P(\tau \mid n_e, \tau_{\text{obs}})$ is true, then our method is equivalent to first inferring FED and then using that distribution to calculate ΔTEC. In Appendix 2.A1 we prove that the expressions in Eqs. 2.24 and 2.26 are equal due to the linear relation between FED and ΔTEC because the sum of two Gaussian RVs is again Gaussian. Most importantly,

this result would not be true if ΔTEC was a non-linear projection of FED.

We refer to this as implicit tomography as opposed to explicit tomography, wherein the FED distribution would be computed first and the ΔTEC computed second (e.g. Jidling et al., 2018). This explains why our kernel is able to accurately predict ΔTEC in regions without nearby calibrators. The computational savings of our approach is many-fold compared with performing explicit tomography, since the amount of memory that would be required to evaluate the predictive distribution of FED everywhere would be prohibitive. Finally, the use of GPs to model ray integrals of a GP scalar field is used in the seismic physics community for performing tomography of the interior of the Earth.

2.6.3 Temporal differential TEC correlations

One clearly missing aspect is the temporal evolution of the ionosphere. In this work we have considered instantaneous realisations of the FED from a spatial GP; however, the inclusion of time in the FED GP is straightforward in principle. One way to include time is by appending a time dimension to the FED kernel, which would mimic internal (e.g. turbulence-driven) evolution of the FED field. Another possibility is the application of a frozen flow assumption, wherein the ionospheric time evolution is dominated by a wind of constant velocity v, so that $n_e(\mathbf{x}, t) = n_e^0(\mathbf{x} - vt)$. Here, n_e^0 represents the FED at time $t = 0$, and n_e is a translation over the array as time progresses. In modelling a real dataset with frozen flow the velocity could be assumed to be *piece-wise* constant in time. We briefly experimented with frozen flow and found hyperparameter optimisation to be sensitive to the initial starting point due to the presence of many local optima far from the ground-truth hyperparameters. We suggest that a different velocity parametrisation might facilitate implementation of the frozen flow approach.

2.6.4 Structure function turnover and anisotropic diffractive scale

The power spectrum is often used to characterise the second-order statistics of a stationary random medium, since according to Bochner's theorem the power spectrum is uniquely related to the covariance function via a Fourier transform. In 1941, Kolmogorov (translated from Russian in Kolmogorov, 1991) famously postulated that turbulence of incompressible fluids with very large Reynolds numbers displays self-similarity. From this assumption, he used dimensional analysis to show that the necessary power spectrum of self-similar turbulence is a power-law with an exponent of -5/3. A convenient related measurable function for the ionosphere is the phase struc-

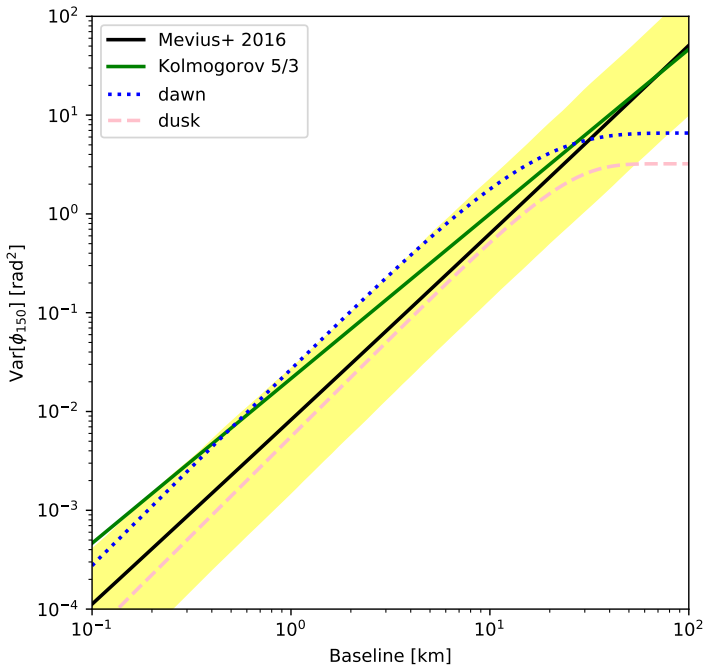


Figure 2.6: Structure functions predicted by our model compared with observations and theory. The dotted and dashed lines show the phase structure function corresponding to the physical kernel, with the dawn and dusk configurations, respectively (see Fig. 2.1). Along side is the predicted structure function of Kolmogorov turbulence with a diffraction scale of 10 km, and the structure function constrained from observations in Mevius et al. (2016) with 1σ confidence region in yellow. We note that Mevius et al. (2016) observe a turnover, but do not characterise it, and therefore we do not attempt to plot it here.

ture function (van der Tol, 2009),

$$D(r) = \langle (\varphi_\nu(R) - \varphi_\nu(r+R))^2 \rangle_R \tag{2.27}$$

$$\triangleq \left(\frac{r}{r_{\text{diff}}}\right)^\beta, \tag{2.28}$$

where the expectation is locally over locations far from the boundaries of the turbulent medium, which is often characterised by an outer scale. The quantity r_{diff} is referred to as the diffractive scale, and is defined as the length where the structure function is 1 rad^2. Under Kolmogorov's theory of 1941, $\beta = 5/3$. Observations from 29 LOFAR pointings constrain β to be 1.89 ± 0.1, slightly higher than predicted by Kolmogorov's theory, and the diffractive scale to range from 5 to 30 km (Mevius et al., 2016).

In Fig. 2.6 the structure functions of the physical kernel are shown for the dawn

and dusk varieties, alongside Kolmogorov's $\beta = 5/3$ and the Mevius et al. (2016) observations. Though not plotted, Mevius et al. (2016) also find that there is a hint of a turnover in the structure functions they observed, which they suggest might be a result of an outer scale in the context of Kolmogorov turbulence. However, these latter authors conclude that longer baselines are needed to properly confirm the turnover and its nature. The dawn and dusk structure functions are nearly parallel with observations, and have turnovers that result because the FED covariance functions decay to zero monotonically and rapidly beyond the HPD. Interestingly, despite the fact that the FED kernels used for the dawn and dusk ionospheres have different spectral shapes, the structure functions have similar slopes. The difference between the dawn and dusk structure functions can be seen in the curvature of their turnovers.

Our model provides an explanation for the observed shape of structure functions, which Kolmogorov's theory of 1941 fails to provide, namely the existence of a turnover, and a slope deviating from five-thirds. Specifically, a turnover requires only FED correlations that are stationary, isotropic, and monotonically decreasing (SIMD). Both the dawn and dusk ionosphere varieties experimented with predict slopes compatible with observations. Moreover, as shown in Appendix 2.A2, our model in conjunction with the SIMD FED kernel is falsifiable by observing a lack of plateau.

Mevius et al. (2016) also observe anisotropy in the measured r_{diff} as a function of pointing direction, and suggest that it is due to FED structures aligned with magnetic field lines (Loi et al., 2015). In total, 12 out of 29 (40%) of their observations show anisotropy unaligned with the magnetic field lines of Earth. We propose a complementary explanation for the anisotropy of diffractive scale, without appealing to magnetic field lines. Our model implies that diffractive scale monotonically decreases with zenith angle. This is a result of the non-stationarity of the physical kernel even if the FED is stationary.

2.6.5 Low-accuracy numerical integration

The numerical integration required to compute Eq. 2.14 is performed using the 2D trapezoid rule. This requires the selection of a number of partitions along the ray. The computational complexity scales quadratically with the number of partitions chosen, and thus a trade-off between accuracy and speed must be chosen. We found the relative error (using the Frobenius norm) to be 80% with two partitions, 20% with three partitions, 10% with four partitions, and 6% with seven partitions. After experimentation it was surprisingly found that two partitions was sufficient to beat all competitor models, and that marginal improvement occurs after five partitions. This suggests that even a low-accuracy approximation of our model encodes enough ge-

ometric information to make it a powerful tool in describing the ionosphere. Ultimately, we chose to use four partitions for our trials.

2.7 Conclusion

In this work, we put forth a probabilistic description of antenna-based ionospheric phase distortions, which we call the physical model. We assumed a single weakly scattering ionosphere layer with arbitrary height and thickness, and free electron density (FED) described by a Gaussian process (GP). We argue that modelling the FED with a GP locally about the mean is a strong assumption due to the small ratio of FED variation to mean as evidenced from ionosphere models. We show that under these assumptions the directly observable ΔTEC must also be a GP. We provide a mean and covariance function that are analytically related to the FED GP mean and covariance function, the ionosphere height and thickness, and the geometry of the interferometric array.

In order to validate the efficacy of our model, we simulated two varieties of ionosphere — a dawn (rough FED) and dusk (smooth FED) scenario — and computed the corresponding ΔTEC for the Dutch LOFAR-HBA configuration over a wide range of experimental conditions including nominal and sparse-and-noisy conditions. We compared this physical kernel to other widely successful competitor GP models that might naively be applied to the same problem. Our results show that we are always able to learn the FED GP hyperparameters and layer height — including from sparse-and-noisy ΔTEC data — and that the layer thickness could likely be learned if a height prior was provided. In general, the physical model is better able to represent observed data and generalises better to unseen data.

Visual validation of the predictive distributions of ΔTEC show that the physical model can accurately infer ΔTEC in regions far from the nearest calibrator. Residuals from the physical model (0.5–1 mTECU) are smaller and less correlated than those of competitor models (2–10 mTECU). In terms of mean equivalent source shift resulting from incorrect predictions, the physical model mean equivalent source shift is approximately half of that of the best competitor model. We show that our model performs implicit tomographic inference at low cost, which is because ΔTEC is a linear projection of FED and the FED is a GP. We suggest possible extensions to incorporate time, including frozen flow and appending the FED spectrum with a temporal power spectrum. Our model provides an alternative explanation for the Mevius et al. (2016) observations: phase structure function slope deviating from Kolmogorov's five-thirds, the turnover on large baselines, and diffractive scale anisotropy.

In the near future, we will apply this model to LOFAR-HBA datasets and per-

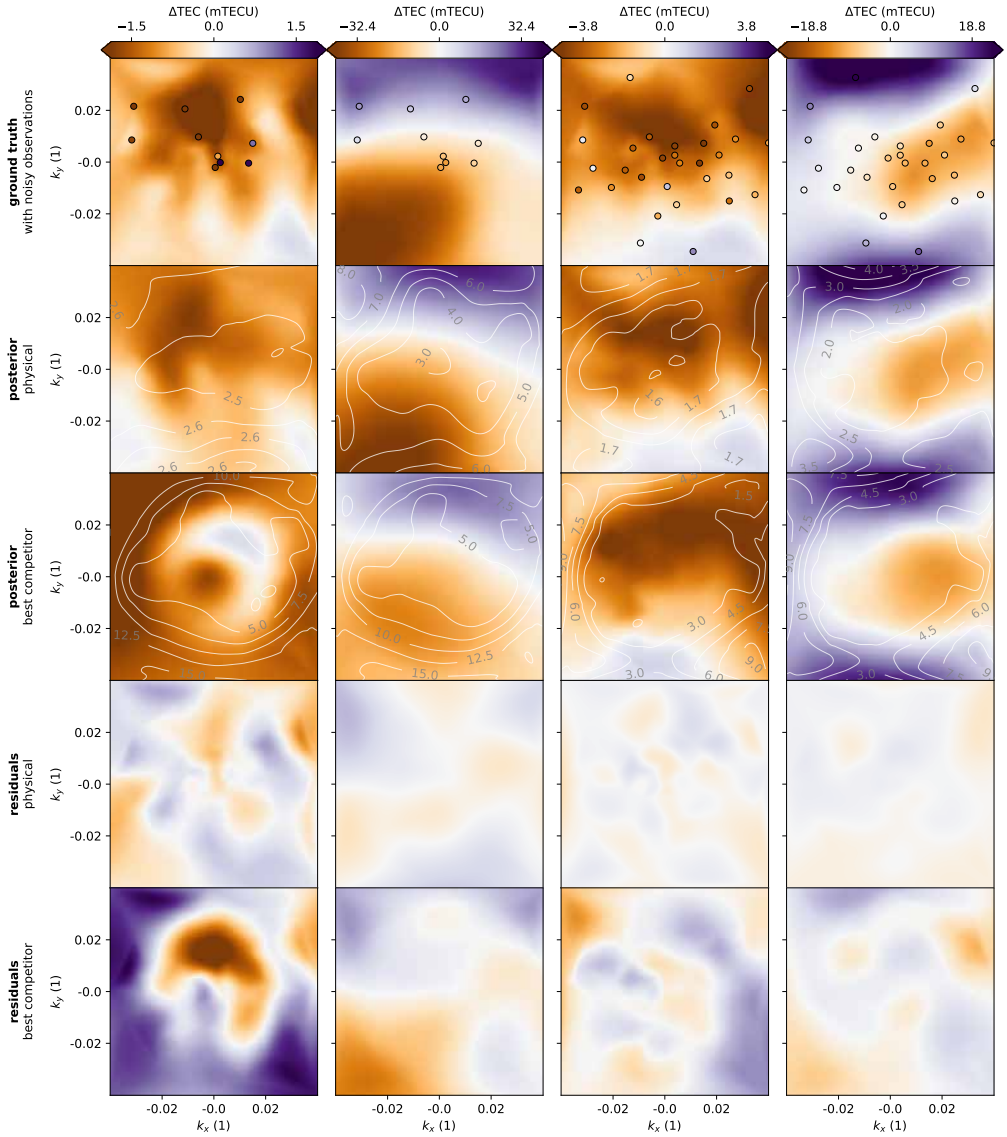


Figure 2.7: Example visual comparison of the predictive performance of our physical model with that of the best competitor model for the dawn ionosphere. *First row:* ground truth ΔTEC overlaid with noisy draws from the ground truth — the observations. *Second and third rows:* posterior mean with uncertainty contours for the physical model and best competitor model respectively. *Fourth and fifth rows:* residuals between posterior means and ground truth for the physical model and best competitor model respectively. *First two columns:* results for a central antenna (near to reference antenna) and a remote station (far from reference antenna), given 10 directions and 2.5 mTECU noise. *Last two columns:* results for a central antenna and a remote station, given 30 directions and 1.6 mTECU noise.

form precise ionospheric calibration for all bright sources in the field of view. It is

envisioned that this will lead to clearer views of the sky at the longest wavelengths, empowering a plethora of science goals.

Author contributions

Joshua came up with the idea of the ionospheric model, and set up a first quantitative outline assuming lognormally distributed ionospheric FEDs. Joshua and Martijn together found the correct Gaussian process equations. Martijn simulated the ionospheric FEDs and ΔTECs, and Joshua compared the Gaussian process regression performance of different kernels. Joshua led the writing of the article, with important contributions from Martijn.

2.A1 Derivation of tomographic equivalence

We now explicitly prove the assertion that Eq. 2.24 is equal to Eq. 2.26, that is,

$$\int P(\tau \mid n_e)\, P(n_e \mid \tau_{\text{obs}})\ \mathrm{d}n_e = \int P(n_e, \tau \mid \tau_{\text{obs}})\ \mathrm{d}n_e. \tag{2.29}$$

We note that we sometimes use the notation $\mathcal{N}[a \mid m_a, C_a]$ which is equivalent to $a \sim \mathcal{N}[m_a, C_a]$.

We define the matrix representation of the DRI operator in Eq. 2.6, $\Delta_* n_e = \{\Delta_{\mathbf{x}}^{\hat{k}} n_e \mid (\mathbf{x}, \hat{k}) \in \mathcal{S}_*\}$, and likewise let Δ be the matrix representation over the index set $\mathcal{S}_{\text{obs}}$. Similarly, the matrix representation of the FED kernel — the Gram matrix — is $K = \{K(\mathbf{x}, \mathbf{x}') \mid \mathbf{x}, \mathbf{x}' \in \mathcal{X}\}$. Using these matrix representation we have the following joint distribution,

$$P(n_e, \tau, \tau_{\text{obs}}) = \mathcal{N}\begin{bmatrix} \bar{n}_e & K & K\Delta_*^T & K\Delta^T \\ 0\,, & \Delta_* K & \Delta_* K \Delta_*^T & \Delta_* K \Delta^T \\ 0 & \Delta K & \Delta K \Delta_*^T & \Delta K \Delta^T + \sigma^2 I \end{bmatrix}. \tag{2.30}$$

Let us first work out the left-hand side (LHS) of Eq. 2.29. Because $\tau = \Delta_* n_e$, and using standard Gaussian identities we have,

$$P(\tau \mid n_e) = \mathcal{N}[\underbrace{\Delta_* K K^{-1}(n_e - \bar{n}_e)}_{\Delta_*(n_e - \bar{n}_e)}, \underbrace{\Delta_* K \Delta_* - \Delta_* K K^{-1} K \Delta_*}_{0}]. \tag{2.31}$$

Similarly, the second distribution on the LHS is,

$$P(n_e \mid \tau_{\text{obs}}) = \mathcal{N}[\bar{n}_e + K\Delta^T(\Delta K\Delta^T + \sigma^2 I)^{-1}\tau_{\text{obs}}, \\ K - K\Delta^T(\Delta K\Delta^T + \sigma^2 I)^{-1}\Delta K]. \tag{2.32}$$

We now can use standard Gaussian identities (e.g. Weiss & Freeman, 2001) to evaluate the integral on the LHS:

$$\int P(\tau \mid n_e)\, P(n_e \mid \tau_{\text{obs}})\ \mathrm{d}n_e = \mathcal{N}[\Delta_* K\Delta^T(\Delta K\Delta^T + \sigma^2 I)^{-1}\tau_{\text{obs}}, \Delta_* K\Delta_*^T \\ - \Delta_* K\Delta^T(\Delta K\Delta^T + \sigma^2 I)^{-1}\Delta K\Delta_*^T]. \tag{2.33}$$

In order to work out the right-hand side (RHS), we simply condition Eq. 2.30 on τ_{obs} and then marginalise n_e by selecting the corresponding sub-block of the Gaussian,

$$P(n_e, \tau \mid \tau_{\text{obs}}) = \mathcal{N}\left[\begin{pmatrix}\bar{n}_e \\ 0\end{pmatrix} + \begin{pmatrix}K\Delta^T \\ \Delta_*^T K\Delta^T\end{pmatrix}\left(\Delta K\Delta^T + \sigma^2 I\right)^{-1}\tau_{\text{obs}}, \right. \\ \begin{pmatrix}\bar{K} & \Delta K\Delta_*^T \\ \Delta_* K\Delta^T & \Delta_* K\Delta_*^T\end{pmatrix} \\ \left. - \begin{pmatrix}K\Delta^T \\ \Delta_* K\Delta^T\end{pmatrix}\left(\Delta K\Delta^T + \sigma^2 I\right)^{-1}\begin{pmatrix}\Delta K & \Delta K\Delta_*^T\end{pmatrix}\right] \tag{2.34}$$

Marginalising over n_e is equivalent to neglecting the sub-block corresponding to n_e. Therefore, the RHS is

$$\int P(n_e, \tau \mid \tau_{\text{obs}})\ \mathrm{d}n_e = \mathcal{N}\left[\Delta_* K\Delta^T\left(\Delta K\Delta^T + \sigma^2 I\right)^{-1}\tau_{\text{obs}}, \Delta_* K\Delta_*^T \right. \\ \left. - \Delta_* K\Delta^T\left(\Delta K\Delta^T + \sigma^2 I\right)^{-1}\Delta K\Delta_*^T\right]. \tag{2.35}$$

■

2.A2 Derivation of the ΔTEC variance function and its limits

We derive the ΔTEC variance function $\sigma^2_{\Delta\text{TEC}}(d)$ for zenith observations ($k = k' = \hat{z}$) by considering a baseline between an antenna-of-interest at $\mathbf{x}_i = \mathbf{x}_j$ and a reference antenna at $\mathbf{x}_0 = 0$. To use the Pythagorean theorem later, we assume that this baseline lies in the plane of the local horizon, i.e. perpendicular to the zenith. Without loss of generality, we can orient the coordinate axes such that this baseline lies along the $\hat{\mathbf{x}}$

direction, so that $\mathbf{x}_i - \mathbf{x}_0 = d\hat{\mathbf{x}}$. Here $d \triangleq ||\mathbf{x}_i||$ is the distance between the two antennae. We then take the general covariance function $K_{\Delta\text{TEC}}\left(\left[\mathbf{x}_i, \mathbf{x}_0, \hat{\mathbf{k}}\right], \left[\mathbf{x}_j, \mathbf{x}_0, \hat{\mathbf{k}}'\right]\right)$, and find that in this particular case

$$\begin{aligned}\sigma^2_{\Delta\text{TEC}}(d) &\triangleq K_{\Delta\text{TEC}}\left(\left[\mathbf{x}_i, \mathbf{x}_0, \hat{\mathbf{z}}\right], \left[\mathbf{x}_i, \mathbf{x}_0, \hat{\mathbf{z}}\right]\right) \qquad (2.36)\\ &= \sum_{p_1=0}^{1}\sum_{p_2=0}^{1}(-1)^{p_1+p_2}\\ &\quad \int_0^b\int_0^b K_{n_e}\left(||\mathbf{x}_{(1-p_1)i} - \mathbf{x}_{(1-p_2)i} + \hat{\mathbf{z}}\left(s_1 - s_2\right)||\right)\,\mathrm{d}s_1\mathrm{d}s_2, \qquad (2.37)\end{aligned}$$

where K_{n_e} is an arbitrary stationary and isotropic kernel (such as the exponentiated quadratic and Matérn $\frac{3}{2}$ kernels considered earlier) for the FED. The two terms where p_1 and p_2 are equal give the same contribution, as do the two terms for which p_1 and p_2 are unequal. By subsequently applying the Pythagorean theorem in this last case (i.e. $p_1 = 0$ and $p_2 = 1$, and vice versa), we find

$$\sigma^2_{\Delta\text{TEC}}(d) = 2\int_0^b\int_0^b K_{n_e}\left(|s_1 - s_2|\right) - K_{n_e}\left(\sqrt{d^2 + \left(s_1 - s_2\right)^2}\right)\mathrm{d}s_1\mathrm{d}s_2. \qquad (2.38)$$

We manipulate this result to obtain a more insightful expression. First, we note the (implicit) presence of three parameters with dimension length: ionospheric thickness b, reference antenna distance d, and FED kernel half-peak distance h. We perform transformations to dimensionless coordinates $u_1 = \frac{s_1}{h}$ and $u_2 = \frac{s_2}{h}$ to reveal that the *shape* — though not the absolute scale — of the function $\sigma^2_{\Delta\text{TEC}}(d)$ is governed only by the length-scale ratios $\frac{b}{h}$ and $\frac{d}{h}$, and the particular functional form of K_{n_e}. Furthermore, for stationary covariance functions, we have $K_{n_e} = \sigma^2_{n_e} C_{n_e}$, where C_{n_e} is the corresponding dimensionless correlation function.

These considerations enable us to express the ΔTEC structure function as a dimensionless, shape-determining double integral appended by dimensionful prefactors; i.e.

$$\begin{aligned}\sigma^2_{\Delta\text{TEC}}(d) = 2\sigma^2_{n_e}h^2\int_0^{\frac{b}{h}}\int_0^{\frac{b}{h}} &C_{n_e}\left(h\left|u_1 - u_2\right|\right)\\ &- C_{n_e}\left(h\sqrt{\left(\frac{d}{h}\right)^2 + \left(u_1 - u_2\right)^2}\right)\mathrm{d}u_1\mathrm{d}u_2. \qquad (2.39)\end{aligned}$$

We first note that the variance of ΔTEC is simply proportional to the variance of n_e.

Secondly, we note that

$$h\left|u_1 - u_2\right| < h\sqrt{\left(\frac{d}{h}\right)^2 + \left(u_1 - u_2\right)^2} \tag{2.40}$$

for any non-zero d, so that

$$C_{n_e}\left(h\left|u_1 - u_2\right|\right) > C_{n_e}\left(h\sqrt{\left(\frac{d}{h}\right)^2 + \left(u_1 - u_2\right)^2}\right) \tag{2.41}$$

for all monotonically decreasing correlation functions C_{n_e} (or, equivalently, covariance functions K_{n_e}). With the integrand always positive, we see that the *integral* must be a strictly increasing function of $\frac{b}{h}$ (which occurs in the integration limits). Therefore, we conclude that for stationary, isotropic, and monotonically decreasing (SIMD) FED kernels with HPD h, the ΔTEC variance increases monotonically with the thickness of the ionosphere b. Simply put: thicker SIMD ionospheres cause larger ΔTEC variations.

Let us now consider three limits of the ΔTEC zenith variance function, that all do not require K_{FED} to decrease monotonically. In the short-baseline limit, i.e. $\frac{d}{h} \to 0$, we have

$$C_{n_e}\left(h\sqrt{\left(\frac{d}{h}\right)^2 + \left(u_1 - u_2\right)^2}\right) \to C_{n_e}\left(h\left|u_1 - u_2\right|\right). \tag{2.42}$$

We therefore find that $\sigma^2_{\Delta\mathrm{TEC}} \to 0$ irrespective of other parameters, recovering that the variance of ΔTEC vanishes near the reference antenna. In the long-baseline limit, i.e. $\frac{d}{h} \gg \frac{b}{h} > 1$, we see that $\sqrt{\left(\frac{d}{h}\right)^2 + \left(u_1 - u_2\right)^2} \approx \frac{d}{h}$, since

$$\left(u_1 - u_2\right)^2 < \left(\frac{b}{h}\right)^2 \ll \left(\frac{d}{h}\right)^2. \tag{2.43}$$

Assuming $C_{n_e}(d) \approx 0$ when $\frac{d}{h} \gg 1$, the integrand reduces to $C_{n_e}\left(h\left|u_1 - u_2\right|\right) - C_{n_e}\left(h \cdot \frac{d}{h}\right) \approx C_{n_e}\left(h\left|u_1 - u_2\right|\right)$. We find that in this case,

$$\sigma^2_{\Delta\mathrm{TEC}} \approx 2\sigma^2_{n_e} h^2 \int_0^{\frac{b}{h}} \int_0^{\frac{b}{h}} C_{n_e}\left(h\left|u_1 - u_2\right|\right) \mathrm{d}u_1 \mathrm{d}u_2. \tag{2.44}$$

This is the plateau value of the ΔTEC variance that our model predicts for the long-baseline limit.

Another way to arrive at the plateau value expression of Equation 2.44 is by considering the statistical properties of TEC first. In a computation analogous to the one for ΔTEC in Section 2.3, one can derive the general TEC covariance function K_{TEC}. The variance of $\tau_i^{\hat{z}}$ (the TEC of antenna i while observing towards the zenith $\hat{z}$) is straightforwardly shown to be

$$\mathbb{V}\left(\tau_i^{\hat{z}}\right) = \sigma_{n_e}^2 h^2 \int_0^{\frac{b}{h}} \int_0^{\frac{b}{h}} C_{n_e}\left(h\left|u_1 - u_2\right|\right) \mathrm{d}u_1 \mathrm{d}u_2. \tag{2.45}$$

We highlight the absence of a dependence on i at the RHS. As a ΔTEC is simply a TEC differenced with a TEC for a reference antenna observing in the same direction, we have

$$\sigma_{\Delta\mathrm{TEC}}^2 = \mathbb{V}\left(\tau_i^{\hat{z}} - \tau_0^{\hat{z}}\right) = \mathbb{V}\left(\tau_i^{\hat{z}}\right) + \mathbb{V}\left(\tau_0^{\hat{z}}\right), \tag{2.46}$$

where the second equality only holds when the TECs are independent. This is exactly the scenario considered in the long-baseline limit. Plugging in Equation 2.45 recovers the plateau level. We can find a general upper bound to the variance of ΔTEC in terms of physical parameters. To this end, we note that the integrand in Equation 2.39 is maximised when, over the full range of integration, the value of the first term is 1 whilst the second term is equal to the infimum of the correlation function. Calling $\inf_{\mathbb{R}}\left\{C_{n_e}(r) : r \in \mathbb{R}_{>0}\right\} \triangleq I$, we find the inequality,

$$\sigma_{\Delta\mathrm{TEC}}^2 \leq 2\sigma_{n_e}^2 h^2 \int_0^{\frac{b}{h}} \int_0^{\frac{b}{h}} 1 - I \,\mathrm{d}u_1 \mathrm{d}u_2 = 2\left(1 - I\right)\sigma_{n_e}^2 b^2. \tag{2.47}$$

For strictly positive FED kernels that decay to zero at large distances (such as the EQ and Matérn kernels considered in this work), we find $\sigma_{\Delta\mathrm{TEC}}^2 \leq 2\sigma_{n_e}^2 b^2$. Kernels resulting in anticorrelated FEDs produce the constraint $\sigma_{\Delta\mathrm{TEC}}^2 \leq 4\sigma_{n_e}^2 b^2$ or tighter. By measuring $\sigma_{\Delta\mathrm{TEC}}(d)$, one can bound the product $\sigma_{n_e} b$ from below. The strongest bound is obtained for large d.

2.A3 Covariance function with frozen flow

Here we give the DTEC covariance function assuming the ionosphere drifts over the array as time progresses, without evolving internally. We derived these results during the development of the original publication, but they were not made part of it.

The covariance between the DTEC RV of a station at $\mathbf{x}_{i_1}$ referenced to one at $\mathbf{x}_{j_1}$

whilst both observe $\hat{k}_{\alpha_1}$ at time t_{β_1}, and the DTEC RV of a station at $\mathbf{x}_{i_2}$ referenced to one at $\mathbf{x}_{j_2}$ whilst both observe $\hat{k}_{\alpha_2}$ at time t_{β_2}, is the sum of four double integrals:

$$K_{\Delta\tau}\left(\begin{bmatrix}\mathbf{x}_{i_1}\\ \mathbf{x}_{j_1}\\ \hat{k}_{\alpha_1}\\ t_{\beta_1}\end{bmatrix}, \begin{bmatrix}\mathbf{x}_{i_2}\\ \mathbf{x}_{j_2}\\ \hat{k}_{\alpha_2}\\ t_{\beta_2}\end{bmatrix}\right) = \sum_{l_1=0}^{1}\sum_{l_2=0}^{1}(-1)^{l_1+l_2}\int_{s_1^-}^{s_1^+}\int_{s_2^-}^{s_2^+} K_{n_e}(\mathbf{y}_1, \mathbf{y}_2)\ \mathrm{d}s_1\mathrm{d}s_2, \quad (2.48)$$

where typically $j_1 = j_2$, K_{n_e} is the FED covariance function, and

$$s_p^{\pm} = \sec\varphi_{\alpha_p}\left(a \pm \frac{b}{2} - \left(\mathbf{x}_{(1-l_p)i_p+l_pj_p} - \mathbf{x}_0 - \mathbf{v}t_{\beta_p}\right)\cdot\hat{z}\right), \quad (2.49)$$

$$\mathbf{y}_p = \mathbf{x}_{(1-l_p)i_p+l_pj_p} - \mathbf{v}t_{\beta_p} + \hat{k}_{\alpha_p}s_p, \quad (2.50)$$

$$\sec\varphi_{\alpha_p} = \left(\hat{k}_{\alpha_p}\cdot\hat{z}\right)^{-1}. \quad (2.51)$$

We model the ionosphere as a layer with thickness b, whose centre lies at a height a above $\mathbf{x}_0$. The zenith is denoted $\hat{z}$. We link DTEC RVs of different times by assuming that the ionosphere is a frozen flow with velocity $\mathbf{v}$.

2.A4 Supporting figures

Figure 2.8 shows 24 Dutch *core* and *remote* LOFAR stations with a simulated ionosphere above them. Figure 2.9 shows simulated ionospheric phase distortions for the same stations. For brevity, these figures did not feature in the original publication; they appear for the first time in this thesis, and illustrate the methods of Sect. 2.4.

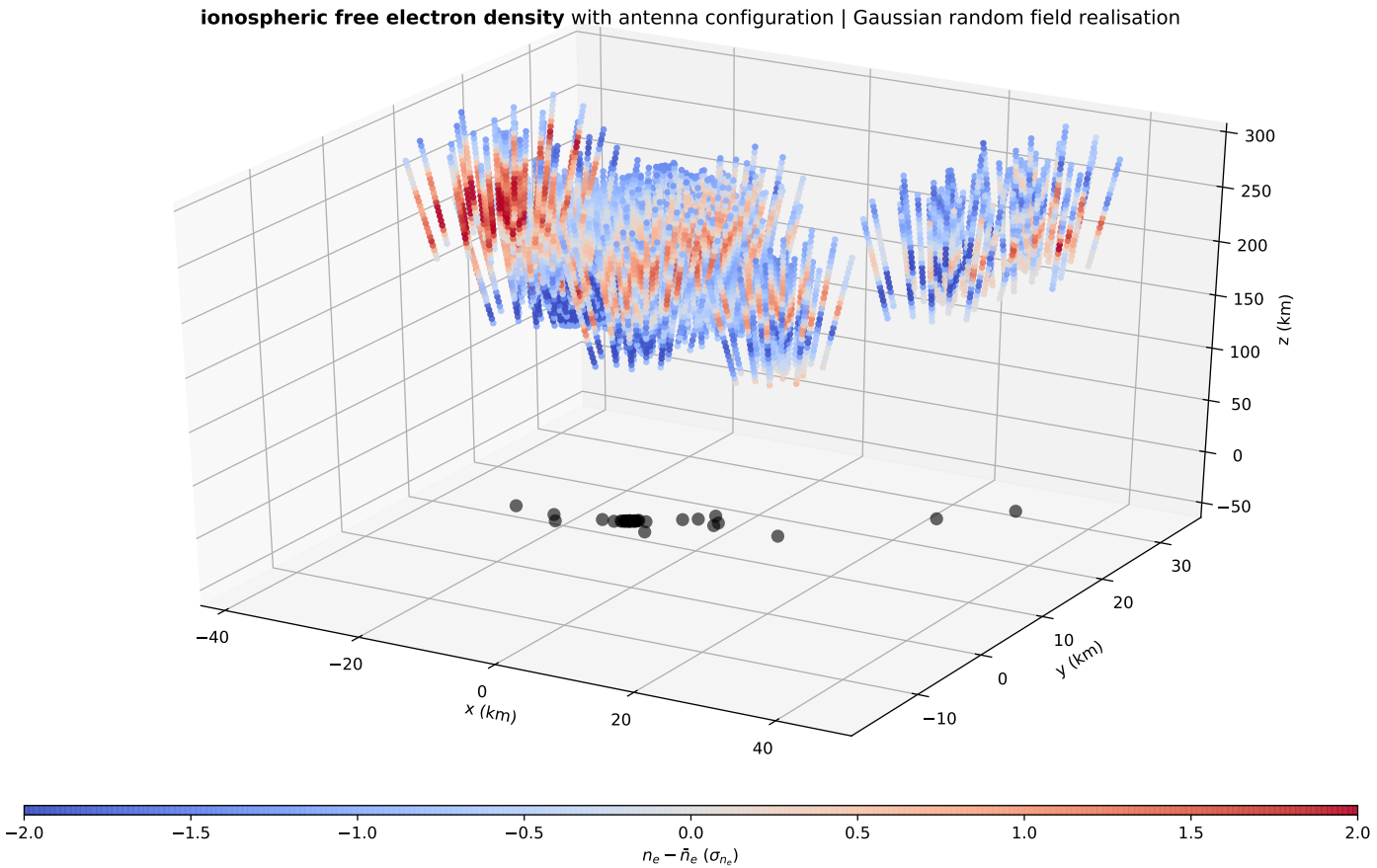


Figure 2.8: Simulation of a three-dimensional Gaussian random field ionosphere above 24 Dutch LOFAR stations (black dots), for which our work presents the optimal calibration method. We show the field $n_e(\mathbf{x}, t)$ along a few sightlines only. The visualisation makes clear that sightlines from the same station pass through similar ionospheric columns, resulting in correlated ΔTECs and phase distortions. For the latter, see Fig. 2.9.

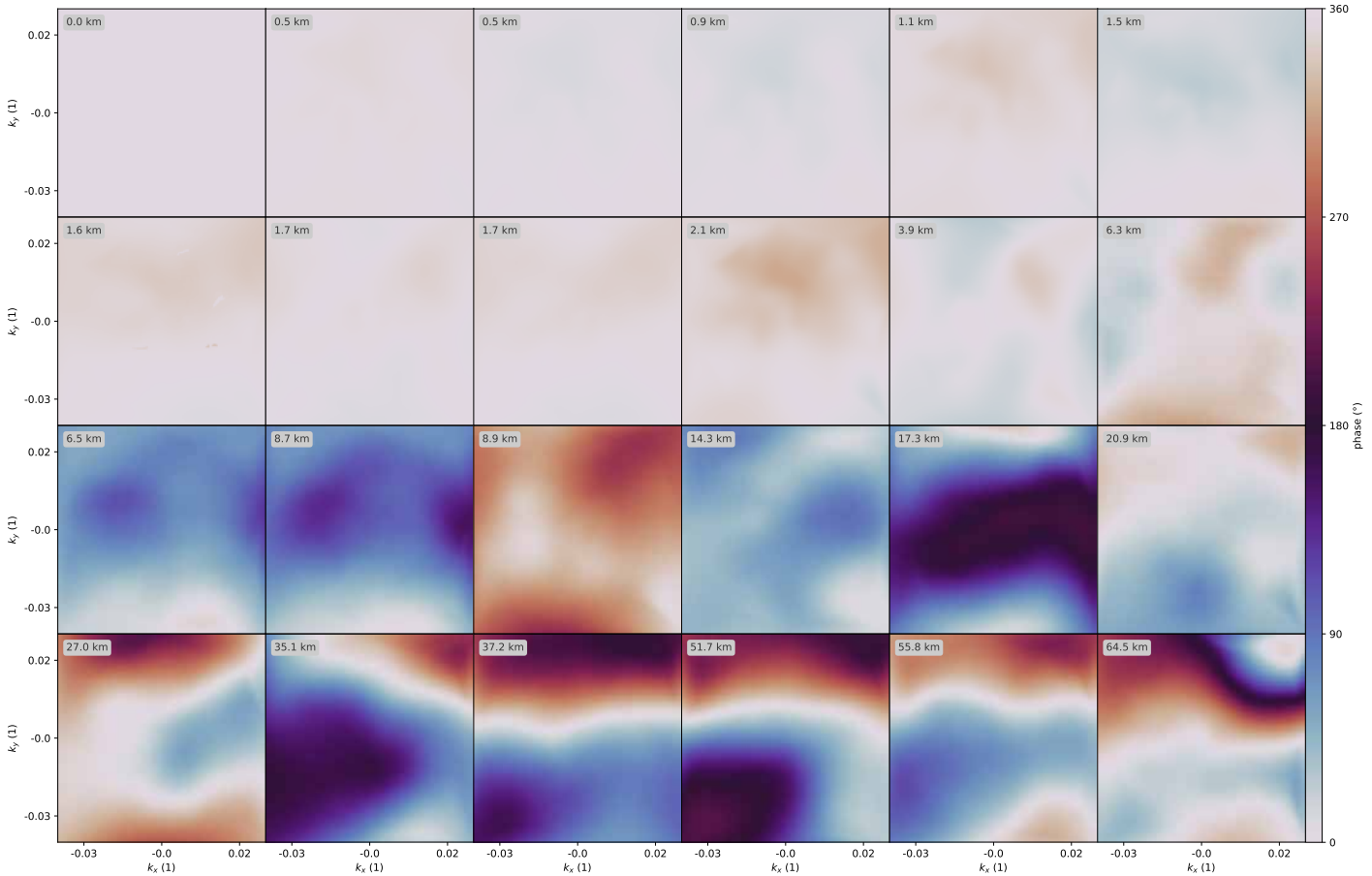


Figure 2.9: Ionospheric phase distortions as a function of sky direction (k_x, k_y) for each station shown in Fig. 2.8, with array centre proximity (legends) marked. Statistically speaking, the further the stations are from the array centre, the larger the phase distortions become. The phase distortions inherit this behaviour from the ΔTEC fields shown in Fig. 2.2, with which the distortions scale linearly.

What I cannot create, I do not understand.

Richard P. Feynman, American physicist, on his Caltech blackboard at the time of his passing (1988)

3

Filamentary Baryons and Where to Find Them: A forecast of synchrotron radiation from merger and accretion shocks in the local Cosmic Web

M. S. S. L. Oei, R. J. van Weeren, F. Vazza, F. Leclercq, A. Gopinath, H. J. A. Röttgering — *Astronomy & Astrophysics*, 662, 87, 2022

Abstract

Context The detection of synchrotron radiation from the intergalactic medium (IGM) that pervades the filaments of the Cosmic Web constitutes an upcoming frontier to test physical models of astrophysical shocks and their radiation mechanisms, trace the missing baryons, and constrain magnetogenesis — the origin and evolution of extragalactic magnetic fields.

Aims The first synchrotron detections of the IGM within filaments have recently been claimed. Now is the time to develop a rigorous statistical framework to predict sky regions with the strongest signal and to move from mere detection to inference, that is to say identifying the most plausible physical models and parameter values from observations.

METHODS Current theory posits that the filament IGM lights up through shocks that originate from large-scale structure formation. With Bayesian inference, we generated a probability distribution on the set of specific intensity functions that represent our view of the merger- and accretion-shocked synchrotron Cosmic Web (MASSCW). We combined the Bayesian Origin Reconstruction from Galaxies (BORG) Sloan Digital Sky Survey (SDSS) total matter density posterior, which is based on spectroscopic observations of galaxies within SDSS DR7, snapshots of Enzo magnetohydrodynamics (MHD) cosmological simulations, a Gaussian random field (GRF), and a ray tracing approach to arrive at the result.

RESULTS We present a physics-based prediction of the MASSCW signal, including principled uncertainty quantification, for a quarter of the sky and up to cosmological redshift $z_{\max} = 0.2$. The super-Mpc 3D resolution of the current implementation limits the resolution of the predicted 2D imagery, so that individual merger and accretion shocks are not resolved. The MASSCW prior can be used to identify the most promising fields to target with low-frequency radio telescopes and to conduct actual detection experiments. We furthermore calculated a probability distribution for the flux density–weighted mean (i.e. sky-averaged) redshift $\bar{\bar{z}}$ of the MASSCW signal up to $z_{\max}$, and found a median of $\bar{\bar{z}} = 0.077$. We constructed a low-parametric analytic model that produces a similar distribution for $\bar{\bar{z}}$, with a median of $\bar{\bar{z}} = 0.072$. Extrapolating the model, we were able to calculate $\bar{\bar{z}}$ for all large-scale structure in the Universe (including what lies beyond $z_{\max}$) and show that, if one only considers filaments, $\bar{\bar{z}}$ depends on virtually one parameter. As case studies, we finally explore the predictions of our MASSCW specific intensity function prior in the vicinity of three galaxy clusters, the Hercules Cluster, the Coma Cluster, and Abell 2199, and in three deep Low-frequency Array (LOFAR) High-band Antennae (HBA) fields, the Lockman Hole, Abell 2255, and the Ursa Major Supercluster.

CONCLUSIONS We describe and implement a novel, flexible, and principled framework for predicting the low-frequency, low-resolution specific intensity function of the Cosmic Web due to merger and accretion shocks that arise during large-scale structure formation. The predictions guide Local Universe searches for filamentary baryons through half of the Northern Sky. Once cosmological simulations of alternative emission mechanisms have matured, our approach can be extended to predict additional physical pathways that contribute to the elusive synchrotron Cosmic Web signal.

Key words: Cosmology: miscellaneous – dark matter – large-scale structure of Universe – Galaxies: clusters: intracluster medium – intergalactic medium – Magnetic

fields – Magnetohydrodynamics (MHD) – Methods: numerical – statistical – Radiation mechanisms: non-thermal – Radio continuum: general – Shock waves

3.1 Introduction

Just after inflation, the Universe' s dark and baryonic matter density functions resembled realisations of nearly constant, isotropic, and stationary Gaussian random fields (GRFs) (Linde, 2008). Due to gravity alone, these fields evolved into the highly inhomogeneous, network-like large-scale structure (LSS) present today (Springel et al., 2005). The late-time Universe consists of two components: voids and the Cosmic Web, which can be further partitioned into sheets, filaments, and (galaxy) clusters.[1] The initial density conditions, through the morphology of the Cosmic Web, determine the spatial distribution of galaxies and some of their internal properties, such as the magnetic field and spin. Hahn et al. (2007), inspired by the seminal work of Zel'dovich (1970), provided a[2] rigorous definition (known as the 'T-web') of the four canonical structure types (voids, sheets, filaments, and clusters) based on the number of positive eigenvalues (0, 1, 2, or 3, respectively) of the tidal field tensor. Their N-body simulations reveal that the prevalence of the structure types (quantified by e.g. volume-filling factors) is evolving, with filaments and sheets disappearing in favour of clusters and voids at the present day.[3]

Of the four structure types, galaxy clusters are most easily studied: in the X-ray, optical and radio bands, for example through thermal bremsstrahlung, gravitational lensing and synchrotron radiation. As the most massive gravitationally bound structures in the Universe thus far, they weigh up to $\sim 10^{15}\ M_{\odot}$, contain up to $\sim 10^{3}$ galaxies, and are pervaded by a dilute ($\sim 10^{3}$ m^{-3}), hot ($\sim 10^{8}$ K) and magnetised ($\sim 1\ \mu$G) hydrogen- and helium-dominated plasma: the intra-cluster medium (ICM) (Cavaliere & Rephaeli, 2011).

At the outskirts of clusters, the ICM transitions into the warm–hot intergalactic medium (WHIM) — the dominant baryonic constituent of filaments. The WHIM is

[1] Some authors use 'Cosmic Web' to refer to filaments exclusively, but in this work, the term is used to refer to all of the late-time Universe excluding the voids. We furthermore differentiate between 'cosmic web' (for the concept in an arbitrary universe), and 'Cosmic Web' (for the concept in ours).

[2] Many other cosmic web classifiers exist, such as DIVA (Lavaux & Wandelt, 2010), V-web (Hoffman et al., 2012), and LICH (Leclercq et al., 2017).

[3] Forero-Romero et al. (2009) subsequently refined Hahn's definition, by counting eigenvalues above a tuneable threshold (rather than zero) related to the gravitational collapse timescale. The partitioning of the cosmic web into the four structure types depends sensitively on the choice of this threshold, and it should therefore always be mentioned when quantitative structure-type properties are stated, such as volume- and mass-filling fractions.

a plasma of nearly primordial chemical composition, but is less dense ($\sim 1 - 10\ \mathrm{m}^{-3}$), cooler ($\sim 10^5 - 10^7$ K) and less magnetised ($\sim 10^{-3} - 10^{-1}$ μG) than the ICM. Compared to clusters, filaments are therefore harder to detect in all three wavelength bands. Despite this, filaments are cosmologically relevant, as simulations predict that the WHIM harbours up to $\sim$ 90% of the Universe's baryons (Cen & Ostriker, 1999; Eckert et al., 2015; de Graaff et al., 2019; Tanimura et al., 2019a).

The formation of filaments (and, indirectly, the galaxy clusters they fuel) has occurred primarily through the influx of dark matter (DM) and cold gas from sheets and voids. Once these free-falling pockets of gas reach a filament's surface, they generate supersonic accretion shocks, with upstream Mach numbers $\mathcal{M}_\mathrm{u} \sim 10^0 - 10^2$ (Ryu et al., 2003). Following Ensslin et al. (1998); Miniati et al. (2001), Hoeft & Brüggen (2007) have proposed that *merger* shocks in *clusters* boost electrons in the high-energy tail of the ICM's Maxwell–Boltzmann distribution to ultrarelativistic velocities via diffusive shock acceleration (DSA). By extension, the Hoeft & Brüggen (2007) model could also describe (again via DSA) how *accretion* shocks in *filaments* boost electrons in the high-energy tail of the WHIM's Maxwell–Boltzmann distribution to ultrarelativistic velocities. The DSA mechanism details how charges diffuse back and forth across the shock front, trapped in a magnetic mirror, and gain speed accordingly (Krymskii, 1977; Axford et al., 1977; Bell, 1978a,b; Blandford & Ostriker, 1978; Drury, 1983; Blandford & Eichler, 1987; Jones & Ellison, 1991; Baring, 1997; Malkov & Drury, 2001). Once released, these high-energy electrons subsequently spiral along the magnetic field lines of the intergalactic medium (IGM), glowing in synchrotron light. It is thought that this accretion-shock-based radiation mechanism provides the dominant contribution to the filaments' synchrotron Cosmic Web (SCW) signal.

Although the Mach numbers of accretion shocks in filaments are higher than those of merger shocks in clusters (where they are $\mathcal{M}_\mathrm{u} \sim 1 - 5$) (Ryu et al., 2003), shocks in filaments remain fainter due to the aforementioned adverse density, temperature and magnetic field strength conditions.[4] This is why observing filaments through synchrotron radiation constitutes an ambitious, futuristic frontier.

Modern radio telescopes, such as the upgraded Giant Metrewave Radio Telescope (uGMRT), the Expanded Very Large Array (EVLA) and the Low-frequency Array High-band Antennae (LOFAR HBA), have enabled detailed studies of particle acceleration in the cluster IGM (e.g. Di Gennaro et al. (2018); Kale (2020); Locatelli et al. (2020); Mandal et al. (2020)) by the detection of radio halos, phoenices and

[4]Precisely how much fainter shocks in filaments are compared to those in clusters, is unknown, because the typical filament IGM magnetic field strength B_IGM and the electron acceleration efficiency ξ_e of shocks — and especially weak ones — is highly uncertain.

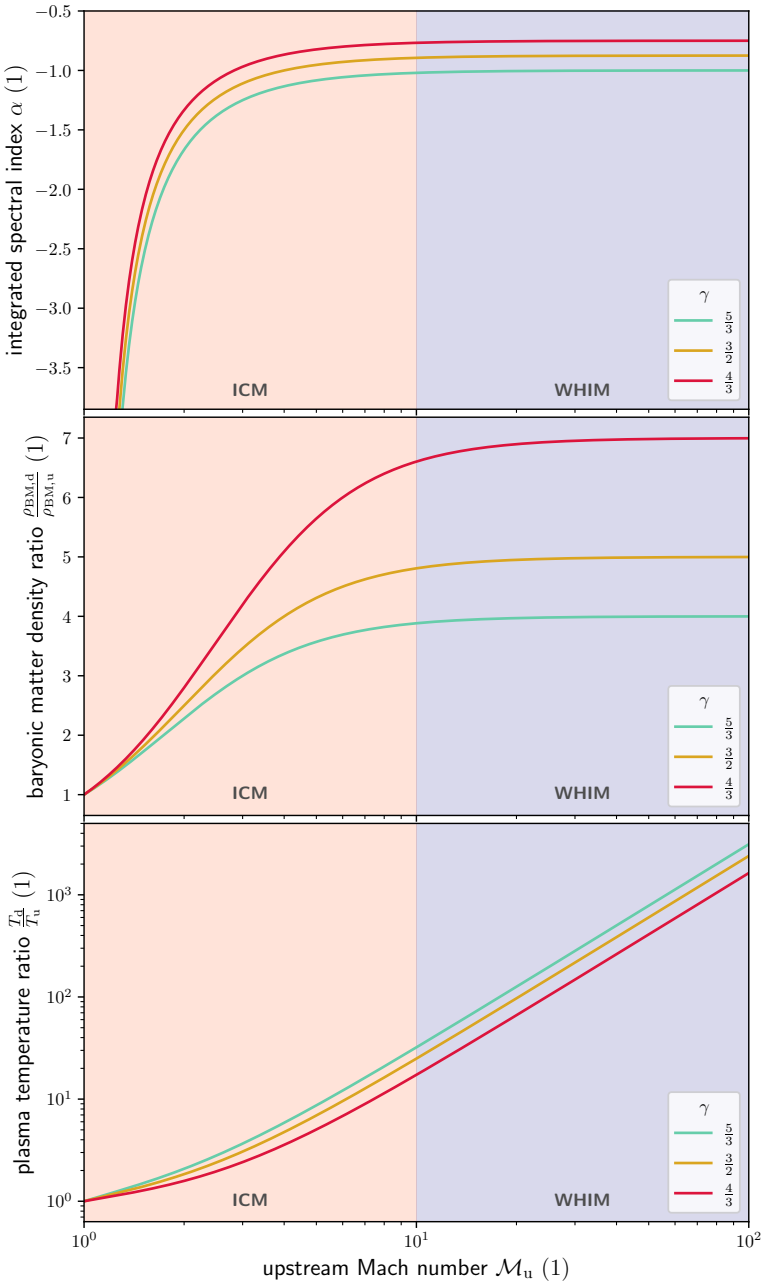


Figure 3.1: Dependence of three central (merger or accretion) shock quantities on upstream Mach number $\mathcal{M}_{\mathrm{u}}$ and adiabatic index γ, as derived from the Rankine–Hugoniot jump conditions in ideal gases. Radiation from $\mathcal{M}_{\mathrm{u}} < 10$ shocks is predominantly generated by electrons that stem from the ICM, whilst radiation from $\mathcal{M}_{\mathrm{u}} > 10$ shocks is predominantly generated by electrons that stem from the WHIM. *Top:* the spectral index α of angularly unresolved shocks, assuming standard DSA. *Middle:* the downstream-over-upstream plasma density ratio. *Bottom:* the downstream-over-upstream plasma temperature ratio.

relics. In contrast, no single shock in filaments has hitherto been observed. Doing so would open up density, temperature, Mach number and magnetic field strength regimes different by orders of magnitude via which astrophysical shock models could be held to the test. For example, the top panel of Fig. 3.1 shows that DSA predicts an (angularly unresolved) synchrotron spectral index $\alpha = -1$ for virtually all filament shocks (with a slight dependency on the adiabatic index γ). Significant deviations from $\alpha = -1$ would falsify standard DSA.[5] A better understanding of astrophysical shocks has ramifications beyond the study of large-scale structure, as possibly similar shocks are found in, amongst other places, accreting X-ray binaries, stellar and pulsar winds, and supernova remnants.

Routine detection of the filament IGM in synchrotron would provide a novel way to address the missing baryon problem — the possible discrepancy between today's mean baryon density as inferred from galaxy surveys versus that predicted by CMB-constrained ΛCDM models.

Direct imaging of the filament IGM in the low-frequency radio window adds spectral diversity to a growing list of methodologies that trace the WHIM, complementing X-ray observations of ionised oxygen (O VII) absorption along the line of sight to quasars (Nicastro et al., 2018), microwave measurements of the thermal Sunyaev–Zel'dovich effect due to hot gas between adjacent galaxies (de Graaff et al., 2019), millimetre searches for the hyperfine spin-flip transition of single-electron nitrogen ions (N VII) (Bregman & Irwin, 2007) and dispersion measurements of localised fast radio bursts (FRBs) (Macquart et al., 2020). Just as direct imaging of the WHIM in the X-ray band (Eckert et al., 2015), detecting baryons through synchrotron emission does not necessitate a special (line-of-sight) geometry. Telescopes such as the LOFAR could thus corroborate the current baryon census, in which still ∼34% is not identified conclusively (de Graaff et al., 2019). Thirdly, the low galaxy number density in filaments means that its IGM, if largely untouched by galactic feedback, can retain information on its initial conditions for billions of years. In particular, MHD simulations (Vazza et al., 2015, 2017) demonstrate that different assumptions for the dominant physical process that drove the growth of magnetism in filaments, lead to different strengths and morphologies of the IGM's magnetic fields today. The authors evaluate three different scenarios for cosmic magnetogenesis: primordial, dynamo, and galactic models, and calculate their evolution from a cosmological redshift of 38 to the present day. The models are calibrated by magnetic field strength measurements of the ICM. In primordial models, seed field fluctuations grow as LSS forma-

[5] Caprioli & Haggerty (2019) describe recent advances in DSA theory, which suggest steeper electron energy and synchrotron spectra for the filaments' strong shocks (i.e. $\alpha < -1$ for $\gamma = 5/3$ and $\mathcal{M}_u > 10$).

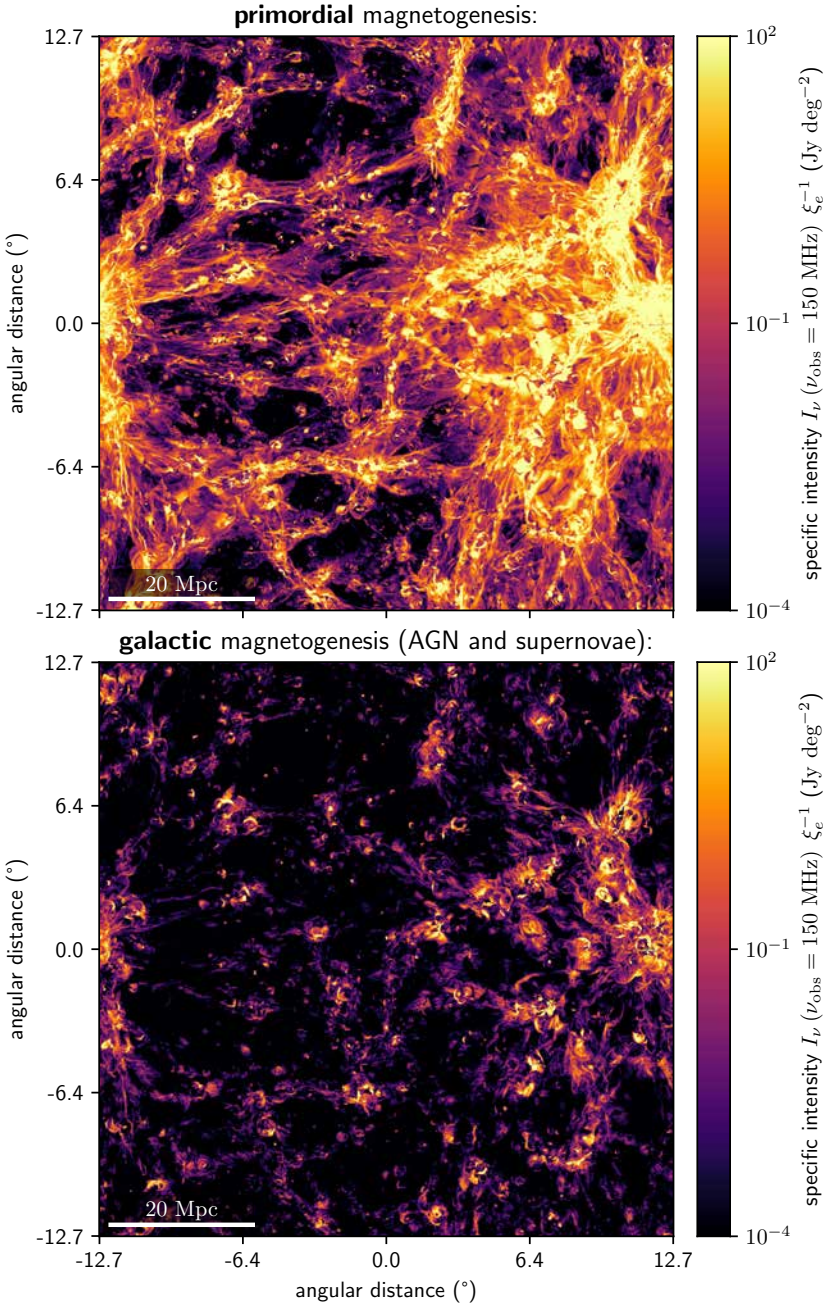


Figure 3.2: Two simulated specific intensity functions at $\nu_{\rm obs} = 150$ MHz, assuming synchrotron radiation from merger and accretion shocks (as in Hoeft & Brüggen (2007)) in LSS at $z = 0.045$. *Top:* primordial scenario for magnetogenesis, starting from $B_{\rm IGM} = 1$ nG in the Early Universe. (In contrast, the simulation underlying this article's predictions starts at $B_{\rm IGM} = 0.1$ nG.) *Bottom:* galactic scenario for magnetogenesis, in which seeding occurs through AGN outflows and supernova winds. See Gheller & Vazza (2020) for details on these scenarios.

tion compresses and rarefies the magnetic field lines. In dynamo models, the seed field is much weaker (10^{-22} T vs. 10^{-13} T), but grows in strength as energy in solenoidal, turbulent gas motion is converted into magnetic energy through small-scale dynamos (Ryu, 2008). In galactic (or 'astrophysical') models, finally, no seed field is assumed, with star formation and outflows from the jets of supermassive black holes (SMBHs) at galactic centres being the dominant contributors to the emergence of the magnetised cosmic web. Fig. 3.2 illustrates that different magnetisation scenarios give rise to morphologically different low-frequency specific intensity functions. Thus, SCW observations could rule out at least some models of magnetogenesis — a puzzle widely considered to be amongst the most significant open problems in cosmology.

Hitherto, several groups have constrained the magnetic field strength of the filament IGM by means of Stokes I low-frequency radio observations. Using Murchison Widefield Array (MWA) Epoch of Reionisation Field 0 data and the equipartition energy condition, Vernstrom et al. (2017) derived a parameter-dependent upper limit of $B_{\mathrm{IGM}} < 0.03 - 1.98$ μG, while Brown et al. (2017) used Parkes 64m Telescope S-PASS data to find a density-weighted upper limit of $B_{\mathrm{IGM}} < 0.13$ μG at the present epoch. Finally, using MWA GLEAM and ROSAT RASS data, Vernstrom et al. (2021) suggest an average magnetic field strength of 30 nG $< B_{\mathrm{IGM}} <$ 60 nG, using both equipartition and inverse Compton arguments. In an upcoming publication (Oei et al., in prep.), we describe and present a LOFAR search, that includes new methodology.

Alternatively, O'Sullivan et al. (2019, 2020); Stuardi et al. (2020) constrain the properties of the IGM's magnetic fields via rotation measure (RM) synthesis applied to LOFAR Two-metre Sky Survey (LoTSS) observations of the lobes of (giant) radio galaxies, finding $B_{\mathrm{IGM}} < 4$ nG in filaments. Vernstrom et al. (2019) have performed a similar analysis with NVSS data, finding $B_{\mathrm{IGM}} < 40$ nG in filaments.

The goal of this article is, first and foremost, to explain a new method for SCW prediction, that yields a probability distribution over the set $\mathrm{Map}\,(\mathbb{S}^2, \mathbb{R}_{\geq 0})$ of specific intensity functions on the 2-sphere $\mathbb{S}^2$. The secondary goal is to demonstrate the method's potential, by showing results for the modern, nearby Universe ($z < 0.2$) over half of the Northern Hemisphere — 25% of the full sky.

In Sect. 3.2, we first provide a general outline of our SCW signal prediction method. We then detail the methodology step-by-step, providing a background to the input data as we proceed. In Sect. 3.3, we develop a simple geometric model of the cosmic web that yields SCW redshift predictions. In Sect. 3.4, we analyse predictions of the main method, and compare these with predictions of the geometric model. Finally, in Sect. 3.5, we discuss caveats of the current work, and give recommendations for future extensions, before we present conclusions in Sect. 3.6.

We adopt a[6] concordance inflationary ΛCDM cosmological model $\mathfrak{M} = (\Omega_{\Lambda,0} = 0.728, \Omega_{\mathrm{DM},0} = 0.227, \Omega_{\mathrm{BM},0} = 0.045, h = 0.702, \sigma_8 = 0.807, n_s = 0.961)$, so that $\Omega_{\mathrm{M},0} := \Omega_{\mathrm{DM},0} + \Omega_{\mathrm{BM},0} = 0.272$ and the Hubble constant $H_0 := h \cdot 100\ \mathrm{km\ s^{-1}\ Mpc^{-1}}$ (Jasche et al., 2015).

3.2 Methods

3.2.1 Overview

For structural clarity, we begin with an overview of our synchrotron Cosmic Web prediction approach.

1. Our starting point is a probability distribution over the total (i.e. baryonic and dark) matter density fields of the modern, nearby Universe.

2. Then, using a cosmological MHD simulation, we established the connection between the total matter density ρ and the merger and accretion shock synchrotron monochromatic emission coefficient (MEC) j_ν, in the form of the conditional probability distribution $P(j_\nu(\mathbf{r}) \mid \rho(\mathbf{r}))$.

3. Next, to create a MEC field, one could independently realise the random variables (RVs) $j_\nu(\mathbf{r}) \mid \rho(\mathbf{r})$ for all $\mathbf{r}$ in the volume of the inferred density fields. However, such an approach would disregard the fact that in reality, adjacent locations have similar physical conditions and thus similar MECs (i.e. the MEC field exhibits spatial correlation). To account for spatial correlation, we generated realisations of an isotropic, stationary, three-dimensional Gaussian random field (GRF) with zero mean and unit variance. By applying the cumulative density function (CDF) of the standard normal RV in point-wise fashion, the GRF transforms into a 'percentile random field' — with values strictly between 0 and 1. This new field inherits the spatial correlation present in the GRF. We then let the percentile random field determine the MEC field, by plugging the percentile scores into the inverse CDF of $j_\nu(\mathbf{r}) \mid \rho = \rho(\mathbf{r})$ for all $\mathbf{r}$ in the volume.

4. After having obtained a merger and accretion shock synchrotron MEC field with spatial correlation, we ray traced through this three-dimensional structure, taking into account effects due to the Universe's expansion. This, finally,

[6] Rather fortunately, the BORG SDSS and Enzo data products combined in this work have been made assuming almost identical cosmological parameters; the set reported is of the BORG SDSS. Authors of future work who strive to achieve perfect self-consistency should use the same cosmological model for the LSS reconstructions as for the MHD simulations.

yields a merger and accretion shock synchrotron Cosmic Web (MASSCW) specific intensity function on the sky.

5. By repeating the above steps over and over, picking different density field and GRF realisations every time, we generated a probability distribution on $\mathrm{Map}\left(\mathbb{S}^2, \mathbb{R}_{\geq 0}\right)$ for the MASSCW signal.

6. With a minor modification, this procedure naturally leads to another probability distribution on $\mathrm{Map}\left(\mathbb{S}^2, \mathbb{R}_{\geq 0}\right)$, which captures the direction-dependent mean cosmological redshift of the MASSCW.

3.2.2 Modern-day total matter density field posterior

A brief history of Bayesian large-scale structure reconstruction

A decade of research in large-scale structure reconstruction from galaxy surveys has culminated in a suite of highly principled, physics-based Bayesian inference techniques that unveil the content, dynamics and history of the local Cosmic Web. These techniques represent the state-of-the-art of LSS reconstruction, and their applications are manifold (see e.g. Jasche & Lavaux (2019)). The application relevant to our purposes is the reconstruction of the modern-day (i.e. $z = 0$) total (i.e. baryonic and dark) matter density field of the local Universe, in the form of a probability distribution over all possible density fields.

Several variations of the same general framework exist. Inspired by the map-level CMB inference methods of Wandelt et al. (2004), some theoretical groundwork common to all these techniques has been developed by Kitaura & Enßlin (2008), Jasche et al. (2010a) and Jasche & Kitaura (2010a), who also demonstrated the practical feasibility of Bayesian LSS reconstruction with the ARGO, ARES and HADES codes, respectively. In Jasche & Wandelt (2013), the authors described BORG (Bayesian Origin Reconstruction from Galaxies), the first algorithm to include physics of structure formation, and provided a proof-of-concept on simulated Sloan Digital Sky Survey Data Release 7 (SDSS DR7) data. In Jasche et al. (2015), the *actual* SDSS DR7 main galaxy sample was analysed with a refined version of BORG, yielding the publically-available BORG SDSS data release. These data form the basis for the SCW prediction method presented in this paper.

The BORG SDSS has hitherto been used for Cosmic Web classification (Leclercq et al., 2015), for the study of galaxy properties as a function of environment (Leclercq et al., 2016) and to unveil the dynamics of DM streams (Leclercq et al., 2017), all in a probabilistic and time-dependent fashion. Further refinements to BORG have been presented alongside applications to new datasets, namely 2M++ (Lavaux & Jasche,

2016; Jasche & Lavaux, 2019) and SDSS3-BOSS (Lavaux et al., 2019). Algorithms related to BORG are ELUCID (Wang et al., 2013, 2014), COSMIC BIRTH (Kitaura et al., 2021) and BARCODE (Bos et al., 2019).

Main principles

Only a brief description of the main ideas underlying these approaches falls within the scope of this paper. The methods all use the crucial insight that the statistical properties of the total density field in the Early Universe are well understood (i.e. Gaussian, with a theoretically predicted and observationally verified covariance function), and that the modern-day total density field relates to the initial field deterministically by means of (approximately Newtonian) gravity.[7] Thus, realisations from today's highly non-Gaussian total density field can be generated by forward modelling the effect of gravity on a collisionless[8] fluid, which is initialised as a Gaussian random field at the time of the CMB.[9] After a CMB-epoch total density GRF has been generated, and evolved into a modern total density field, a Poisson point process — with an intensity function that attempts to capture the relation between galaxy locations and the surrounding matter distribution — is used to calculate the likelihood of finding galaxies at their measured locations assuming the true underlying total density field is the one currently considered. As suggested before, spectroscopic galaxy surveys, with hundreds of thousands of galaxies pinpointed in three-dimensional comoving space, form the typical input data. Most notably, the likelihood for the modern total density field also fixes the likelihood for the initial total density field from which the modern one was evolved. Because the prior on the initial total density fields is Gaussian, the initial total density field can not only be assigned a likelihood, but also a posterior probability. Hamiltonian Monte Carlo (HMC) Markov Chains provide the means to explore the high-dimensional posterior distribution of initial total density fields consistent with the galaxy data. A posterior distribution over modern-day total density fields, which is used in this article, is generated as a by-product of this process.

[7] On super-Mpc scales, baryonic effects such as gas and radiation pressure play a subdominant role in structure formation, and can therefore be safely ignored.

[8] The assumption of a collisionless fluid is apt for DM, but only approximate for BM.

[9] Because simulating gravity with an N-body simulation from the CMB to the present day is computationally expensive, and the method requires this process to be repeated thousands of times, all authors resort to methods that approximate Newtonian gravitational evolution. In order of increasing accuracy (and numerical cost), they invoke either first-order Lagrangian perturbation theory (i.e. the Zel'dovich approximation (ZA; Zel'dovich, 1970)), second-order Lagrangian perturbation theory (2LPT) or particle mesh (PM) models, which approach the N-body solution.

BORG SDSS

The BORG SDSS modern-day total density field posterior used in this article is based on optical galaxy data from the Sloan Digital Sky Survey (SDSS; York et al., 2000; Strauss et al., 2002) Legacy Survey (Abazajian et al., 2009). This catalogue provides right ascensions φ, declinations θ, spectroscopic redshifts and Petrosian r-band apparent magnitudes m_r of about a million galaxies in the Northern Galactic Cap (NGC), which covers a region of 7646 deg^2 of Northern Sky, and three stripes of the Southern Galactic Cap (SGC), which together cover 386 deg^2. In these regions, the SDSS Legacy Survey is spectroscopically complete for galaxies with $m_r < 17.77$.

The BORG SDSS, which is based on 372,198 NGC galaxies (see Jasche et al. (2015) for selection details), infers structures in the density field of $\sim$3 Mpc h^{-1} and larger. Each realisation of the posterior is a cube containing 256^3 voxels, that represents a region $\mathcal{R} \subset \mathbb{R}^3$ with comoving volume $(750 \text{ Mpc h}^{-1})^3$. Naturally, $\mathcal{R}$ is chosen to correspond to the half of the Northern Sky that contains the NGC (up to a cosmological redshift $z \sim 0.2$).

3.2.3 Relating total matter density to synchrotron monochromatic emission coefficient

Next, to convert total matter density fields ρ to proper synchrotron monochromatic emission coefficient (MEC) fields j_ν, we established the relation between ρ and j_ν (including its variability). In radiative transfer theory, the MEC quantifies the amount of radiative energy released per unit of time, volume, frequency and solid angle (Rybicki & Lightman, 1986).

The accretion shock Ansatz

Cosmological simulations demonstrate that accretion shocks during large-scale structure formation are ubiquitous, and dominate the thermalisation of kinetic energy of baryons falling onto filaments. Given the suite of known particle acceleration processes, and the fact that such shocks are almost certain to exist, shock acceleration is expected to be the biggest contributor to the synchrotron emission from filaments. This paper follows this hypothesis, by only considering the merger and accretion shock contribution to the SCW.

Even under the Ansatz that merger and accretion shocks drive the SCW signal, considerable uncertainty surrounding the correct physical description remains. For example, it is an open question whether the electrons that eventually radiate in synchrotron light originate from the high-energy tail of the thermal Maxwell–Boltzmann

distribution, or are rather accelerated to cosmic ray (CR) energies by SMBHs in the centres of galaxies, and then flung out via jets into the IGM. Apart from uncertainty in the source of energetic electrons, the complex nature of plasma physics also makes it hard to establish how already-energetic electrons attain ultra-relativistic energies in magnetised shock fronts, with the theory of diffusive shock acceleration (DSA) being just one of multiple scenarios. In DSA, CRs are accelerated by repeated crossings of the shock front, which acts as a magnetic mirror (Malkov & Drury, 2001; Xu et al., 2020). Thus, we stress that the functional form of the SCW MEC, which we require to establish a connection with the total matter density field, depends on ill-constrained assumptions surrounding the exact radiation mechanism initiated by the accretion shocks.

The Hoeft & Brüggen (2007) model

Hoeft & Brüggen (2007) have derived an analytic expression for the synchrotron power density of cosmological shock waves, assuming that the radiating electrons exclusively originate from the high-energy tail of the thermal Maxwell–Boltzmann distribution, and that the electron energy spectrum at the shock front is well-described by DSA. As the DSA-based formulae of Hoeft & Brüggen (2007) have been partially successful in explaining observations of shocks in the ICM (e.g. van Weeren et al. (2019); Locatelli et al. (2020)), we postulate that the same formulae can describe synchrotron emission due to accretion shocks onto filaments.[10]

We assume that the power density P_ν of a single shock obeys Eq. 32 of Hoeft & Brüggen (2007). Let the shock surface area be A, and let the effective width of the downstream region be $\langle y \rangle$ (intuitively, this is the thickness of the shock in the direction perpendicular to the surface). The effective shock volume is $V = A\langle y \rangle$. If shocks radiate isotropically, the MEC j_ν is direction-independent. However, like ρ, j_ν is a volume-averaged quantity, so that it depends on the scale on which the averaging occurs. Index all shocks in the universe, so that $P_{\nu,i}(z, \nu)$ is the power density of shock i at cosmological redshift z and emission frequency ν, and A_i and $\langle y \rangle_i$ are its surface area and effective width.[11] Moreover, let $\mathcal{R}_i(z) \subset \mathbb{R}^3$ be the region of space occupied by this shock at redshift z; the Lebesgue measure of $\mathcal{R}_i$ is $V_i = \int_{\mathcal{R}_i} \mathrm{d}\mathbf{r} = A_i \langle y \rangle_i$. Under the approximation that the shock emission is homogeneous within $\mathcal{R}_i$, the average

[10] Araya-Melo et al. (2012) also took this approach and used the MareNostrum simulation to establish that, under these assumptions, filaments at a redshift of 0.15 should produce a flux density of 10^{-1} μJy at 150 MHz.

[11] Because the (cosmology-dependent) function $z(t)$ is strictly decreasing and thus invertible, we can use z as a time coordinate.

total MEC within an arbitrary region $\bar{\mathcal{R}} \subset \mathbb{R}^3$ with Lebesgue measure $\bar{V} = \int_{\bar{\mathcal{R}}} \mathrm{d}\mathbf{r}$ is

$$j_\nu(z,\nu) = \frac{1}{4\pi\,\mathrm{sr}} \frac{1}{\bar{V}} \sum_i P_{\nu,i}(z,\nu) \frac{\int_{\mathcal{R}_i(z)\cap\bar{\mathcal{R}}} \mathrm{d}\mathbf{r}}{V_i}. \tag{3.1}$$

The MEC of a *single* shock (thus dropping indices) located at $\mathbf{r}$ at redshift z, and averaged over its own effective volume, is

$$\begin{aligned} j_\nu(\mathbf{r},z,\nu) &= \frac{1}{4\pi\,\mathrm{sr}} \frac{P_\nu}{V} \\ &= \frac{445\,\mathrm{Jy}}{\mathrm{deg}^2\,\mathrm{Mpc}} \left(\frac{\langle y\rangle}{\mathrm{Mpc}}\right)^{-1} \frac{n_{e,\mathrm{d}}(\mathbf{r},z)}{10\,\mathrm{m}^{-3}} \frac{\xi_e}{10^{-2}} \left(\frac{\nu}{150\,\mathrm{MHz}}\right)^{\alpha(\mathbf{r},z)} \\ &\left(\frac{T_\mathrm{d}(\mathbf{r},z)}{10^8\,\mathrm{K}}\right)^{\frac{3}{2}} \frac{\left(\frac{B_\mathrm{d}}{\mu\mathrm{G}}\right)^{1-\alpha(\mathbf{r},z)}(\mathbf{r},z)}{\left(\frac{B_\mathrm{d}}{\mu\mathrm{G}}\right)^2(\mathbf{r},z) + \left(\frac{B_\mathrm{CMB}}{\mu\mathrm{G}}\right)^2(z)} \left(\frac{150}{1400}\right)^{\alpha(\mathbf{r},z)} \Psi(\mathbf{r},z). \end{aligned} \tag{3.2}$$

Here, $n_{e,\mathrm{d}}$ is the downstream electron number density, T_d is the downstream plasma temperature, α is the integrated spectral index of the associated synchrotron emission, B_d is the downstream plasma magnetic field strength, B_CMB is the CMB magnetic field strength, and Ψ is a dimensionless quantity with a strong dependence on the upstream Mach number $\mathcal{M}_\mathrm{u}$ and a weak dependence on T_d (as can be seen in Fig. 4 of Hoeft & Brüggen (2007)). Ψ approaches unity for high Mach numbers, and so for the WHIM, where the upstream Mach numbers are expected to be high, $\Psi(\mathbf{r},z) \approx 1$. Like Ψ, α does not depend on the spacetime coordinate $(\mathbf{r},z)$ directly, but rather via $\mathcal{M}_\mathrm{u}$ and the adiabatic index γ. Concretely, for $\mathcal{M}_\mathrm{u} > 1$,

$$\alpha = \alpha(\mathcal{M}_\mathrm{u}(\mathbf{r},z), \gamma(\mathbf{r},z)) = \frac{\frac{1}{4}(1-3\gamma)\mathcal{M}_\mathrm{u}^2 - 1}{\mathcal{M}_\mathrm{u}^2 - 1}. \tag{3.3}$$

Note that $\alpha \to -\infty$ as $\mathcal{M}_\mathrm{u} \to 1+$, and that $\alpha \to {}^1\!/_4(1-3\gamma)$ as $\mathcal{M}_\mathrm{u} \to \infty$. The top panel of Fig. 3.1 illustrates this behaviour. The electron acceleration efficiency ξ_e quantifies the fraction of the shock's thermal energy that is used to accelerate suprathermal electrons. Unfortunately, ξ_e is unknown for WHIM shocks, but a comparison to supernova remnant (SNR) shocks suggests $\xi_e = 0.05$ (Keshet et al., 2004).

Enzo simulations

To capture, in a statistical sense, the relationship between the total matter density and the Hoeft & Brüggen (2007) monochromatic emission coefficient, we turned to MHD simulations. In particular, we used snapshots (i.e. 3D spatial fields at constant time) of the largest uniform-grid cosmological MHD simulation to date (Vazza et al., 2019), which is based on the Enzo code (Bryan et al., 2014). These cubic snapshots cover a comoving volume of $(100\,\mathrm{Mpc})^3$ with 2400^3 voxels, yielding a (comoving) resolution of $41\frac{2}{3}$ kpc per voxel edge. The simulations recreate the evolution of the baryonic and dark matter density functions ρ_{BM} and ρ_{DM} — as well as (thermo)dynamic quantities[12] such as the gas temperature T, magnetic field strength B and gas velocity v — under Newtonian gravity. (However, the effects of the expansion of the Universe as predicted by general relativity are still incorporated.) No galactic physics is included. Shocks in the snapshots can be identified by searching for temperature and velocity jumps, as described in Vazza et al. (2009).[13] For shocked voxels, we simply took $\langle y \rangle$ to be the voxel edge length, and calculated α and Υ after establishing $\mathcal{M}_{\mathrm{u}}$. With the exception of ξ_e, which we assumed to be a constant throughout, the simulation thus enables us to compute all factors on the RHS of Eq. 3.2.

Taking their product for the $z = 0.025$ snapshot, we obtained j_ν, and compared it to $\rho := \rho_{\mathrm{BM}} + \rho_{\mathrm{DM}}$ to study their relationship.[14]

[12] As long as one considers a single shock, space (at a fixed time) can be classified into an upstream region and a downstream region, with associated temperatures (T_{u} and T_{d}, respectively) and magnetic field strengths (B_{u} and B_{d}, respectively). This distinction is less meaningful once one considers multiple shocks at the same time: a given location could then be upstream for some shocks, and downstream for others. Therefore, cosmological simulations do not evolve an upstream or downstream temperature field, but just a *general* temperature field T. Analogously, cosmological simulations maintain only *general* magnetic field component fields B_x, B_y and B_z (and thus a magnetic field strength field $B := \sqrt{B_x^2 + B_y^2 + B_z^2}$), without upstream and downstream distinction.

[13] From the middle panel of Fig. 3.1, it is clear that shocks also induce a jump in gas density. However, this jump is modest, saturating at e.g. a factor 4 for $\gamma = 5/3$. By contrast, the jump in temperature, seen in the bottom panel of the same figure, can be several orders of magnitude — in the WHIM, at least.

[14] By repeating our analysis for snapshots of other redshifts, one could study the time evolution of this relationship. However, as the Hoeft & Brüggen (2007) model is likely only a rough description of the actual synchrotron emission mechanism in filaments, we currently consider such level of detail superfluous.

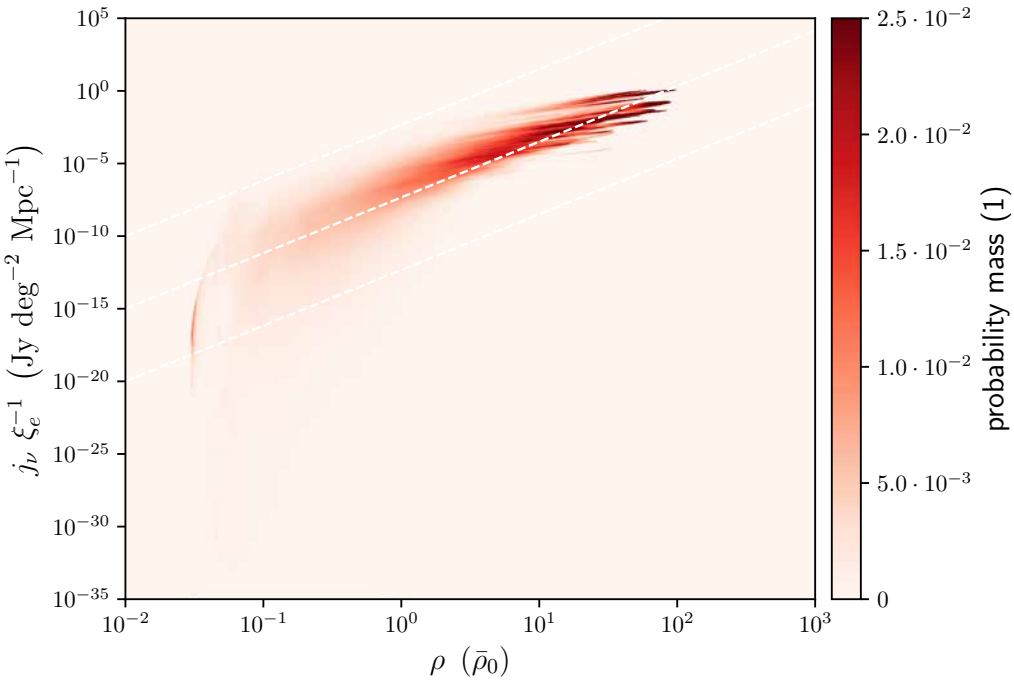


Figure 3.3: Conditional probability distributions $P(j_\nu \,|\, \rho)$ of synchrotron monochromatic emission coefficient j_ν (at rest-frame (emission) frequency $\nu = 150$ MHz) given total matter density ρ for the Hoeft & Brüggen (2007) formalism, as derived from $z = 0.025$ snapshots of cosmological MHD simulations by Vazza et al. (2019). The dashed lines indicate a $j_\nu \propto \rho^{23/6}$ (single shock) scaling relation for various proportionality constants. ρ is shown relative to the current-day mean total matter density $\bar{\rho}_0 = \Omega_{\mathrm{M},0}\rho_{c,0}$, while the weakly constrained electron acceleration efficiency ξ_e is divided out from j_ν. Each conditional is numerically approximated by a probability mass function with 1000 bins, whose edges vary by a constant factor and span 40 orders of magnitude in j_ν.

Conditional probability distribution

In order to convert the BORG SDSS total matter density fields into synchrotron MEC fields, we needed to predict j_ν from ρ — ideally including uncertainty. Because the BORG SDSS has a comoving resolution of 4.17 Mpc per voxel edge compared to Enzo's $4.17 \cdot 10^{-2}$ Mpc per voxel edge, we blurred both j_ν and ρ with a Gaussian kernel whose standard deviation is half of the comoving resolution ratio. We then treated each (j_ν, ρ) pair of the blurred fields as a draw from the joint distribution $P(j_\nu, \rho)$. By binning the data, we performed a simplistic form of kernel density estimation (KDE). Next, we calculated the conditional $P(j_\nu \,|\, \rho)$ from the joint by dividing it by the marginal $P(\rho) := \int_0^\infty P(j_\nu, \rho)\, \mathrm{d}j_\nu$.

The result is the set of probability distributions shown in Fig. 3.3. We overplot the single shock scaling relation expected for the WHIM regime: $j_\nu \propto \rho^{23/6}$. See Ap-

pendix 3.A2 for a derivation.

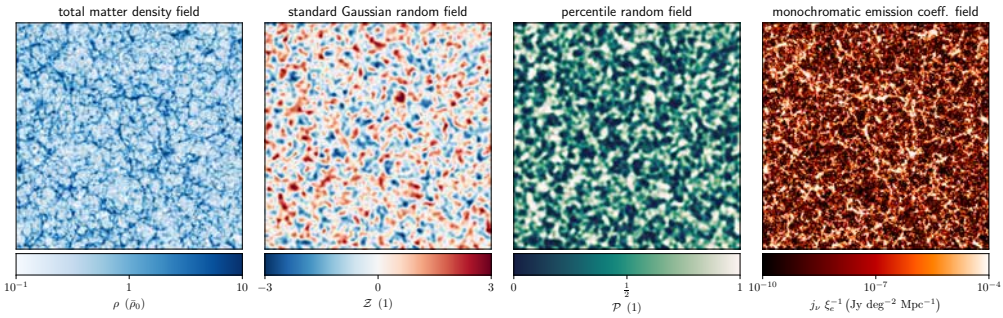


Figure 3.4: Overview of the procedure for converting a probability distribution over total (i.e. baryonic and dark) matter density fields into a probability distribution over monochromatic emission coefficient (MEC) fields. We show a fixed slice through $\mathcal{R}$, of 750 Mpc h^{-1} × 750 Mpc h^{-1} for the density and MEC fields, and a subregion of 150 Mpc h^{-1} × 150 Mpc h^{-1} for the Gaussian and percentile random fields (to more clearly illustrate the bijective mapping between the two). *Left:* BORG SDSS density ρ sample, in multiples of the current-day mean total matter density $\bar{\rho}_0 = \Omega_{\mathrm{M},0}\,\rho_{\mathrm{c},0}$. *Middle left:* Gaussian random field $\mathcal{Z}$ sample using a squared exponential (SE) kernel with length-scale $l_{\mathrm{SE}} = 2$ Mpc. *Middle right:* the corresponding percentile random field $\mathcal{P}$ sample. *Right:* the corresponding MEC j_ν sample at an emission frequency of $\nu = 150$ MHz, with the weakly-constrained electron acceleration efficiency ξ_e divided out.

3.2.4 Generating monochromatic emission coefficient fields

In principle, we could — for each BORG SDSS sample $\rho(\mathbf{r})$ — convert to $j_\nu(\mathbf{r})$ by drawing from $P(j_\nu \,|\, \rho = \rho(\mathbf{r}))$ independently for every $\mathbf{r} \in \mathcal{R}$. However, this is clearly suboptimal, as accurate SCW MEC fields exhibit spatial correlation up to megaparsec scales, both under the merger and accretion shock Ansatz as well as in a turbulence scenario (Govoni et al., 2019; Brunetti & Vazza, 2020).

To generate spatially correlated draws from our conditional probability distribution, we used realisations of a Gaussian random field $\mathcal{Z}$ over $\mathcal{R}$. In general, the statistical properties of a GRF are determined by its mean, and its covariance function or *kernel*. In our case, we set the mean to 0, and chose an isotropic and stationary kernel, as suggested by the Cosmological Principle. From the variety of remaining choices commonly used in machine learning (ML), we picked[15] the squared-exponential (SE)

[15] The choice of kernel should ideally reflect the morphology of merger and accretion shocks. However, given the current spatial resolution of the BORG SDSS density fields and our requirement that the kernel be isotropic (disregarding the relation between shock morphology and local LSS orientation), the approach is already so approximate that no particular kernel is clearly preferred over the others.

kernel $K_{\mathrm{SE}} : \mathcal{R} \times \mathcal{R} \to \mathbb{R}_{\geq 0}$ (Rasmussen & Williams, 2006):

$$K_{\mathrm{SE}}\left(\mathbf{r}, \mathbf{r}'\right) = \sigma_{\mathrm{SE}}^2 \exp\left(-\frac{||\mathbf{r} - \mathbf{r}'||_2^2}{2l_{\mathrm{SE}}^2}\right), \text{ so that}$$
$$\mathrm{Cov}\left(\mathcal{Z}\left(\mathbf{r}\right), \mathcal{Z}\left(\mathbf{r}'\right)\right) = K_{\mathrm{SE}}\left(\mathbf{r}, \mathbf{r}'\right) \text{ with } \mathbf{r}, \mathbf{r}' \in \mathcal{R}. \tag{3.4}$$

The variance σ_{SE}^2 allows one to endow the GRF with dimensionality and scale it at will. As our GRF only served to draw spatially correlated MEC samples, we set $\sigma_{\mathrm{SE}} = 1$. The lengthscale l_{SE} is the characteristic scale of spatial correlation, and forms an important model choice. Some of the longest coherent shocks observed thus far are the Toothbrush Relic in galaxy cluster 1RXS J0603.3+4214 (van Weeren et al., 2012) and the Sausage Relic in galaxy cluster CIZA J2242.8+5301 (Di Gennaro et al., 2018), both with a largest linear size (LLS) of ∼2 Mpc. Although the majority of observed shocks seem smaller (suggesting $l_{\mathrm{SE}} < 2$ Mpc), the presence of noise in radio imagery obfuscates the full extent of shock fronts, biasing measurements towards small LLSs. As a compromise, we chose $l_{\mathrm{SE}} = 2$ Mpc.

As is customary in cosmological simulations — where GRFs with isotropic and stationary kernels are used to initialise the Early Universe matter density fields — we generated our zero mean, unit covariance GRF realisations on the BORG SDSS voxel grid using Fourier analysis. This necessitated calculating the power spectrum of the kernel, which in the case of the 3D SE kernel is

$$P_{\mathrm{SE}}\left(\mathbf{k}\right) = \left(2\pi l_{\mathrm{SE}}^2\right)^{\frac{3}{2}} \exp\left(-2\pi ||\mathbf{k}||_2^2 l_{\mathrm{SE}}^2\right). \tag{3.5}$$

Finally, Fourier-generated GRFs on a finite (numerical) grid require an appropriate normalisation; we used Parseval's theorem to find the correct factor.

This procedure generated a 'standard normal' GRF $\mathcal{Z}$, in the sense that each location's RV is standard normal: $\mathcal{Z}\left(\mathbf{r}\right) \sim \mathcal{N}\left(0, 1\right)$ for all $\mathbf{r} \in \mathcal{R}$. Next, we applied the CDF Φ of the standard normal to the GRF in point-wise fashion (i.e. voxel-wise, in practice), thereby creating a 'percentile random field' $\mathcal{P}$:

$$\mathcal{P}\left(\mathbf{r}\right) := \Phi\left(\mathcal{Z}\left(\mathbf{r}\right)\right), \text{ where } \Phi\left(x\right) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(\frac{-y^2}{2}\right) \mathrm{d}y. \tag{3.6}$$

This field has values strictly between 0 and 1, and inherits the spatial correlations present in $\mathcal{Z}$.

Let $\Xi_\rho\left(x\right) : \mathbb{R}_{\geq 0} \to [0, 1]$ be the CDF of the RV $j_\nu \,|\, \rho$, and let $\Xi_\rho^{-1}\left(y\right) : [0, 1] \to$

$\mathbb{R}_{\geq 0}$ be its inverse. Then the final MEC field is given by

$$j_\nu(\mathbf{r}) := \Xi^{-1}_{\rho(\mathbf{r})}\left(\mathcal{P}(\mathbf{r})\right), \text{ for all } \mathbf{r} \in \mathcal{R}. \tag{3.7}$$

Because ρ and $\mathcal{P}$ are random fields on $\mathcal{R}$, so is j_ν. A graphical summary of the procedure described in this subsection, and defined in Eq. 3.7, is shown in Fig. 3.4.

3.2.5 Generating specific intensity functions

To find the specific intensity of the SCW on the sky, I_ν, we simulated the passage of light rays through our 3D MEC fields via ray tracing. Conveniently, projecting a ray's 4D null geodesic in a pure Friedmann–Lemaître–Robertson–Walker (FLRW) metric onto 3D comoving space results in a straight line. We provide a brief derivation in Appendix 3.A3. By assuming an exact FLRW metric, we neglected all spacetime deformations due to local (large-scale structure) energy density fluctuations. Our results thus do not feature gravitational lens effects around massive clusters.

We estimated I_ν in a sky patch by sampling many directions $\hat{r}_i$ within it, and simulating the passage of light through our reconstructed SCW in comoving space for each such direction (or 'ray') $\hat{r}_i$.

From proper monochromatic emission coefficient to observer's specific intensity

Next, we establish the relation between the observed specific intensity I_ν of a ray that has travelled through the SCW to Earth from direction $\hat{r}$, and the MEC j_ν along its path.

Be $j_\nu : \mathbb{R}^3 \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{>0} \to \mathbb{R}_{\geq 0}$ the function that assigns to each 3-tuple $(\mathbf{r}, z, \nu)$ containing a comoving locus $\mathbf{r}$, a cosmological redshift z (which represents time) and an EM wave frequency ν, the proper (rather than comoving) MEC $j_\nu(\mathbf{r}, z, \nu)$ (with SI units $\mathrm{W\,m^{-3}\,Hz^{-1}\,sr^{-1}}$). Conceptually, we must differentiate between the specific intensity of large-scale structure at the time of emission ('there and then'), and at today's observing epoch ('here and now'). Today's quantity I_ν for direction $\hat{r}$ and observing frequency ν_{obs} is

$$I_\nu(\hat{r}, \nu_{\mathrm{obs}}) = \frac{c}{H_0}\int_0^\infty \frac{j_\nu\left(r(z)\,\hat{r}, z, \nu_{\mathrm{obs}}(1+z)\right)}{(1+z)^4 E(z)} \mathrm{d}z. \tag{3.8}$$

We provide a brief derivation, alongside explicit expressions for $r(z)$ and $E(z)$, in Appendix 3.A4.

For each specific intensity function that we wished to generate, we now drew $M = 10^6$ ray directions uniformly from the π sr lune that covers the SDSS-constrained half of the Northern Sky. For each ray, we calculated the corresponding specific intensity by combining Eq. 3.2 and 3.8.

Note that this requires evaluating the MEC field at a range of emission frequencies. Numerically, we realised this by building both the Enzo-derived conditional probability distribution $P(j_\nu \,|\, \rho)$ and the MEC field sample (using the same density field sample and GRF sample) at two emission frequencies ν: $\nu = \nu_{\text{obs}}$ and $\nu = \nu_{\text{obs}}\,(1 + z_{\text{max}})$. Next, for each voxel, we calculated a spectral index from the two MEC fields by assuming a power-law spectrum, and used it to find the MEC at the emission frequency $\nu = \nu_{\text{obs}}\,(1 + z)$ needed given the voxel's cosmological redshift z.

We repeated this process many times by selecting 1000 density samples from the BORG SDSS posterior (we discarded the first 2110 samples due to burn-in, and thinned the Monte Carlo Markov chain by a factor 10) and generating one independent and identically distributed (IID) GRF sample for each. (There is no compelling reason to thin the chain and we could have used multiple GRF samples per density sample; the current approach is merely to limit data storage, compute time and energy usage.) Thus, we numerically realised $1000 \cdot 1 = 1000$ specific intensity functions. Together, they form a probability distribution over MASSCW specific intensity functions over a quarter of the sky.

3.2.6 Generating redshift functions

Our methodology can also be used to characterise the redshift properties of the MASSCW.

Specific intensity–weighted mean redshift

A ray's total specific intensity is found by summing up the specific intensity contributions from all voxels along its path, each of which has a known redshift. Therefore, a ray's mean redshift is found by weighing each voxel's redshift by the corresponding specific intensity contribution, before dividing by the sum of such contributions. Concretely, we define the specific intensity–weighted mean redshift $\bar{z}$ at observing frequency ν_{obs} of a ray $\hat{r}$ passing through MEC field j_ν to be

$$\bar{z}\,(\hat{r}, \nu_{\text{obs}}) := I_\nu^{-1}\,(\hat{r}, \nu_{\text{obs}}) \int_0^\infty z\,\frac{\mathrm{d}I_\nu}{\mathrm{d}z}\,(\hat{r}, \nu_{\text{obs}}, z)\;\mathrm{d}z. \tag{3.9}$$

By iterating over total matter density and GRF samples, we generated a probability distribution over the function $\bar{z}(\hat{r}, \nu_{\text{obs}})$, in exact analogy to the generation of the distribution over specific intensity functions $I_\nu(\hat{r}, \nu_{\text{obs}})$.

One might wonder what the specific intensity–weighted mean redshift of the MASSCW signal is for a randomly picked direction on the sky. The set $\{\bar{z}(\hat{r}_i, \nu_{\text{obs}}) \mid i \in \{1, 2, \ldots, M\}\}$, with rays $\hat{r}_i$ drawn from a uniform distribution over the sky, can be viewed as a random sample from a specific intensity–weighted mean redshift random variable $\bar{Z}$. An RV is fully characterised by its CDF. Each MASSCW prior realisation generates another empirical CDF (ECDF) for $\bar{Z}$.[16]

Flux density–weighted mean redshift

One might also wonder what the mean redshift of the MASSCW signal is — not for a single direction, but overall. Due to strong attenuation with redshift, the MASSCW specific intensity is usually 'high' for directions with a 'low' specific intensity–weighted mean redshift, and vice versa. This means that, although the signal could originate from far away for most of the sky (given sufficiently small cluster and filament volume-filling fractions), the sky-averaged redshift can still be low. We define the flux density–weighted mean redshift of the MASSCW signal at observing frequency ν_{obs} to be

$$\bar{\bar{z}} := \left(\int_{\mathbb{S}^2} I_\nu(\hat{r}, \nu_{\text{obs}}) \, \mathrm{d}\Omega \right)^{-1} \int_{\mathbb{S}^2} \int_0^\infty z \, \frac{\mathrm{d}I_\nu}{\mathrm{d}z}(\hat{r}, \nu_{\text{obs}}, z) \, \mathrm{d}z \, \mathrm{d}\Omega, \tag{3.10}$$

where $\bar{\bar{z}}$ is shorthand for $\bar{\bar{z}}(\nu_{\text{obs}})$.

3.3 Redshift predictions from geometric cosmic web model

3.3.1 Overcoming the redshift limitation

Although we can calculate the specific intensity function $I_\nu(\hat{r}, \nu_{\text{obs}})$, the specific-intensity–weighted mean redshift function $\bar{z}(\hat{r}, \nu_{\text{obs}})$, the single-direction redshift RV $\bar{Z}$ ECDF and the flux density–weighted mean redshift $\bar{\bar{z}}(\nu_{\text{obs}})$ for each realisation of the MASSCW prior, all four quantities suffer from the fact that the BORG SDSS reconstructions stop at a redshift $z_{\text{max}} = 0.2$. It is of prime interest to know to what extent such redshift limitations, which we anticipate will become less stringent in the

[16] The probability distribution over ECDFs of $\bar{Z}$ thus obtained is less informative than the probability distribution over functions $\bar{z}$, because it disposes of the directional correlations in specific intensity–weighted mean redshift $\bar{z}(\hat{r}, \nu_{\text{obs}})$.

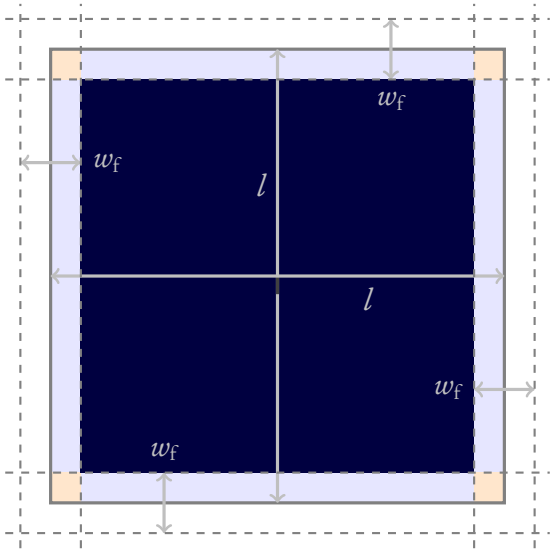


Figure 3.5: Geometry of the simple model, in which we treat the cosmic web as a collection of randomly displaced cubic unit cells. We show one face (grey square) of a unit cell, with edges of comoving length l. Filaments of typical comoving width w_f and length $l_f := l - w_f$ cover the light-blue–shaded region of the face, and extend to the faces of neighbouring unit cells (grey dashed lines). Where filaments meet, galaxy clusters reside (orange-shaded regions), each of which is also connected to a filament oriented away from the observer. The central part of the face represents a sheet (dark blue), behind which a large void looms (not depicted). The figure is not to scale: filament widths are exaggerated compared to realistic filament lengths.

future, affect our inferences. For example, if most of the real MASSCW signal originates from $z > 0.2$, the predictive power of our specific intensity function distribution would be limited.

Here we introduce a simple analytical model that allows us to calculate both the distribution of $\bar{Z}$, and $\bar{\bar{z}}$, for an arbitrary value of $z_{\max}$. The model reproduces the results calculated from the MASSCW prior, if it is equally limited to $z_{\max} = 0.2$. This provides tentative evidence that the model captures the essential elements of the MASSCW signal, so that it might be used for extrapolation. We thus use the model to predict the distribution of $\bar{Z}$ and $\bar{\bar{z}}$, for the case that describes the actual Universe: $z_{\max} \to \infty$.

3.3.2 Model formulation

To retrieve MASSCW redshift properties, we propose a simple geometric model. The key idea is that the dominant MASSCW redshift for an arbitrary sightline can be calculated via a weighted sum of the redshifts of the clusters and filaments it passes through, where the weights are set by the geometry and MECs of the LSS pierced.

The redshifts of the structures a sightline pierces through depend on the particular LSS realisation the observer is immersed in. However, the cosmological principle dictates that LSS is statistically similar everywhere, and thus typical geometric cosmic web parameters must exist (and can be retrieved from numerical simulations; see for example Gheller et al. (2015)).

Here we assume that sightlines pass through an arrangement of identical cubic unit cells — ignoring morphological variation among clusters, filaments, sheets and voids — with comoving edge length l. The edges are surrounded by square cuboids, which represent filaments, of typical comoving width w_f and length $l_\mathrm{f} := l - w_\mathrm{f}$ (and thus volume $w_\mathrm{f}^2 l_\mathrm{f}$).[17] Galaxy clusters reside where filaments meet, and — in this simplistic model — are cubes with comoving edge length w_f (and thus volume w_f^3). By far the largest component of a unit cell is its central void, a cubical region of comoving edge length l_f (and thus volume l_f^3). Finally, unit cells contain sheets, which fill the regions bounded by filaments and voids. They are of typical comoving length l_f and thickness w_f (and thus volume $l_\mathrm{f}^2 w_\mathrm{f}$). Fig. 3.5 depicts a face of a unit cell. If we were to assume a perfect crystal structure for the cosmic web, some sightlines would never encounter clusters or filaments whilst others would consistently do so at regular (comoving) intervals. To avoid this unphysical scenario, we assume that our sightline-of-interest enters every unit cell it encounters at a random position on the face of incidence. For additional simplicity, we assume that the sightline always hits such faces perpendicularly.

If one would choose a point-of-incidence on the face shown in Fig. 3.5 in uniform fashion, the sightline would hit a sheet and then a void (dark-blue-shaded area) with probability

$$p_{\mathrm{s-v}}\left(\frac{w_\mathrm{f}}{l}\right) = \frac{(l - w_\mathrm{f})^2}{l^2} = \left(1 - \frac{w_\mathrm{f}}{l}\right)^2 . \tag{3.11}$$

Note that $p_{\mathrm{s-v}} \to 1$ when $\frac{w_\mathrm{f}}{l} \to 0$, and $p_{\mathrm{s-v}} \to 0$ when $\frac{w_\mathrm{f}}{l} \to 1$, as required. Likewise, the probability that the sightline hits a filament only (light-blue-shaded area) is

$$p_\mathrm{f}\left(\frac{w_\mathrm{f}}{l}\right) = \frac{2w_\mathrm{f}\,(l - w_\mathrm{f})}{l^2} = 2\frac{w_\mathrm{f}}{l}\left(1 - \frac{w_\mathrm{f}}{l}\right) . \tag{3.12}$$

Finally, the probability that the sightline hits a cluster first, and then the filament

[17] Often, filaments are modelled as cylinders. In such cases, the length of the path of a sightline through a filament depends on the exact point of incidence on the unit cell boundary. However, in our simplistic geometric model, this effect does not arise.

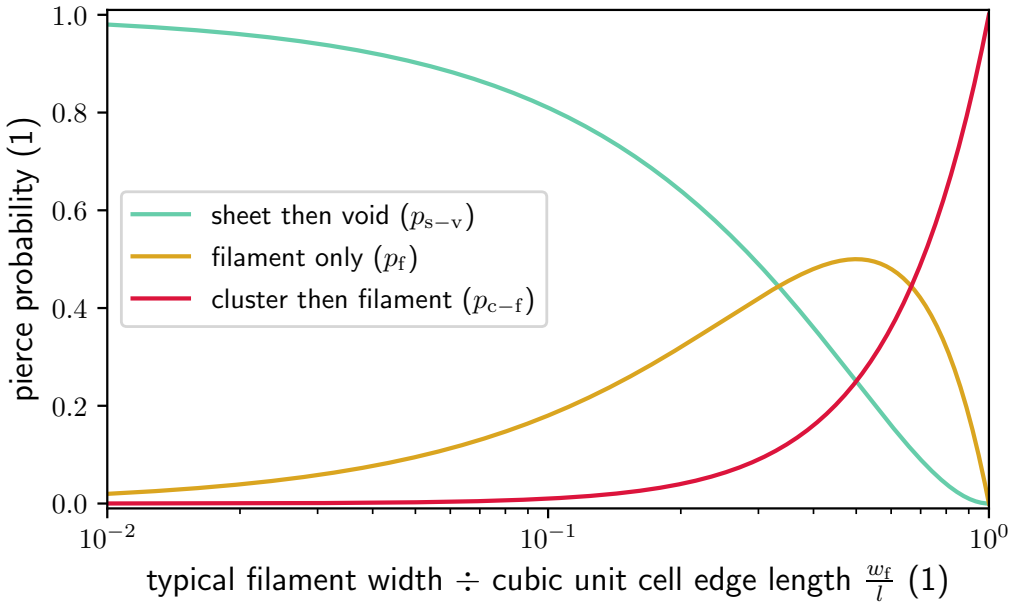


Figure 3.6: When a sightline enters a cubic unit cell of the cosmic web perpendicularly to a face (as in Fig. 3.5) and at a random position on it, one of three different LSS piercing events occurs. We show the probabilities for each event as a function of the ratio of the typical filament width and cubic unit cell edge length (see Eqs. 3.11–3.13).

behind it (orange-shaded area) is

$$p_{\mathrm{c-f}}\left(\frac{w_{\mathrm{f}}}{l}\right) = \left(\frac{w_{\mathrm{f}}}{l}\right)^2 . \tag{3.13}$$

Fig. 3.6 depicts these pierce probabilities as a function of the ratio between filament width and unit cell edge length $\frac{w_{\mathrm{f}}}{l}$. For a comparison between the LSS volume-filling factors (VFFs) of this simple model and those from cosmological simulations, see Appendix 3.A5.

Pick a sightline, and successively label each unit cell boundary this sightline crosses with a natural number $n \in \mathbb{N}_{\geq 1}$. If the first boundary crossing occurs at a comoving distance d_1, then the n-th crossing happens at comoving distance $d_n = d_1 + (n-1)\,l$. Let $\mathfrak{M}$ be the cosmological model of preference, and let $z_{\mathfrak{M}}(d)$ denote the function that converts comoving distance to cosmological redshift under this model. The unit cell boundary crossings occur at redshifts $z_n := z_{\mathfrak{M}}(d_n)$ for all $n \in \mathbb{N}_{\geq 1}$.

Because we want to be able to calculate the dominant redshift of the MASSCW signal produced by filaments and clusters up to a given cosmological redshift $z_{\max}$ only, we introduce a parameter N. Intuitively, $N \in \mathbb{N}_{\geq 1}$ is the label of the last boundary crossing within the LSS considered. More formally, $N := \max\{n \in \mathbb{N}_{\geq 1} \mid z_n <$

$z_{\max}\}$.

Let $X_n(\nu_{\text{obs}})$ be the RV denoting the contribution to the MASSCW specific intensity at observing frequency ν_{obs} picked up by the sightline during the n-th full unit cell crossing (that is whilst travelling through the n-th *newly entered* unit cell; the unit cell the observer resides in does not contribute). To retain low complexity, we assign the complete specific intensity contribution from the n-th crossing to the redshift of incidence z_n. It is useful to regard X_n as an instance of a more general random variable: $X_n(\nu_{\text{obs}}) := X(\nu_{\text{obs}}, z_n)$. $X(\nu_{\text{obs}}, z)$ is the RV with support $(0, \Delta I_{\nu,\text{f}}(\nu_{\text{obs}}, z), \Delta I_{\nu,\text{c-f}}(\nu_{\text{obs}}, z))$, and corresponding probabilities $(p_{\text{s-v}}, p_{\text{f}}, p_{\text{c-f}})$. Note that we assume that sheets and voids have a vanishing contribution to the MASSCW signal. Here,

$$\Delta I_{\nu,\text{f}}(\nu_{\text{obs}}, z) := w_{\text{f}} j_{\nu,\text{f}}(\nu_{\text{obs}}(1+z), z)(1+z)^{-4} \tag{3.14}$$

$$\begin{aligned}\Delta I_{\nu,\text{c-f}}(\nu_{\text{obs}}, z) := (&w_{\text{f}} j_{\nu,\text{c}}(\nu_{\text{obs}}(1+z), z) \\ &+ l_{\text{f}} j_{\nu,\text{f}}(\nu_{\text{obs}}(1+z), z))(1+z)^{-4}.\end{aligned} \tag{3.15}$$

The factor $(1+z)^{-4}$ follows from Eq. 3.51: the specific intensity as an integral over comoving radial distance. As we assume LSS to occur at regularly-spaced radial comoving distances d_n, this integral can be approximated by a Riemann sum, where each term equals the integrand multiplied by $\Delta r = l$.

Although redshift-dependent, we assume the MEC to be constant (at a given emission frequency) within — but different between — clusters and filaments of the same unit cell. Respectively, $j_{\nu,\text{f}}(\nu, z)$ and $j_{\nu,\text{c}}(\nu, z)$ represent the typical proper filament and cluster SCW MECs at emission frequency ν and cosmological redshift z. We use the dimensionless parameter $\mathcal{C}$ to denote the typical cluster-to-filament-SCW MEC ratio at some reference emission frequency $\nu = \nu_{\text{ref}}$ and $z = 0$: $j_{\nu,\text{c}}(\nu_{\text{ref}}, 0) := \mathcal{C} j_{\nu,\text{f}}(\nu_{\text{ref}}, 0)$. To retain minimal complexity, we propose to describe the spectral and temporal dependencies of both $j_{\nu,\text{f}}(z)$ and $j_{\nu,\text{c}}(z)$ as power laws in $\frac{\nu}{\nu_{\text{ref}}}$ and $(1+z)$:

$$j_{\nu,\text{f}}(\nu, z) = j_{\nu,\text{f}}(\nu_{\text{ref}}, 0)\left(\frac{\nu}{\nu_{\text{ref}}}\right)^{\alpha_{\text{f}}}(1+z)^{\beta_{\text{f}}} \tag{3.16}$$

$$j_{\nu,\text{c}}(\nu, z) = j_{\nu,\text{c}}(\nu_{\text{ref}}, 0)\left(\frac{\nu}{\nu_{\text{ref}}}\right)^{\alpha_{\text{c}}}(1+z)^{\beta_{\text{c}}}, \tag{3.17}$$

with $\alpha_{\text{f}}, \beta_{\text{f}}, \alpha_{\text{c}}$ and β_{c} constants. Under this choice of MEC function parametrisation,

the expectation value of $X(\nu_{\text{obs}}, z)$ becomes

$$\mathbb{E}\left(X\left(\nu_{\text{obs}}, z\right)\right) = p_{\text{f}} \Delta I_{\nu,\text{f}}\left(\nu_{\text{obs}}, z\right) + p_{\text{c-f}} \Delta I_{\nu,\text{c-f}}\left(\nu_{\text{obs}}, z\right) \tag{3.18}$$

$$= (1+z)^{-4} \left(\frac{w_{\text{f}}}{l}\right)^2 j_{\nu,\text{f}}\left(\nu_{\text{ref}}, 0\right) \cdot$$

$$\left(3 l_{\text{f}} \left(\frac{\nu_{\text{obs}}}{\nu_{\text{ref}}}\right)^{\alpha_{\text{f}}} (1+z)^{\alpha_{\text{f}}+\beta_{\text{f}}} + \mathcal{C} w_{\text{f}} \left(\frac{\nu_{\text{obs}}}{\nu_{\text{ref}}}\right)^{\alpha_{\text{c}}} (1+z)^{\alpha_{\text{c}}+\beta_{\text{c}}}\right). \tag{3.19}$$

3.3.3 Specific intensity–weighted mean redshift

Consider the random vector $[X_1, X_2, \ldots, X_N]^T$. Whenever $\max\{X_1, X_2, \ldots, X_N\} > 0$, we can define the MASSCW specific intensity–weighted mean redshift RV

$$\bar{Z}\left(\nu_{\text{obs}}\right) := \frac{\sum_{n=1}^{N} z_n X_n\left(\nu_{\text{obs}}\right)}{\sum_{n=1}^{N} X_n\left(\nu_{\text{obs}}\right)}. \tag{3.20}$$

We stress that $\bar{Z}$, being a deterministic function of the RVs $\{X_1, X_2, \ldots, X_N\}$, is itself an RV. It quantifies the variety of specific intensity–weighted mean redshifts that could occur in *different* LSS realisations (with the same model parameters) for a *fixed* direction (sightline) of the sky. Conversely, it could be interpreted as representing the redshift variety that occurs in a *fixed* LSS realisation for *different* directions of the sky, provided that the sky contains many (almost) independent patches of LSS. This second interpretation invites for a comparison between the specific intensity–weighted mean redshift distribution of our MASSCW prior skies, and the distribution of $\bar{Z}$.

Because $X_n\left(\nu_{\text{obs}}\right)$ features in both numerator and denominator of the fraction in Eq. 3.20, redshift-independent multiplicative factors cancel out. Therefore, $j_{\nu,\text{f}}\left(\nu_{\text{ref}}, 0\right)$ need not be specified; also, $\bar{Z}$ becomes ν_{obs}-independent if $\alpha_{\text{f}} = \alpha_{\text{c}}$. In general, the distribution of $\bar{Z}$ is determined by the cosmological model $\mathfrak{M}$ and 9 additional parameters (w_{f}, l, d_1, $z_{\max}$, α_{f}, β_{f}, α_{c}, β_{c} and $\mathcal{C}$). Using d_1, we can force the filaments and clusters nearest to the observer to occur at a fixed comoving distance. Alternatively, one could randomise the observer's position with respect to the unit cells by choosing $d_1 \sim \text{Uniform}\,(0, l)$, making it an RV. In such case, 8 parameters remain that determine the distribution of $\bar{Z}$.

3.3.4 Flux density–weighted mean redshift

The geometric model also allows to calculate the flux density–weighted mean (that is sky-averaged) MASSCW redshift $\bar{\bar{z}}\left(\nu_{\text{obs}}\right)$. We did so by considering $M \in \mathbb{N}_{\geq 1}$ sightlines (instead of one) and their associated specific intensity–weighted mean redshift

RVs $\{\bar{Z}_1, \bar{Z}_2, ..., \bar{Z}_M\}$ in the sense of Eq. 3.20. In total, these RVs depend on a set of $N \cdot M$ discrete RVs X_{nm}, which can be partitioned into N subsets of M IID RVs: $X_{nm} \sim X_n$. We find $\bar{\bar{z}}$ by summing the elements of $\{\bar{Z}_1, \bar{Z}_2, ..., \bar{Z}_M\}$, each weighted by the corresponding specific intensity, and dividing the result by the sum of these weights; all whilst we let $M \to \infty$:

$$\bar{\bar{z}}\left(\nu_{\text{obs}}\right) := \lim_{M\to\infty} \frac{\sum_{m=1}^{M}\sum_{n=1}^{N} z_n X_{nm}\left(\nu_{\text{obs}}\right)}{\sum_{m=1}^{M}\sum_{n=1}^{N} X_{nm}\left(\nu_{\text{obs}}\right)} \tag{3.21}$$

$$= \lim_{M\to\infty} \frac{\sum_{n=1}^{N}\left(z_n M^{-1}\sum_{m=1}^{M} X_{nm}\left(\nu_{\text{obs}}\right)\right)}{\sum_{n=1}^{N}\left(M^{-1}\sum_{m=1}^{M} X_{nm}\left(\nu_{\text{obs}}\right)\right)} \tag{3.22}$$

$$= \frac{\sum_{n=1}^{N} z_n \mathbb{E}\left(X_n\left(\nu_{\text{obs}}\right)\right)}{\sum_{n=1}^{N} \mathbb{E}\left(X_n\left(\nu_{\text{obs}}\right)\right)}. \tag{3.23}$$

Here we use the fact that the limit of a ratio of two sequences equals the ratio of the sequences' limits (provided that the denominator sequence does not converge to 0), and that the sample mean $M^{-1}\sum_{m=1}^{M} X_{nm} \to \mathbb{E}\left(X_n\right)$ as $M \to \infty$, for all $n \in \{1, 2, ..., N\}$. Like before, as $\mathbb{E}\left(X_n\left(\nu_{\text{obs}}\right)\right)$ features in both numerator and denominator of the fraction in Eq. 3.23, redshift-independent multiplicative factors cancel out. Thus, $\bar{\bar{z}}$ is ν_{obs}-independent when $\alpha_{\text{f}} = \alpha_{\text{c}}$.

Eq. 3.19 makes clear that clusters dominate over filaments at cosmological redshift z when

$$\mathcal{C} \gg 3\frac{l_{\text{f}}}{w_{\text{f}}}\left(\frac{\nu_{\text{obs}}}{\nu_{\text{ref}}}\right)^{\alpha_{\text{f}}-\alpha_{\text{c}}}(1+z)^{\alpha_{\text{f}}-\alpha_{\text{c}}+\beta_{\text{f}}-\beta_{\text{c}}}. \tag{3.24}$$

At $\nu_{\text{obs}} = \nu_{\text{ref}}$ and $z = 0$, this inequality reduces to

$$\mathcal{C} \gg 3\left(\frac{l}{w_{\text{f}}} - 1\right), \tag{3.25}$$

which is amply satisfied (see Sect. 3.3.5 for parameter estimates). Near $\nu_{\text{obs}} = \nu_{\text{ref}}$ and for z of order unity, the inequality continues to hold as long as $\alpha_{\text{f}} - \alpha_{\text{c}} + \beta_{\text{f}} - \beta_{\text{c}}$ is a number of order unity at most. In addition, most redshifts high enough that the inequality is violated, do not contribute meaningfully to $\bar{\bar{z}}$. Therefore, $\bar{\bar{z}}$ around $\nu_{\text{obs}} = \nu_{\text{ref}}$ is typically dominated by clusters, and $\bar{\bar{z}}\left(\nu_{\text{obs}}\right) \approx \bar{\bar{z}}$. As long as Eq. 3.24 holds, $\bar{\bar{z}}$ becomes insensitive to variations in $\mathcal{C}$, w_{f}, α_{f} and β_{f}; setting $z_{\max} = \infty$, just 3 significant parameters remain: l, d_1 and $\alpha_{\text{c}} + \beta_{\text{c}}$.

Finally, it is of significant interest to calculate $\bar{\bar{z}}$ for the filaments' SCW signal only,

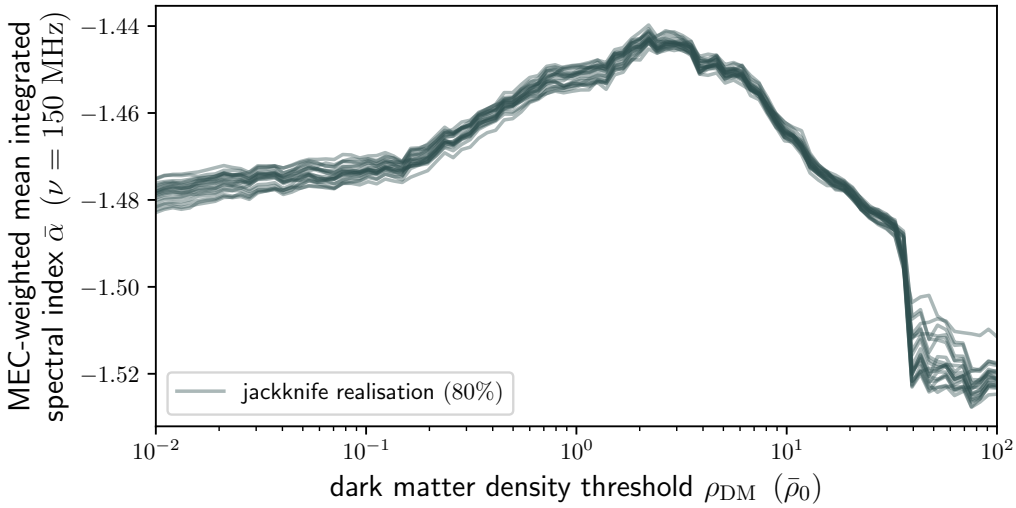


Figure 3.7: Monochromatic-emission-coefficient–weighted mean integrated spectral index $\bar{\alpha}$ at an emission frequency of $\nu = 150$ MHz, including all shocks that occur at locations where the dark matter (DM) density ρ_{DM} is below some threshold. Larger thresholds incorporate shocks from a wider range in ρ_{DM}, eventually including shocks from both filament and cluster environments. To provide an idea of the uncertainty in $\bar{\alpha}$, we use jackknife realisations that each contain 80% of the shocked voxels present in our Enzo snapshot at $z = 0.025$.

discarding the dominant SCW signal from merger and accretion shocks in galaxy clusters. To find a filament-only $\bar{\bar{z}}$ with the geometric model, we just set $\mathcal{C} = 0$. This simplifies the expression for $\mathbb{E}\left(X\left(\nu_{\mathrm{obs}}, z\right)\right)$, and reduces the number of relevant parameters considerably. When $\mathcal{C} = 0$, $\bar{\bar{z}}$ becomes not only independent of α_{c} and β_{c}, but also of w_{f} and ν_{obs}:

$$\bar{\bar{z}}\left(\nu_{\mathrm{obs}}\right) = \bar{\bar{z}} = \frac{\sum_{n=1}^{N} z_n \left(1 + z_n\right)^{-4+\alpha_{\mathrm{f}}+\beta_{\mathrm{f}}}}{\sum_{n=1}^{N} \left(1 + z_n\right)^{-4+\alpha_{\mathrm{f}}+\beta_{\mathrm{f}}}}. \tag{3.26}$$

As before, if we set $z_{\max} = \infty$, $\bar{\bar{z}}$ is a function of just three parameters: l, d_1 and $\alpha_{\mathrm{f}}+\beta_{\mathrm{f}}$, in this case.

3.3.5 Parameter estimates

This subsection provides concrete parameter estimates, which are necessary to evaluate the model in practice.

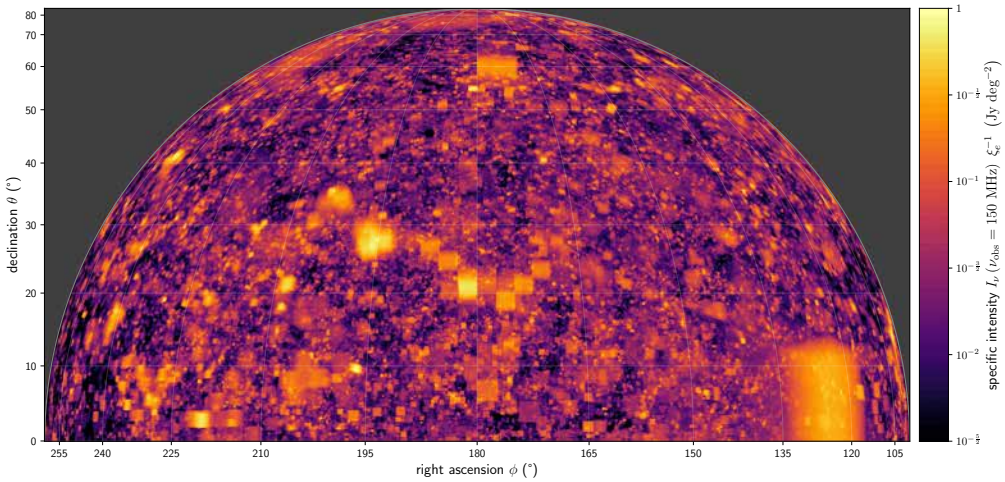
Cosmological simulations indicate $w_{\mathrm{f}} \sim 10^0 - 10^1$ Mpc, while $l \sim 10^1 - 10^2$ Mpc. The ratio of these parameters is constrained too: Table 3.1 shows that a VFF comparison between the geometric model and cosmological simulations favours $\frac{w_{\mathrm{f}}}{l} \sim 10^{-1}$.

From Fig. 3.3, we estimated $\mathcal{C} \sim 10^5 - 10^7$ for $\nu_{\rm ref} = 150$ MHz, assuming filament environments are characterised by $\rho \sim \bar{\rho}_0$, and cluster environments are characterised by $\rho \sim 10^2\,\bar{\rho}_0$ (at the $\sim$3 Mpc h^{-1} resolution, at least). Technically, Fig. 3.3 does not show $P(j_\nu\,|\,\rho)$ for $z = 0$ (the redshift at which $\mathcal{C}$ is defined to be the cluster-to-filament SCW MEC ratio), but for $z = 0.025$. However, for our purposes, the evolution of this relationship is likely of negligible importance.

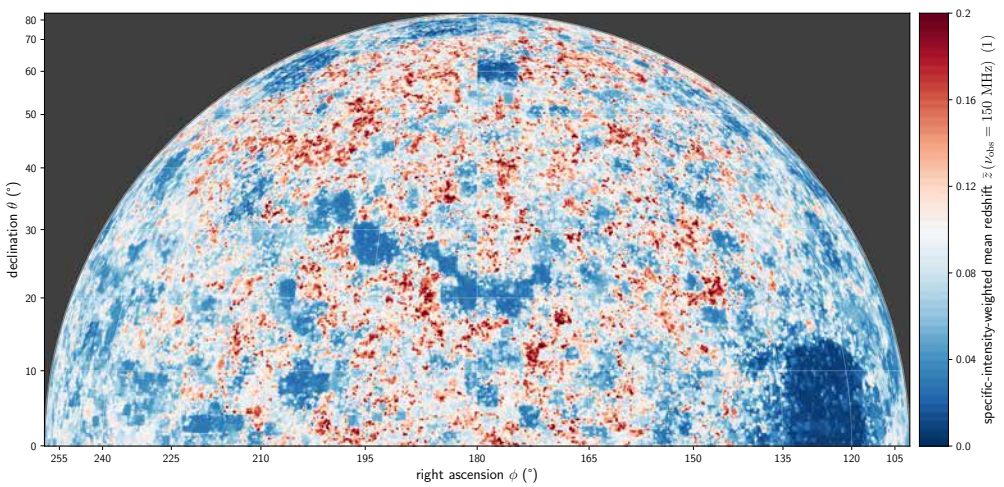
To find appropriate values for $\alpha_{\rm f}$ and $\alpha_{\rm c}$, we calculated the MEC-weighted mean integrated spectral index $\bar{\alpha}$ from our Enzo simulation snapshot as a function of shock environment. Fig. 3.7 shows the result when weighing by MECs at emission frequency $\nu = 150$ MHz; in general, $\bar{\alpha} = \bar{\alpha}(\nu)$. Nevertheless, it seems that $\alpha_{\rm f} = \alpha_{\rm c} = -3/2$ are reasonable choices.

Next, to find $\beta_{\rm f}$ and $\beta_{\rm c}$, we revisited the single-shock MEC expression of Eq. 3.2. Inverse Compton (IC) scattering to the CMB contributes a factor $(1+z)^{-4}$ to the single-shock MEC as $j_\nu \propto B_{\rm CMB}^{-2}$ and $B_{\rm CMB} \propto (1+z)^2$ (see Appendix 3.A2). Further factors of $(1+z)$ follow by considering the typical redshift evolution of the proper BM density $\rho_{\rm BM}$, proper magnetic field strength B, proper shock velocity v relative to the upstream IGM and *comoving* shock number density; all for both filament and cluster environments. In the linear regime of structure formation within an Einstein–de Sitter universe, the density contrast is proportional to $(1+z)^{-1}$. As the proper *mean* matter density is proportional to $(1+z)^3$, the proper BM density in filaments is expected to be proportional to $(1+z)^2$. In clusters, the density contrast grows more strongly; numerical simulations suggest that the density contrast is roughly proportional to $(1+z)^{-2}$. Likewise, this would imply that the proper BM density in clusters is proportional to $(1+z)$. Assuming that the magnetic field strength only evolves due to field line compression, we have $B \propto \rho_{\rm BM}^{2/3}$ (see Appendix 3.A2).[18] In filaments then, $B_{\rm f} \propto (1+z)^{4/3}$, whilst in clusters $B_{\rm c} \propto (1+z)^{2/3}$. As the single-shock MEC is proportional to $B^{1-\alpha}$, the proper magnetic field strength approximately contributes a factor $(1+z)^{10/3}$ for filaments, and a factor $(1+z)^{5/3}$ for clusters. Lacking further knowledge, we applied Occam's razor and assumed that the proper shock velocity and comoving shock number density do not evolve with redshift.

Collecting factors of $(1+z)$, we found $\beta_{\rm f} = 4/3$ and $\beta_{\rm c} = -4/3$. We stress that these values are highly uncertain and should be used as indications only.



(a)



(b)

Figure 3.8: Realisation of the merger- and accretion-shocked synchrotron Cosmic Web (MASSCW) priors at observing frequency $\nu_{\rm obs} = 150$ MHz. We show exactly ¼ of the total sky. *Top:* specific intensity I_ν. *Bottom:* specific intensity–weighted mean (cosmological) redshift $\bar{z}$.

3.4 Results

3.4.1 MASSCW specific intensity function distribution: general results

All the results given are for $\nu_{\rm obs} = 150$ MHz. At the BORG SDSS LSS reconstruction resolution, realisations from our MASSCW specific intensity function distribu-

[18] Ignoring magnetogenesis by outflows from AGN and supernovae, this Ansatz likely underestimates the actual magnetic field strength growth with time.

tion exhibit a factor $\sim 10^3$ of specific intensity variation over the sky, stemming from the highly variable power and localised nature of merger and accretion shocks in the Cosmic Web. These variations clearly appear in Fig. 3.8a, where we show a single realisation of $I_\nu(\hat{r}, \nu_{\rm obs})$ over the full lune. Some sharp-edged structures are visible: these are due to the voxelised nature of the BORG SDSS density field samples. The effect is most pronounced for high-MEC voxels close to the observer, which result in bright patches that span hundreds of square degrees.

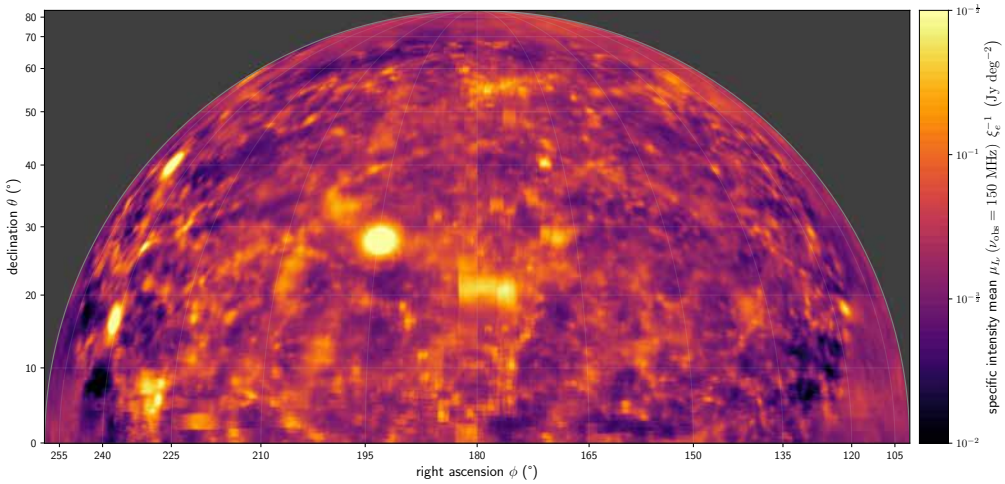
In Fig. 3.8b, we show the corresponding MASSCW specific intensity–weighted mean redshift function realisation $\bar{z}(\hat{r}, \nu_{\rm obs})$. All redshifts are within $[0, 0.2]$, because the BORG SDSS reconstructions are limited to $z_{\rm max} = 0.2$. A comparison with Fig. 3.8a reveals that generally, sky patches of high MASSCW specific intensity are due to structures at low redshift, and vice versa.

In Fig. 3.9a and 3.9b, we show the mean specific intensity function $\mu_{I_\nu}(\hat{r}, \nu_{\rm obs})$ and mean specific intensity–weighted mean redshift function $\mu_{\bar{z}}(\hat{r}, \nu_{\rm obs})$ at $\nu_{\rm obs} = 150$ MHz for each of the 10^6 ray-traced directions, pooling all 10^3 realisations. Both of these functions represent a summary statistic calculated from each ray's marginal I_ν- and $\bar{z}$-distribution (that is the distribution obtained by marginalising out — from the joint (prior) distribution — the RVs of all directions but one). We calculated the mean specific intensity after removing, for each marginal distribution separately, the lowest 1% and highest 1% of values.

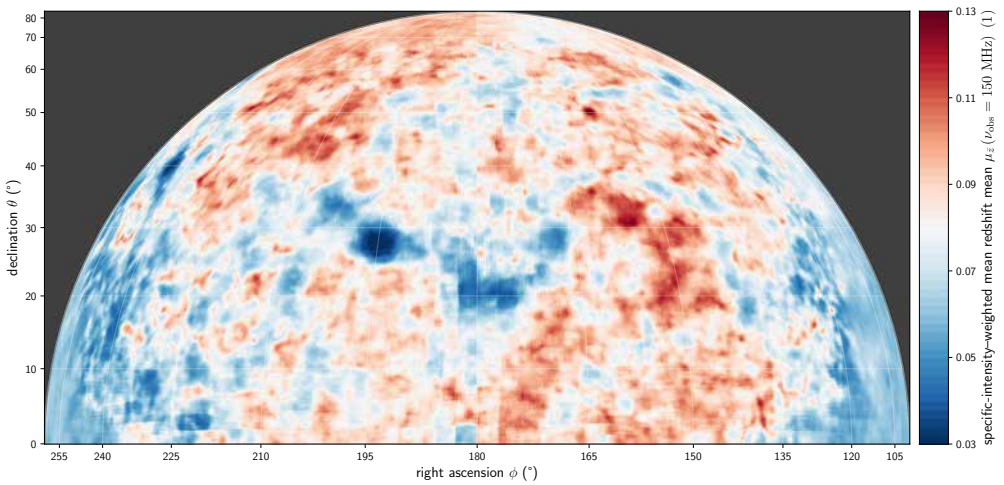
Three bright spots stand out; these are (in increasing order of declination) the Hercules Cluster, the Coma Cluster, and Abell 2199. Note that we have not included the specific intensity contribution of radio halos around galaxy clusters, which are of different origin: turbulent reacceleration (Brunetti et al., 2001; Petrosian, 2001). As observations suggest that the radio halo contribution usually dominates over the merger and accretion shock contribution, our results cannot be directly compared to actual galaxy cluster images.

The median specific intensity function $m_{I_\nu}(\hat{r}, \nu_{\rm obs})$ (not shown) strongly resembles the mean specific intensity function, but is smaller over the whole lune. This is indicative of skewed (single-direction) marginal specific intensity distributions. The $5^{\rm th}$ to $95^{\rm th}$ percentile mean-to-median ratios span the interval $(1.3, 2.1)$; the median mean-to-median ratio is 1.5.

Probabilistic approaches also allow for quantification of prediction uncertainty. In Fig. 3.14a and Fig. 3.14b, we present both an absolute and relative measure of spread, again calculated from the MASSCW specific intensity prior marginals. Fig. 3.14a shows the standard deviation $\sigma_{I_\nu}(\hat{r}, \nu_{\rm obs})$ after performing the same filtering as was done for the mean. The resulting function is highly similar to the mean: directions that are brighter on average also tend to have larger absolute prediction uncertainties.



(a)



(b)

Figure 3.9: Per-direction mean of the MASSCW priors at observing frequency $\nu_{\rm obs} = 150$ MHz, exhibiting less variability than the single realisation of Fig. 3.8a and 3.8b (note the smaller colour ranges). Three famous galaxy clusters stand out; these are (in increasing order of declination) the Hercules Cluster, the Coma Cluster, and Abell 2199. *Top:* specific intensity mean. *Bottom:* specific intensity–weighted mean (cosmological) redshift mean.

Of natural interest is also the relative prediction uncertainty $\sigma_{I_\nu}\mu_{I_\nu}^{-1}$, which we show in Fig. 3.14b. In this sense, the specific intensity function can be most accurately predicted around galaxy clusters and superclusters, reaching $\sigma_{I_\nu} \sim 60\%\, \mu_{I_\nu}$. As expected, the relative prediction uncertainty is highest outside the SDSS DR7 coverage (i.e. near the $\varphi = 90°$ and $\varphi = 270°$ edges of the lune). It is also high in regions

within the SDSS DR7 coverage that have low average brightness, inverting the trend that characterises the absolute prediction uncertainty.

3.4.2 MASSCW specific intensity function distribution: outskirts of massive galaxy clusters

Upon approaching the virial radius of a galaxy cluster from a connected filament, the WHIM transitions into the ICM, and the IGM's magnetic field grows stronger. Especially in models such as Hoeft & Brüggen (2007), in which the synchrotron-emitting electrons originate from the thermal pool, these cluster outskirts constitute the most promising targets to find synchrotron emission from the filament IGM. Therefore, we also show zoom-ins of three realisations from our MASSCW specific intensity function prior near three massive galaxy clusters: the Coma Cluster, the Hercules Cluster, and Abell 2199. In Fig. 3.10, we show I_ν over spherical domes with an angular radius of 12°, whilst Fig. 3.15 shows the specific intensity–weighted mean redshift for the same realisations and sky regions. For the Coma Cluster, two filaments are discernible: one in northeastern, and one running in western direction. For the Hercules cluster, the realisations suggest one northern and two southbound filaments. Finally, for Abell 2199 a southbound filament is evident.

By inspecting the mean redshift functions, one can verify that the identified structures indeed lie at the cluster redshift, rather than being structures closer by or further away that appear connected to the clusters in chance alignments.

3.4.3 MASSCW specific intensity function distribution: deep LOFAR HBA fields

With the LOFAR HBA, several deep ($\geq$ 50-hr) observations have been conducted that complement the wide-field, 8-hr approach of the LOFAR Two-metre Sky Survey (LoTSS; Shimwell et al., 2019). Under thermal-noise–limited conditions, deep fields are of prime interest to search for a signature of the filament SCW. For this reason, in Fig. 3.11, we show the MASSCW specific intensity prior single-direction (i.e. marginal) medians for three such deep fields.[19] We generate imagery on arbitrary Flexible Image Transport System (FITS) grids by first calculating the marginal medians for each traced ray, and then applying a Voronoi tessellation to achieve a prediction for every pixel.

The Lockman Hole (for which a 100-hr dataset is available (Tasse et al., 2021)) is a field known for its relatively low Milky Way column densities of neutral hydrogen

[19] For these close-ups, we present the median instead of the mean to emphasise that our probabilistic approach enables the calculation of a variety of useful summary statistics.

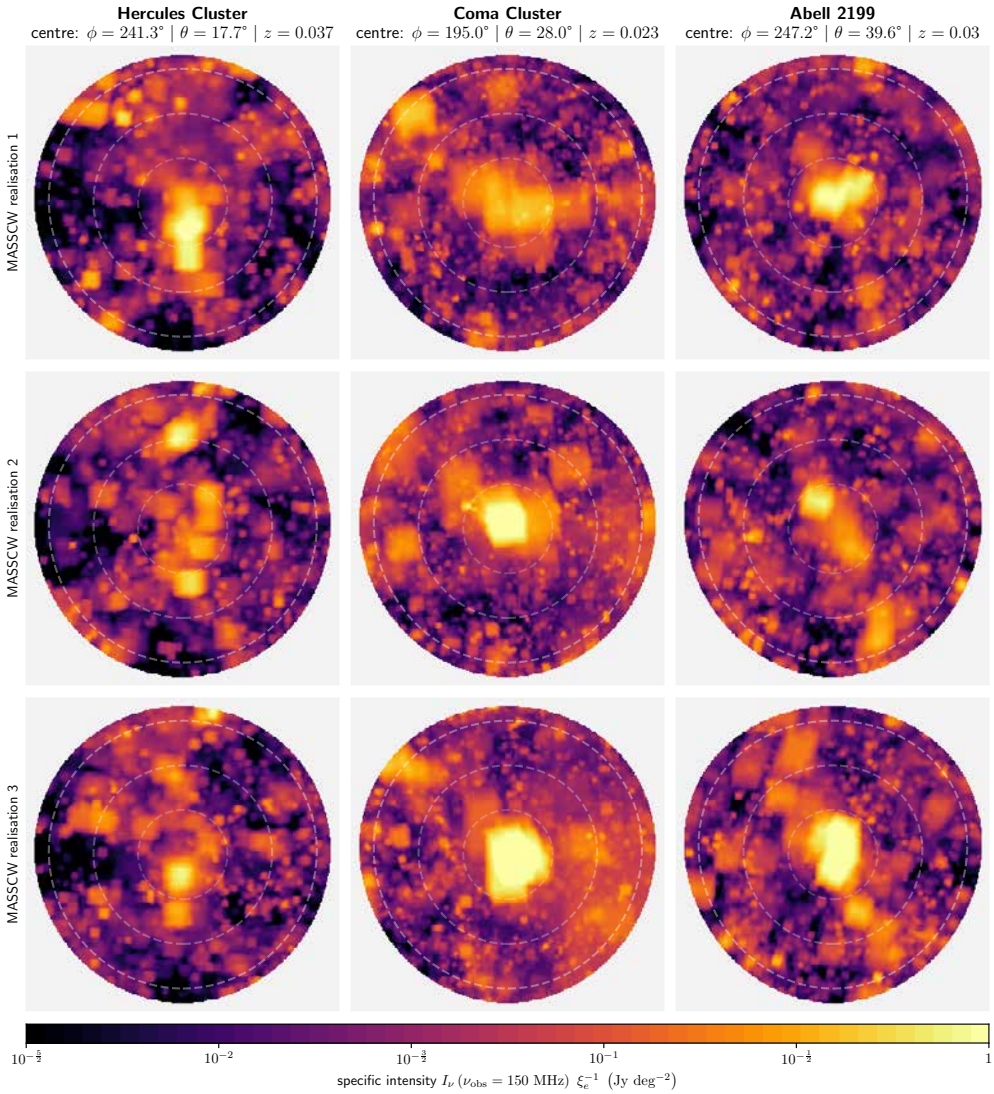


Figure 3.10: Three realisations (rows) of the merger- and accretion-shocked synchrotron Cosmic Web (MASSCW) specific intensity prior at observing frequency $\nu_{\rm obs} = 150$ MHz, showing zoom-ins around three massive Northern Sky galaxy clusters (columns). The zoom-ins show caps of the celestial sphere of 10° radius. The dashed circles are at 3°, 6° and 9° from the cluster centre. Note that the usual radio halos that permeate galaxy clusters are not shown; these are caused by turbulent reacceleration, and we only show the merger and accretion shock contribution. The ∼3 Mpc h^{-1} resolution of the 3D total matter density and monochromatic emission coefficient (MEC) fields limits the resolution of the 2D specific intensity fields, so that individual shocks and ∼1 Mpc MEC - total matter density (anti)correlations are not resolved.

and dust (Lockman et al., 1986), making it favourable for study in the extreme UV and soft X-ray bands, amongst others. The potential for multi-wavelength synergy

has made deep observations of the Lockman Hole a LOFAR HBA priority (Mahony et al., 2016). Abell 2255 (for which a 75-hr dataset is available (Botteon et al., 2022)) is a merging galaxy cluster (Feretti et al., 1997; Botteon et al., 2020a) at $z = 0.08$ that is part of the North Ecliptic Pole (NEP) Supercluster (Mullis et al., 2001; Shim et al., 2011). In this dynamic environment, merger and accretion shocks could light up the filaments connected to Abell 2255 (Pizzo et al., 2008). Lastly, the Ursa Major Supercluster (for which a 50-hr dataset is being assembled) at $z = 0.06$ stands out as the most prominent structure in our MASSCW specific intensity prior after the galaxy clusters shown in Fig. 3.10 and 3.15. For this reason, it has been selected for deep LOFAR HBA observations.

3.4.4 Specific intensity–weighted mean redshift

In the top panel of Fig. 3.12, we put our geometric model to the test by comparing a predicted specific intensity–weighted mean redshift CDF $F_{\bar{Z}}(\bar{z})$ (green dotted line) to 10^2 randomly drawn ECDFs from our 10^3 MASSCW prior realisations (translucent black lines), at $\nu_{\rm obs} = 150$ MHz. Both prior realisations and the model reach a cumulative probability ('fraction of the sky') of 1 at $\bar{z} = 0.2$: the BORG SDSS samples do not feature LSS beyond $z_{\rm max} = 0.2$ and the model is restricted accordingly. The model CDF is constructed from 10^5 draws from RV $\bar{Z}$ (see Eq. 3.20).[20] We adopt parameters suggested by Sect. 3.3.5: $w_{\rm f} = 5$ Mpc, $l = 50$ Mpc, $d_1 = 25$ Mpc, $\mathcal{C} = 10^6$ at $\nu_{\rm ref} = 150$ MHz, $\alpha_{\rm f} = \alpha_{\rm c} = -3/2$, $\beta_{\rm f} = 4/3$ and $\beta_{\rm c} = -4/3$. Without further parameter tuning, the model CDF reproduces the trend revealed by the majority of prior ECDFs. This correspondence is evidence that the geometric model in general (and not just for some highly specific choice of parameters) captures the main features of the MASSCW, and motivates calculating the distribution of $\bar{Z}$ for the *true* sky, which features LSS beyond $z_{\rm max} = 0.2$. For the observational study of filaments, it is of greatest interest to calculate $F_{\bar{Z}}(\bar{z})$ when the cluster contribution to the MASSCW is ignored, as it would otherwise dominate. In the bottom panel of Fig. 3.12, we therefore show $F_{\bar{Z}}(\bar{z})$ according to the geometric model for $z_{\rm max} = \infty$ and $\mathcal{C} = 0$. When $\mathcal{C} = 0$, $\alpha_{\rm c}$ and $\beta_{\rm c}$ become irrelevant. Four parameters remain: $\alpha_{\rm f} + \beta_{\rm f}$, l, $w_{\rm f}$ and d_1. The first of these has by far the most effect on $F_{\bar{Z}}(\bar{z})$, as evidenced by the differently coloured curves (we keep $\alpha_{\rm f} = -3/2$ constant and vary $\beta_{\rm f}$, but varying $\alpha_{\rm f}$ and keeping $\beta_{\rm f}$ constant would lead to identical results). We also vary the purely geometric parameters l, $w_{\rm f}$ and d_1, but in such a way that $w_{\rm f} = \frac{1}{10} l$ (as suggested by comparing the model's VFFs to those from cosmological simulations)

[20] In a strict sense, given the finite number of realisations in our numerical approximation, this function is also an ECDF. However, for all practical purposes, it can be regarded as a CDF.

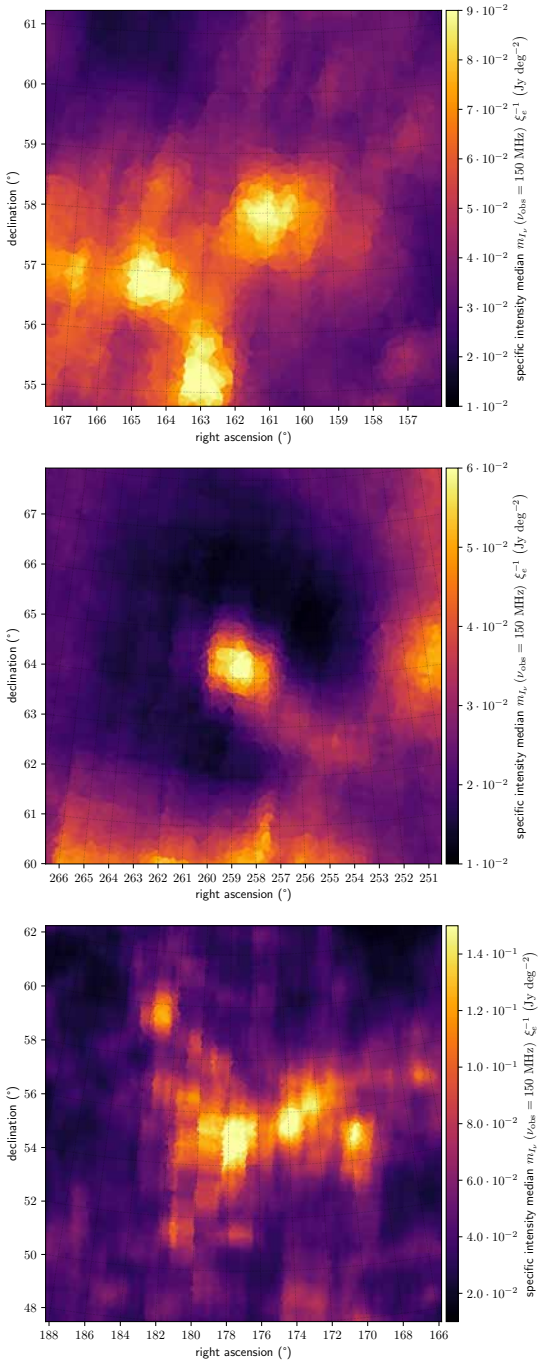


Figure 3.11: Merger- and accretion-shocked synchrotron Cosmic Web (MASSCW) specific intensity prior marginal medians at observing frequency $\nu_{\rm obs} = 150$ MHz for three deep LOFAR HBA fields. Individual shocks should not be discernible in these statistical aggregates. The colour scales share the same lower bound. *Top:* Lockman Hole. *Middle:* Abell 2255. *Bottom:* Ursa Major Supercluster.

and $d_1 = \frac{1}{2}l$ (i.e. the observer is always put at the centre of a cubic unit cell). When $l \in \{50\text{ Mpc}, 75\text{ Mpc}, 100\text{ Mpc}\}$ is increased, and w_f and d_1 accordingly, modest changes in $F_{\bar{Z}}(\bar{z})$ occur (dash-dotted and dotted lines): the distribution of $\bar{Z}$ attains larger spread.

$F_{\bar{Z}}(\bar{z})$ provides detailed information. For example, the model with $\beta_\mathrm{f} = 4/3$ and $l = 50$ Mpc predicts that for $\sim 50\%$ of the sky ($\sim$21, 000 sq. deg.), the MASSCW signal has a mean redshift of 0.35 or lower, whilst for $\sim 80\%$ of the sky ($\sim$33, 000 sq. deg.), the mean redshift is 0.42 or lower.

3.4.5 Flux density–weighted mean redshift

The flux density–weighted mean redshift $\bar{\bar{z}}$ constitutes the most concise measure of the typical MASSCW signal epoch. In the top panel of Fig. 3.13, we show a distribution (shaded grey) over $\bar{\bar{z}}$ generated via KDE (Gaussian kernel, $\sigma_\mathrm{KDE} = 4 \cdot 10^{-4}$) from our 10^3 MASSCW prior samples. The median is $\bar{\bar{z}} = 0.077$ (solid grey); however, we stress that this is because only LSS up to $z_\mathrm{max} = 0.2$ is included. We present an example model (solid green) that reproduces the median $\bar{\bar{z}}$, with main parameters $z_\mathrm{max} = 0.2$, $l = 75$ Mpc, $d_1 = 60$ Mpc, $\alpha_\mathrm{c} = -3/2$ and $\beta_\mathrm{c} = -4/3$. (Adopting $\mathcal{C} = 10^6$ at $\nu_\mathrm{ref} = 150$ MHz as suggested by Sect. 3.3.5, clusters dominate over filaments, and the other parameters play a very minor role.) To explore the sensitivity of $\bar{\bar{z}}$ on the parameters, we generate $5 \cdot 10^5$ parameter sets by drawing from wide uniform distributions. We draw $l \sim \text{Uniform}\,(50\text{ Mpc}, 100\text{ Mpc})$, $d_1 \sim \text{Uniform}\,(0, l)$, $w_\mathrm{f} \sim \text{Uniform}\,(5\text{ Mpc}, 10\text{ Mpc})$, $\log_{10}(\mathcal{C}) \sim \text{Uniform}\,(5, 7)$, $\alpha_\mathrm{f}, \alpha_\mathrm{c} \sim \text{Uniform}\,(-3/2 - 1/4, -3/2 + 1/4)$, $\beta_\mathrm{f} \sim \text{Uniform}\,(4/3 - 2, 4/3 + 2)$ and $\beta_\mathrm{c} \sim \text{Uniform}\,(-4/3 - 2, -4/3 + 2)$. However, the resultant distribution for $\bar{\bar{z}}$ is mostly restricted to the relatively small range of $(0.06, 0.08)$. We stress that the two distributions are conceptually distinct and are not meant to be compared directly.

In the bottom panel of Fig. 3.13, we show $\bar{\bar{z}}$ for filaments only ($\mathcal{C} = 0$) and $z_\mathrm{max} = \infty$. The prediction then becomes a function of three parameters: $\alpha_\mathrm{f} + \beta_\mathrm{f}$, l and d_1. The dependencies on both l and d_1 are weak, and $\bar{\bar{z}}$ is thus, to good approximation, determined by just a *single* parameter. Unfortunately, this parameter is currently highly uncertain, and should be constrained with upcoming cosmological simulations. The various coloured curves suggest that the impact of d_1 on $\bar{\bar{z}}$ increases when l is larger. As the unit cells grow (so that their number density drops $\propto l^{-3}$), $\bar{\bar{z}}$ becomes more sensitive to contributions of individual boundary crossings (filaments), including the first. Filament specific intensity contributions decrease quickly (especially for low values of $\alpha_\mathrm{f} + \beta_\mathrm{f}$) and monotonically with distance, so that the distance of the first filament to the observer (e.g. nearby, $d_1 = 10\%\, l$, or far away, $d_1 = 90\%\, l$) is of some

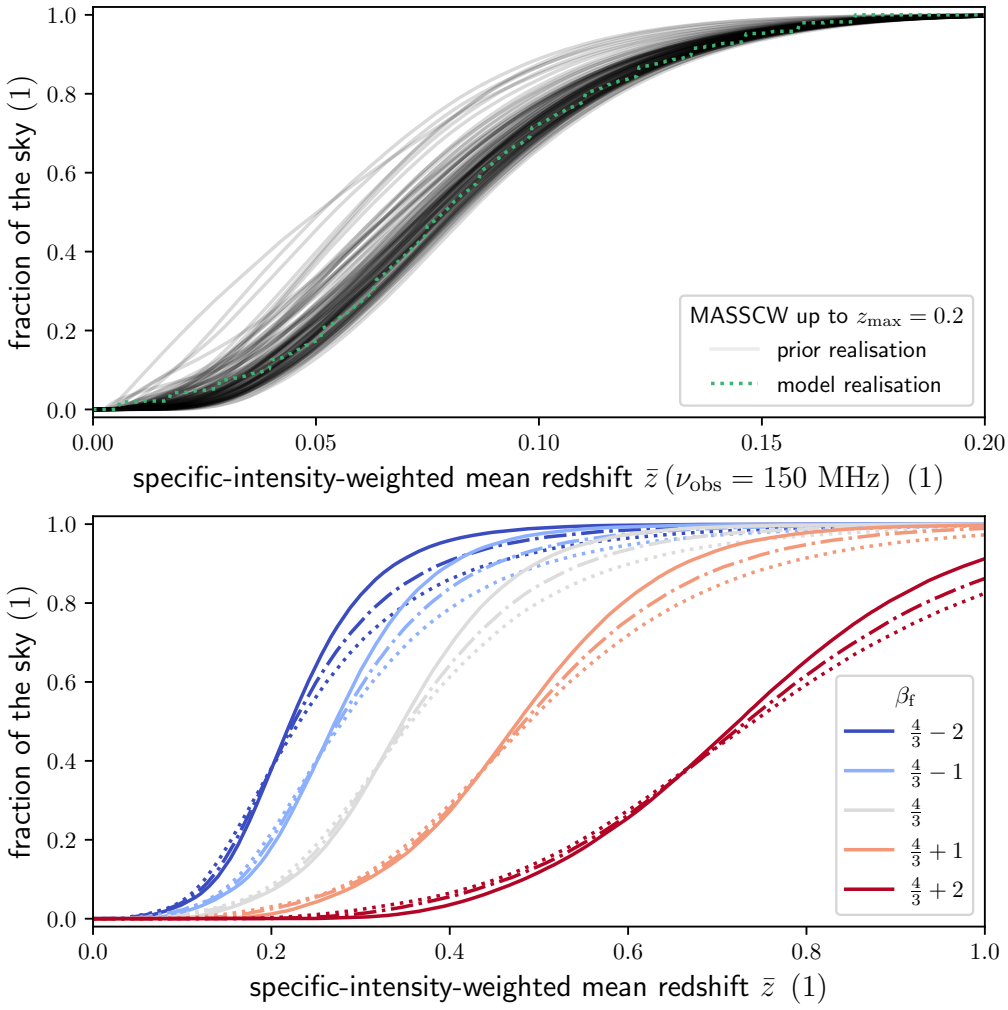


Figure 3.12: MASSCW specific intensity–weighted mean redshift RV $\bar{Z}$ results. *Top:* distributions of $\bar{Z}$, showing 100 randomly selected ECDFs from our prior (grey curves), and a geometric model CDF (green curve). Both data and model consider LSS up to redshift $z_{\max} = 0.2$ only. The other geometric model parameters are: $w_{\rm f} = 5$ Mpc, $l = 50$ Mpc, $d_1 = 25$ Mpc, $\mathcal{C} = 10^6$, $\alpha_{\rm f} = \alpha_{\rm c} = -3/2$, $\beta_{\rm f} = 4/3$ and $\beta_{\rm c} = -4/3$. *Bottom:* geometric model CDFs of $\bar{Z}$ for filaments only ($\mathcal{C} = 0$), though including all of them ($z_{\max} = \infty$). By far the most influential parameter is also the most uncertain one; we therefore vary $\beta_{\rm f}$ over a plausible range. We also vary $l \in \{50\text{ Mpc}, 75\text{ Mpc}, 100\text{ Mpc}\}$, and by extension also $w_{\rm f}$ and d_1, by forcing $w_{\rm f} = \frac{1}{10} l$ and $d_1 = \frac{1}{2} l$. As l increases, the unit cell number density decreases; we symbolise this with more sparsely drawn curves. We keep $\alpha_{\rm f} = -3/2$ constant.

importance.

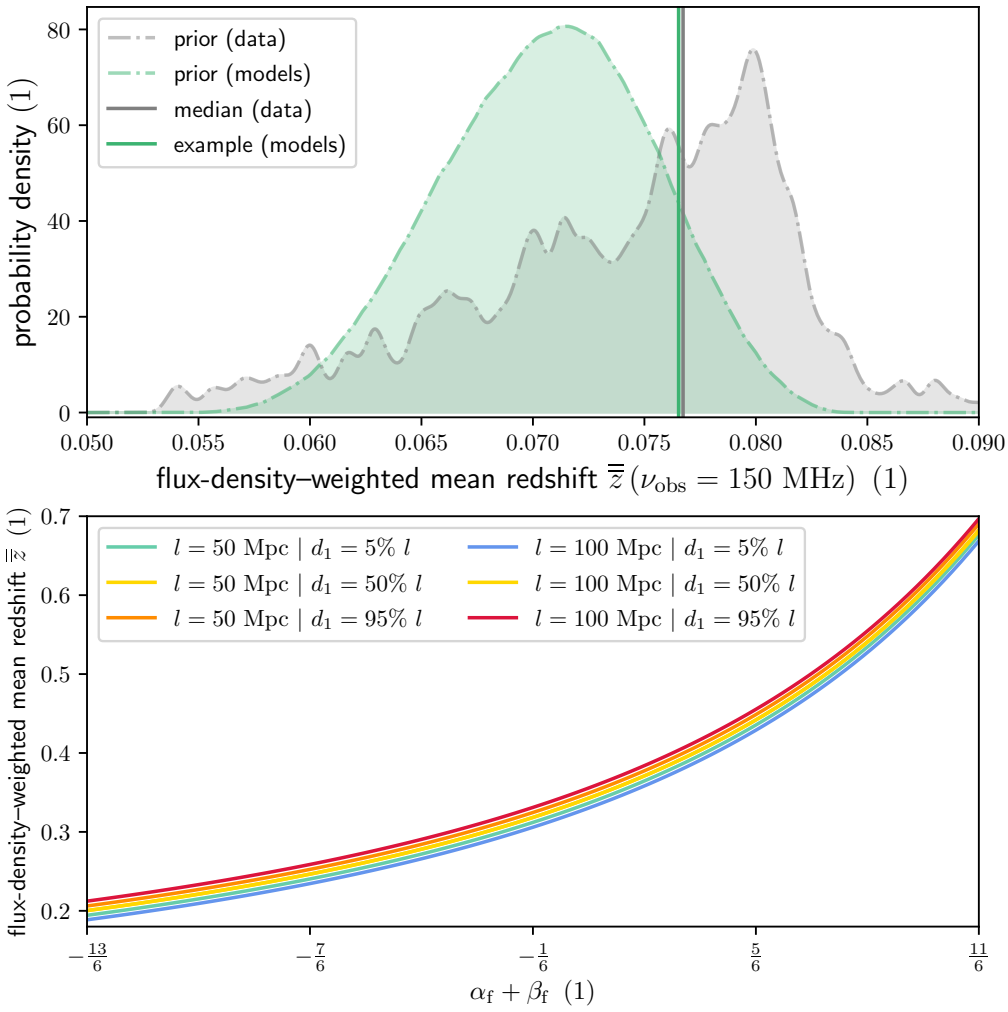


Figure 3.13: MASSCW flux density–weighted mean redshift $\bar{\bar{z}}$ results. *Top:* distribution of $\bar{\bar{z}}$ calculated via KDE from our prior's 10^3 realisations of I_ν and $\bar{z}$ (shaded grey), which are limited to $z_{\rm max} = 0.2$. The median is $\bar{\bar{z}} = 0.077$ (solid grey). Furthermore, we show the variety of geometric model predictions for $\bar{\bar{z}}$ when similarly limited to $z_{\rm max} = 0.2$, assuming flat priors on all model parameters (shaded green; see text for details). The median is $\bar{\bar{z}} = 0.072$. As a concrete example, we show the flux density–weighted mean redshift of a single model that reproduces the data median (solid green). This model is given by $l = 75$ Mpc, $d_1 = 60$ Mpc and $\beta_{\rm c} = -4/3$. *Bottom:* predictions for $\bar{\bar{z}}$ considering filaments *only*, varying all three relevant geometric model parameters: $\alpha_{\rm f} + \beta_{\rm f}$, l and d_1. The two $d_1 = 50\%\ l$ curves are virtually indistinguishable.

3.5 Discussion

3.5.1 Inter-cluster radio bridges: evidence for turbulence rather than shocks

In the last few years, LOFAR HBA observations around 140 MHz have revealed massive structures connecting clusters of galaxies at the onset of merging. These structures, or 'bridges', likely are compressed filaments of the Cosmic Web. The first *tentative* discovery (which has since been confirmed) of an inter-cluster bridge was in Abell 1758 (of length $\sim$2 Mpc) (Botteon et al., 2018, 2020b), whilst the first unambiguous detection was in Abell 399–401 (of length $\sim$3 Mpc) (Govoni et al., 2019). The latter ridge features higher densities ($\sim 3 \cdot 10^2$ m^{-3}), temperatures ($\sim 7 \cdot 10^7$ K), and magnetic field strengths ($\sim$1 μG) than are expected to exist in typical filaments.

These two inter-cluster bridges cannot be, and are not expected to be, faithfully reproduced by our MASSCW prior. Firstly, both bridges do not lie in the reconstructed volume: Abell 1758 lies at a redshift of $z = 0.279$, beyond the current redshift range ($z < 0.2$) of BORG SDSS reconstructions; Abell 399–401 has a favourable redshift of $z = 0.07$, but falls outside of the SDSS DR7 footprint. Another complication is that bridges of $2-3$ Mpc extent are smaller than the current BORG SDSS resolution of $\sim$3 Mpc h^{-1}, and would therefore only barely be identifiable if they fell within the reconstructed volume.

Notwithstanding practical difficulties, Govoni et al. (2019) show that the Hoeft & Brüggen (2007) model alone cannot explain the LOFAR HBA data. Significantly, this discovery has cast doubt on the widespread hypothesis — also adopted in this article — that accretion shocks generate the dominant contribution to the filaments' SCW signal. Other mechanisms, like turbulence (Brunetti & Vazza, 2020), could play an important role. Cosmological simulations that model turbulence in the sense of Brunetti & Vazza (2020) that would allow a recalculation of the SCW prior under this alternative emission mechanism, do not yet exist. We stress that our methods, although presently employed to generate MASSCW predictions, are of general nature, and can be used to explore alternative (combinations of) SCW emission mechanisms in the future.

3.5.2 Generalisation to other wavelength bands

We have presented our methodology in the context of predicting the Cosmic Web's contribution to the *radio* sky. However, the methodology does not rely on any special property of radio emission. Apart from comparing different emission mechanisms for the *same* wavelength band (as discussed in the previous subsection), the methodol-

ogy could also be extended to predicting the Cosmic Web's contribution to the sky in *other* wavelength bands. With missions such as XRISM and Athena on the horizon, an extension into the X-ray window is of prime interest. As a demonstration, we have tentatively calculated the thermal bremsstrahlung component of the Cosmic Web's contribution to the X-ray sky. For plain simplicity, we do not invoke snapshots from cosmological simulations, but rather assume the proportionalities $\rho_{\mathrm{BM}}(\rho) = \frac{\Omega_{\mathrm{BM},0}}{\Omega_{\mathrm{M},0}}\rho$ and $T(\rho_{\mathrm{BM}}) = T_{\mathrm{ref}}\frac{\rho_{\mathrm{BM}}}{\rho_{\mathrm{BM,ref}}}$ (with $T_{\mathrm{ref}} = 10^6$ K and $\rho_{\mathrm{BM,ref}} = m_p\ \mathrm{m}^{-3}$, where m_p is the proton rest mass). Furthermore, we assume a hydrogen-helium plasma of primordial chemical composition, and use the Gaunt factor regimes from Novikov & Thorne (1973). We show results in Fig. 3.16, for exactly the same sky regions as in Fig. 3.11. Note that although the existence of a thermal bremsstrahlung component from the WHIM to the X-ray sky is uncontested and features relatively low uncertainties, it is likely that the oxygen line emission component dominates. We have not considered this component due to its relative complexity and distance to our work's main focus, but propose its calculation as a promising direction for future research.

3.5.3 Total matter density reconstructions: resolution, coverage and depth

The voxelised nature of density reconstructions like the BORG SDSS causes large-scale ($\sim$10°) blocky shapes in specific intensity and specific intensity–weighted mean redshift function realisations that do not represent plausible real-life morphologies. The angular scales at which these discontinuities occur, depend on the distances between the responsible (high-MEC) voxels and the observer. Future density reconstructions that are run at higher resolution will contain fewer of these problematic patterns. Furthermore, as hinted at in the previous paragraph, a modest resolution improvement (i.e. by a factor of order unity) would allow MASSCW predictions to contain inter-cluster bridges.

An improved resolution would also give relevance to the method of generating spatially correlated MEC fields as described in Sect. 3.2.4. Currently, the BORG SDSS voxel length is larger than the shock correlation length $l_{\mathrm{SE}} = 2$ Mpc, so that the MEC fields obtained by our method are only marginally different from MEC fields where each voxel's MEC is drawn independently from the appropriate conditional probability distribution.

A promising idea that is not pursued in this work, is to interpolate the density fields (by e.g. doubling or tripling the number of voxels along each dimension), so that it becomes possible to generate merger and accretion shocks of the appropriate size. Note that this approach would not add new small-scale structure to the density

reconstructions, but merely ensures that the size of shocks generated stochastically are determined by the length scale l_{SE}, instead of by the density reconstruction resolution.

A straightforward next step would be to use density reconstructions of the remaining $3/4$ of the sky to complete the MASSCW prior: generating predictions for the Southern Sky is relevant for SKA searches of the filament SCW. This could be done with the already available BORG 2M++ (Jasche & Lavaux, 2019), although these reconstructions remain shallower than those of the BORG SDSS, with $z_{\mathrm{max}} = 0.1$ instead of $z_{\mathrm{max}} = 0.2$.

Reconstructions that push beyond $z_{\mathrm{max}} = 0.2$ would also improve our predictions. The yield of such an extension depends on the filament-only specific intensity–weighted mean redshift CDF $F_{\bar{Z}}(\bar{z})$ for $z_{\mathrm{max}} = \infty$. However, the single most influential parameter governing $F_{\bar{Z}}(\bar{z})$, $\alpha_{\mathrm{f}}+\beta_{\mathrm{f}}$, is currently ill-constrained.[21] We therefore do not yet know to what extent deeper reconstructions would improve our MASSCW priors.

3.5.4 Cosmological simulations

The MASSCW predictions presented in this work are based on the statistical relationship between total matter density and the MASSCW MEC as inferred from Enzo cosmological simulations. However, in absence of tight observational constraints, these cosmological simulations must assume one of many magnetogenesis scenarios. The snapshots used in this work assume a primordial magnetogenesis scenario, starting from $z = 45$ with a seed magnetic field with a uniform comoving strength of 0.1 nG. More complex spectral energy distributions of primordial magnetic fields are however not excluded by present constraints from the CMB (e.g. Vazza et al. (2021b) and references therein). Future work should explore the effect of different choices in the magnetogenesis scenario landscape on the MASSCW predictions, which could see a systematic change in specific intensity by an order of magnitude for filament-dominated directions.

A minor additional uncertainty comes from the fact that the $(100\ \mathrm{Mpc})^3$ cube used in this work is not yet large enough to fully capture the density–MASSCW MEC relation in a statistically exhaustive manner. This modest problem could be alleviated by appending the joint probability distribution from which the conditional probability distributions shown in Fig. 3.3 are derived with data from more simulation runs.

Finally, in this work, RAM limitations have necessitated discarding shocks with upstream Mach numbers below 2 in our shock identification procedure. This Mach number cut means that the MECs assigned to voxels are lower bounds. Future in-

[21] Interestingly, its determination could be done with existing cosmological simulations.

clusion of the shocks now omitted could increase the MASSCW specific intensity functions by a (direction-dependent) factor of order unity.

3.5.5 Observational considerations

Using our predictions, we can evaluate the observational prospects of detecting the filament SCW — and the filament MASSCW in particular. As evinced by radio bridge detections (Botteon et al., 2018; Govoni et al., 2019; Botteon et al., 2020b) and a statistical all-sky (or close to all-sky) detection (Vernstrom et al., 2021), both special-geometry *and* global observations of the filament SCW are already possible with modern low-frequency radio telescopes. The more interesting question therefore is whether detections are possible on an intermediate level — that of individual regions of large-scale structure — so that spatially resolved measurements of the intergalactic magnetic field strength $B_{\rm IGM}$ in filaments come within reach. In this section, we therefore explore the observational prospects and challenges of detecting the filament SCW around individual massive galaxy clusters, such as the Hercules Cluster, the Coma Cluster, Abell 2199 and Abell 2255, and around larger LSS complexes, such as the cluster triple in the Lockman Hole and the Ursa Major Supercluster. We assume that the MASSCW signal is the dominant contributor (in fact, the *only* contributor) to the SCW signal of these LSS regions.[22] For associated MASSCW predictions, see Fig. 3.10 and 3.11.

Making firm observability forecasts is in the first place hampered by the fact that there are three major unknowns in MASSCW predictions: firstly, the strength of the seed magnetic field in the Early Universe combined with the dominant process by which this field has evolved into magnetic fields in filaments today (see Sect. 3.1); secondly, the filling factor of shocks of appropriate strength and obliquity to trigger DSA; and, thirdly, the magnitude of the electron acceleration efficiency ξ_e in filaments (see Sect. 3.2.3). From Fig. 3.10, we see that the specific intensity function in the direction of filaments around massive low-redshift galaxy clusters, at $\nu_{\rm obs} = 150\,{\rm MHz}$ and degree-scale resolution, reaches $I_\nu \sim 10^{-1}\,{\rm Jy\,deg^{-2}}$ for $\xi_e = 1$. If $\xi_e \sim 10^{-2}$ in filaments, as in Keshet et al. (2004)'s SNR shocks, the actual specific intensity is $I_\nu \sim 10^{-3}\,{\rm Jy\,deg^{-2}}$. A conservative estimate of the uncertainty in ξ_e is an order of magnitude, yielding a range of specific intensity estimates $I_\nu \sim 10^{-4}$–$10^{-2}\,{\rm Jy\,deg^{-2}}$.

In this work, we have considered the MASSCW signal without contaminants. However, in most directions, the Milky Way is the dominant contributor to the specific

[22] As discussed in Sect. 3.5.1, the MASSCW signal is unlikely to be the dominant contributor to the SCW signal *in radio bridges*; in the case of the all-sky search of Vernstrom et al. (2021), the dominant emission mechanism remains unknown.

intensity function at $\nu_{\rm obs} \sim 10^2$ MHz, and actual observational attempts to detect the SCW should therefore avoid the Galactic Plane, the North Polar Spur (Salter, 1983), and other bright synchrotron Milky Way features. The extent to which the Milky Way hampers a SCW detection, depends on the typical angular scales of the synchrotron Milky Way, the SCW, and those measurable by the interferometer. Oei et al. (prepa) have made compact-source–subtracted, low-resolution (60″ and 90″) images with the LoTSS DR2 (Shimwell et al., 2022) that reveal the Milky Way's specific intensity function at $\nu_{\rm obs} = 144$ MHz up to degree scales. (By the lack of baselines shorter than 68 metres, larger scales are resolved out.) These images have most power on the degree scale and show specific intensity variations of $\sigma_{I_\nu} \sim 10^0$–10^1 Jy deg^{-2} in the off-Galactic plane region. Meanwhile, the scales at which a sky region's SCW specific intensity function has most power depend on the distances to the region's most massive large-scale structures and vary per dominant emission mechanism assumed, with power on larger scales for turbulence compared to merger and accretion shocks. A discrepancy in dominant scales between the specific intensity functions of the synchrotron Milky Way and the SCW can be leveraged to bolster the prospects of a SCW detection. Doing so appears important, because already at the 8-hr depth of the LoTSS DR2, the Milky Way's specific intensity dominates over the thermal noise. By using an inner (u, v)-cut that *removes* most Milky Way emission (but that *retains* most SCW emission), Milky Way–induced specific intensity variations can be reduced by one to two orders of magnitude,[23] leaving a Milky Way contamination of

[23] This estimate follows from considering the spherical harmonics angular power spectrum C_ℓ of the interferometrically observed (and thus large-angular-scale–deprived) synchrotron Milky Way. We model C_ℓ as a power law (La Porta et al., 2008) from degree ℓ_0 onwards:

$$C_\ell(\ell) = \begin{cases} 0 & \ell < \ell_0; \\ C_\ell(\ell_0)\left(\frac{\ell}{\ell_0}\right)^\beta & \ell \geq \ell_0, \end{cases} \tag{3.27}$$

where $\beta < -1$. Generally, given an angular power spectrum, the total power P is

$$P = \sum_{\ell=0}^{\infty} C_\ell(\ell)\,. \tag{3.28}$$

We model the imposition of an additional inner (u, v)-cut as the removal of power on all angular scales up to (but excluding) ℓ_0', with $\ell_0' \geq \ell_0$. The (negative) relative change in total power caused by the (u, v)-cut is

$$\frac{P' - P}{P} = \frac{\sum_{\ell_0'}^{\infty} \ell^\beta}{\sum_{\ell_0}^{\infty} \ell^\beta} - 1 = \frac{\zeta(-\beta, \ell_0')}{\zeta(-\beta, \ell_0)} - 1, \tag{3.29}$$

where ζ is the Hurwitz zeta function. For $\beta \in [-3, -2]$ (e.g. La Porta et al., 2008; Ghosh et al., 2012;

$\sigma_{I_\nu} \sim 10^{-2}$–$10^{0}$ Jy deg^{-2}. (The optimal (u, v)-cut choice for a particular sky region can be derived by comparing the angular power spectrum of the synchrotron Milky Way with that of the region's SCW predictions.) For depths such that the Milky Way remains the dominant noise source, the signal-to-noise ratio for a solid angle of a square degree centred around a massive filament thus is 10^{-4} in a very pessimistic case, 10^{-1} in a fairly optimistic case, and 10^{0} in a very optimistic one. The signal-to-noise ratio grows with the square root of the number of such solid angles considered, and thus linearly with the angular diameter of the observed region. Notably, in the fairly optimistic case, we need to observe a region of 10^{1} degree diameter to achieve a signal-to-noise ratio of order unity (for example three). For example, the Ursa Major Supercluster (see the bottom panel of Fig. 3.11) is a region of roughly the required extent. In case the MASSCW specific intensity is an order of magnitude weaker, that is of order $I_\nu \sim 10^{-3}$ Jy deg^{-2}, the sky region required to detect the MASSCW is roughly the entire sky.

Galaxies populate filaments and generate synchrotron radiation. If they are not masked or removed from the observed imagery, their presence could mimic a SCW detection signal in cross-correlation experiments with low-spatial-resolution MASSCW predictions. (However, at high spatial resolution, Hodgson et al. (2021) show that the SCW and synchrotron emission from galaxies can be separated, as these signals trace the LSS matter distribution in different ways.) This issue underlines the importance of deep observations that feature low thermal noise levels, although — in contrast to the usual situation — in order to minimise systematic rather than random errors. As noted before, for LoTSS DR2 observations, Milky Way contamination dominates over thermal noise, rendering thermal noise largely irrelevant — at least prior to the suppression of degree scales. However, low thermal noise levels do allow for a more thorough subtraction of the galaxy population in filaments and thus help control an important systematic effect.

3.5.6 Independence of random fields

In this work, we have outlined how a specific intensity random field $I_\nu(\hat{r}, \nu_{\rm obs})$ can be generated from a percentile random field $\mathcal{P}(\mathbf{r})$ (or, equivalently, a Gaussian random field $\mathcal{Z}(\mathbf{r})$) and a total matter density random field $\rho(\mathbf{r})$; $\hat{r} \in \mathbb{S}^2$, $\nu_{\rm obs} \in \mathbb{R}_{>0}$, and $\mathbf{r} \in \mathcal{R}$. The function f that maps the two input random fields to the output random field is deterministic, informed by conditional probability distributions extracted from cosmological simulations, and non-linear in both arguments. Symbolically, $I_\nu = f(\mathcal{P}, \rho)$. Implicitly, our approach has been to sample ρ from the BORG

Sims et al., 2016), $\ell_0 \sim 10^2$ and $\frac{\ell_0'}{\ell_0} \sim 10^0$, we find relative total power changes of -90% to -99%.

SDSS posterior, and $\mathcal{P}$ from another, *independent* distribution (which is fixed by the distribution of $\mathcal{Z}$, and thus specified by the covariance function K_{SE}). There is a physical scenario in which this is justified. The joint distribution for the input random fields $P(\mathcal{P}, \rho)$ can be written as the product of a conditional and marginal: $P(\mathcal{P}, \rho) = P(\mathcal{P} \mid \rho)\, P(\rho)$. Thus, the central assumption that underlies our sampling approach is that $P(\mathcal{P} \mid \rho) \approx P(\mathcal{P})$, so that $P(\mathcal{P}, \rho) \approx P(\mathcal{P})\, P(\rho)$. Our approximation is thus that the density field does not inform where, given a set of locations *with the same density*, high (or, equivalently, low) shock emission is more likely to occur.

The real world will violate this assumption to some extent. For example, a point in the outskirts of a galaxy cluster could be as dense as a point along the central axis of a prominent filament; however, the cluster point still likely has a different MASSCW MEC probability distribution than the filament point. One reason could be the presence of passing merger shocks in cluster outskirts; another could be the higher typical speed by which accretion shocks crash onto clusters, compared to the typical speed by which they hit filaments.

3.5.7 Spectral indices in 3D and 2D

In Sect. 3.2.5, we calculate I_ν by generating j_ν at different emission frequencies with the same percentile random field. This approach implicitly assumes that if shocks were to be ordered by their MEC, the ordering remains invariant over emission frequency range $[\nu_{\mathrm{obs}}, \nu_{\mathrm{obs}}\,(1 + z_{\mathrm{max}})]$. Does this assumption correspond to a plausible physical scenario? Let the MECs of two shocks at some reference emission frequency $\nu = \nu_{\mathrm{ref}}$ be j_1 and j_2, and let $j_1 > j_2$ without loss of generality. Assume the existence of some function $\alpha = \alpha\,(j_\nu)$, that assigns integrated spectral indices to shocks based on their MEC at $\nu = \nu_{\mathrm{ref}}$. At emission frequency ν, the MEC ranks of the shocks are the same as at ν_{ref} if and only if

$$j_1 \left(\frac{\nu}{\nu_{\mathrm{ref}}}\right)^{\alpha(j_1)} > j_2 \left(\frac{\nu}{\nu_{\mathrm{ref}}}\right)^{\alpha(j_2)} \quad \text{for all}\, j_1, j_2 \in \mathbb{R}_{\geq 0},\, j_1 > j_2. \tag{3.30}$$

The rewritten inequality

$$\frac{j_1}{j_2} > \left(\frac{\nu}{\nu_{\mathrm{ref}}}\right)^{\alpha(j_2) - \alpha(j_1)} \quad \text{for all}\, j_1, j_2 \in \mathbb{R}_{\geq 0},\, j_1 > j_2 \tag{3.31}$$

suggests $\alpha\,(j_2) - \alpha\,(j_1) \leq 0$, or $\alpha\,(j_1) \geq \alpha\,(j_2)$ for $\nu > \nu_{\mathrm{ref}}$: $\alpha\,(j_\nu)$ must be a monotonically increasing function. Analogously, for $\nu < \nu_{\mathrm{ref}}$, we find that $\alpha\,(j_\nu)$ must be a monotonically decreasing function. These scenarios are visualised in Fig. 3.17. Us-

ing the Enzo simulation data used in this work, we explore the spectral index - MEC relation at $\nu_{\rm ref} = 180$ MHz, and find a general downward trend for spectral index as a function of MEC at this frequency. However, the relation is scattery and therefore not fully described by a monotonically decreasing function $\alpha\left(j_\nu\right)$; this implies that shock MEC ranks do change when varying the emission frequency. Although our approach to determining spectral indices remains approximate, compared to other uncertainties in our methodology, the error thus introduced is likely of minor importance.[24]

Our formalism allows for the generation of the function $I_\nu\left(\hat{r}, \nu_{\rm obs}\right)$ at two (or more) different observing frequencies. In turn, this enables the calculation of spectral indices for specific intensity rather than MEC (i.e. 'in 2D' instead of 'in 3D'), emulating the type of spectral analysis routinely performed by observational astronomers. We caution that the procedure for MEC spectral index assignment used in this work does not respect the full diversity of spectral behaviour present in the Enzo simulation, and instead forces MEC spectral indices to approach the MEC-weighted mean. In turn, this also causes specific intensity spectral index variations to be biased low. Future work should adapt this procedure so that a plausible specific intensity spectral index prior can be added to the potent suite of predictions that follow simultaneously from our methodology.

3.6 Conclusions

In this work, we describe and implement the first methodology to produce a (prior) probability distribution over specific intensity functions representing the synchrotron cosmic web (SCW) of the Local Universe. We assume merger and accretion shocks to be the main generators of the SCW, and assume a primordial magnetogenesis scenario for the evolution of magnetic fields in the IGM. However, the methodology is general enough to explore alternative physical hypotheses in the future. Our prior can be used to guide and verify observational attempts to detect the SCW with low-frequency radio telescopes such as the LOFAR and the SKA.

1. Using BORG SDSS total matter density reconstructions and Enzo cosmological simulations, we have built a prior distribution that is informative over half of the Northern Sky, and that has a $\sim 0.6^\circ$ resolution for LSS at $z = 0.1$. Although not a fundamental limitation of the methodology, the current reso-

[24] A conceptually correct way to address this problem would be to realise both a spectral index and a MEC at some fixed emission frequency ($\nu = \nu_{\rm obs}$, say), using $P\left(\alpha, j_\nu \,|\, \rho\right) = P\left(\alpha \,|\, j_\nu, \rho\right) P\left(j_\nu \,|\, \rho\right)$. The conditional probability distribution $P\left(\alpha \,|\, j_\nu, \rho\right)$ could be learnt from the Enzo simulation data used in this work, too.

lution is not high enough to resolve individual merger and accretion shocks. Typically, filaments near massive structures give $I_\nu\ \xi_e^{-1} \sim 10^{-1}$ Jy deg^{-2} at $\nu_{\rm obs} = 150$ MHz; ξ_e is the highly uncertain electron acceleration efficiency. Even at the $\sim$3 Mpc h^{-1} reconstruction resolution, our MASSCW prior indicates that the specific intensity for a given direction is highly uncertain (with a typical standard deviation being $\sim$100% of the mean) due to uncertainty regarding the presence and highly variable nature of shock emission along the line-of-sight. We present (marginal median) MASSCW specific intensity predictions for three deep LOFAR HBA fields: the Lockman Hole, Abell 2255, and the Ursa Major Supercluster.

2. With a simple geometric model of cubic unit cells, we have calculated both the distribution of the specific intensity–weighted mean redshift RV $\bar{Z}$, as well as the flux density–weighted mean redshift $\bar{\bar{z}}$ of the MASSCW signal for LSS reconstructions up to $z_{\rm max} = 0.2$. We obtain results that closely resemble those found numerically from our data-driven MASSCW prior, whose construction is much more involved. Encouraged by this, we present filament-only geometric model predictions for $\bar{\bar{z}}$ that include all LSS (i.e. $z_{\rm max} = \infty$). These predictions are highly insensitive to plausible variations in model parameters l and d_1, demonstrating that $\bar{\bar{z}}$ is effectively determined by a single parameter: the sum of the typical MEC-weighted filament spectral index $\alpha_{\rm f}$ and the MEC - cosmological redshift power law exponent $\beta_{\rm f}$. Its future determination will characterise the completeness of the MASSCW predictions put forth in this work.

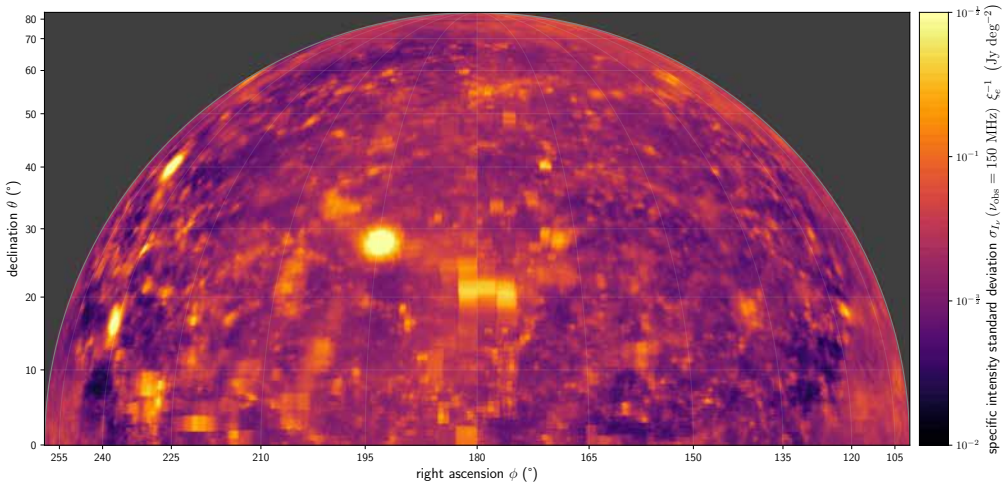
In an optimistic case, our prior already reveals a great deal about filamentary baryons — and where to find them.

M.S.S.L. Oei and R.J. van Weeren acknowledge support from the VIDI research programme with project number 639.042.729, which is financed by the Dutch Research Council (NWO). F. Vazza acknowledges support from the ERC STG MAGCOW (714196) from the H2020. F. Leclercq acknowledges funding from the Imperial College London Research Fellowship Scheme. The cosmological simulations used in this work were produced with the Enzo code (enzo-project.org) and run on the Piz-Daint supercluster at LSCS (Lugano) under project 's701' with F. Vazza as P.I..

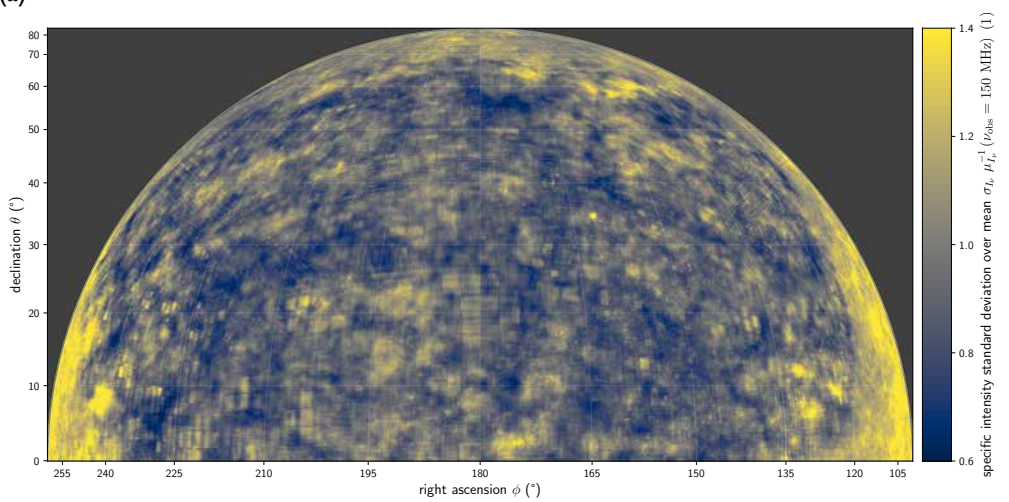
M.S.S.L. Oei warmly thanks Andrea Botteon, Matthias Hoeft, Vincent Icke, Josh Albert, Lara Anisman, Jacob Bakermans and Jesse van Oostrum for helpful discussions. *In dearest memory of Grandma Maria, who brightened my universe with so much love and laughter.*

3.A1 Additional figures

In this appendix, we list four figures referenced in the main text, but relegated for structural clarity.



(a)



(b)

Figure 3.14: MASSCW priors allow for a quantification of prediction uncertainty. Here we show both an absolute and a relative measure of spread for the single-direction specific intensity distributions (i.e. marginals) at $\nu_{\rm obs} = 150$ MHz. For each direction, we discard data outside the $1 - 99\%$ percentile range. *Top:* marginal standard deviation (absolute uncertainty), which closely resembles the marginal mean of Fig. 3.9a. *Bottom:* marginal standard deviation over mean (relative uncertainty), which reveals an inverted trend.

3.A2 Single-shock synchrotron MEC–total matter density scaling relation

In this appendix, we derive a synchrotron power density–total matter density scaling relation for individual shocks with high upstream Mach numbers in cluster outskirts

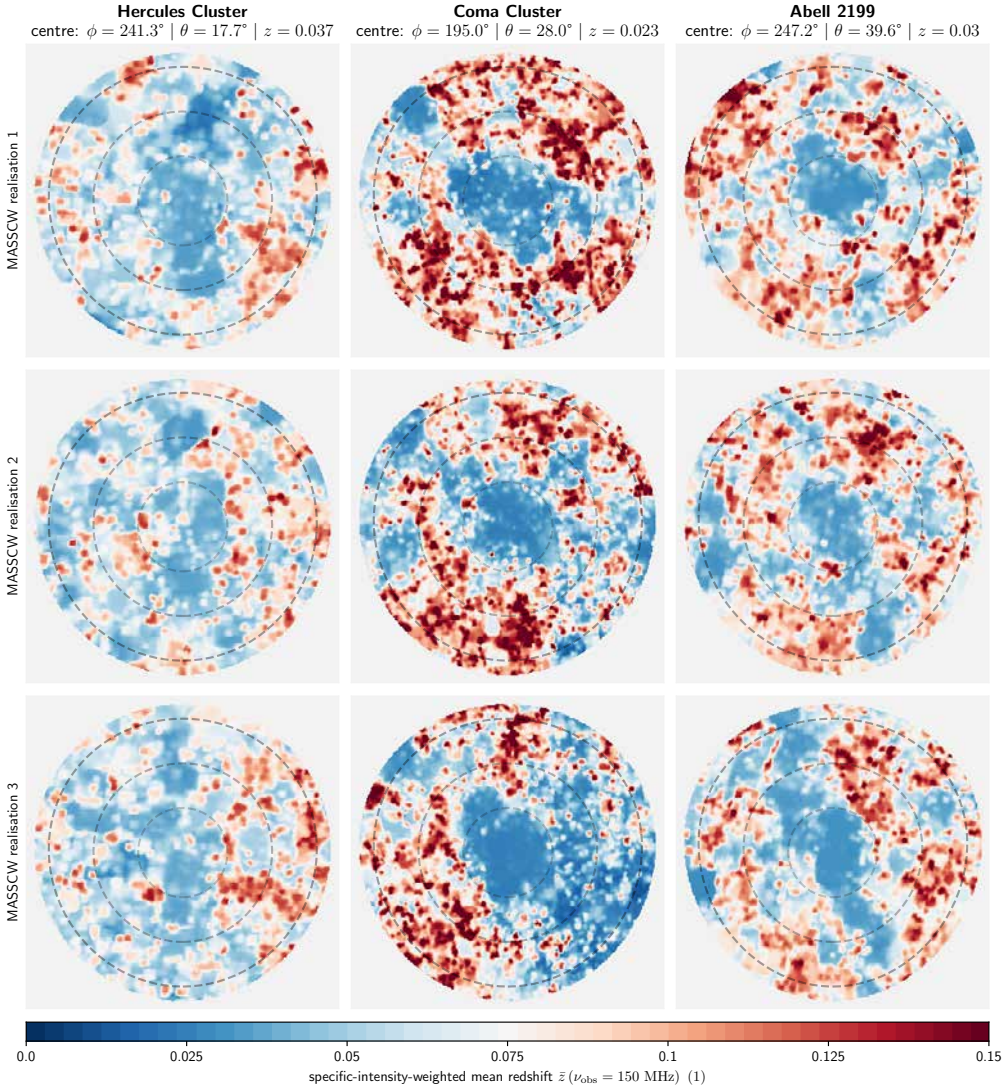


Figure 3.15: Same as Fig. 3.10, but now for specific intensity–weighted mean redshift instead of specific intensity.

and filaments, assuming the Hoeft & Brüggen (2007) model and $\gamma = 5/3$. Because the power density and MEC of a single shock are proportional, this immediately also yields the desired single-shock synchrotron MEC–total matter density scaling relation.

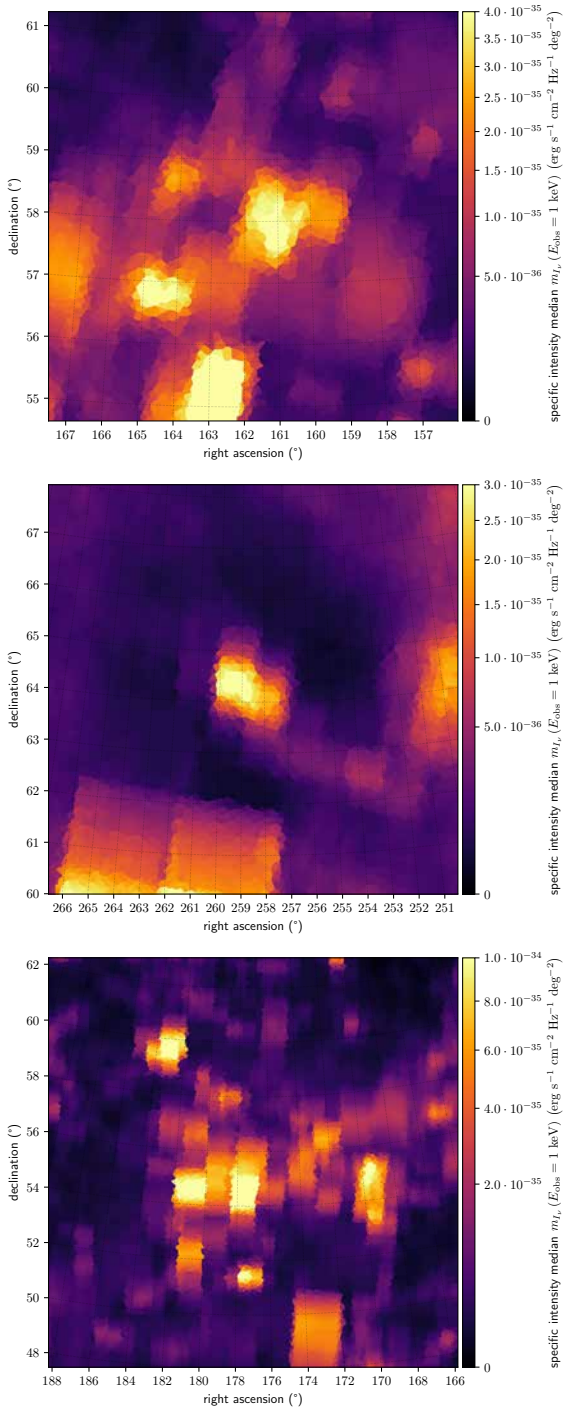


Figure 3.16: Thermal bremsstrahlung specific intensity prior marginal medians at $E_{\rm obs} = 1$ keV for three deep LOFAR HBA fields. For merger and accretion shock synchrotron predictions of the same regions, see Fig. 3.11. *Top:* Lockman Hole. *Middle:* Abell 2255. *Bottom:* Ursa Major Supercluster.

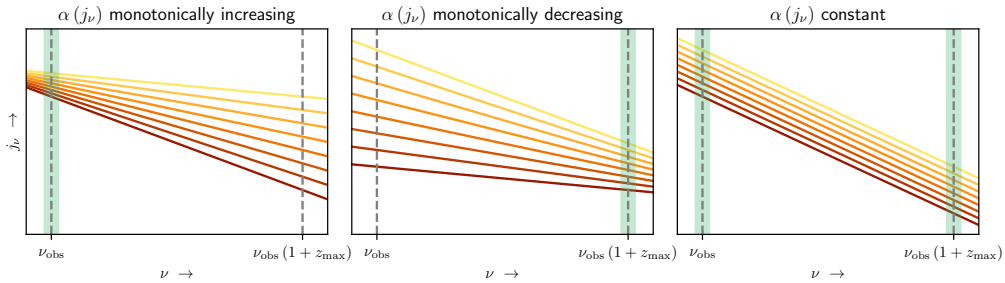


Figure 3.17: Three scenarios for simple power-law synchrotron spectra of shocks that are consistent with our methodology. Due to cosmological redshifting, the calculation of the specific intensity I_ν at observing frequency $\nu_{\rm obs}$ necessitates knowing the monochromatic emission coefficient (MEC) j_ν at a range of emission frequencies ν (see Eq. 3.8). In this work, we generate j_ν at emission frequencies $\nu = \nu_{\rm obs}$ and $\nu = \nu_{\rm obs}\,(1 + z_{\rm max})$, where $z_{\rm max}$ is the maximum cosmological redshift of the large-scale structure reconstructions. We do so using the same Gaussian random field realisation $\mathcal{Z}$ (and thus percentile random field realisation $\mathcal{P}$), implicitly assuming that shocks retain their MEC percentile rank over the frequency range $[\nu_{\rm obs}, \nu_{\rm obs}\,(1 + z_{\rm max})]$. This assumption holds in the three scenarios sketched above. Each solid line represents a shock with constant percentile rank (over the range shown, at least); the graphs are drawn with logarithmic scaling. Regarding the sign of the spectral index, we use the convention $j_\nu \propto \nu^\alpha$. *Left:* the integrated spectral index α is a monotonically increasing function of j_ν at $\nu = \nu_{\rm obs}$ ('brighter shocks have flatter spectra'). *Middle:* the integrated spectral index α is a monotonically decreasing function of j_ν at $\nu = \nu_{\rm obs}\,(1 + z_{\rm max})$ ('brighter shocks have steeper spectra'). *Right:* the integrated spectral index α is a j_ν-independent constant (at all frequencies); this is a limiting case of the previous two.

3.A2.1 Temperature and the speed of sound

Note that $\mathcal{M}_{\rm u} := \frac{v_{\rm u}}{c_{\rm s,u}}$, where $v_{\rm u}$ is the shock velocity relative to the upstream plasma, and $c_{\rm s,u}$ is the speed of sound in the upstream plasma. The Newton–Laplace equation for an ideal gas predicts $c_{\rm s,u} \propto \sqrt{T_{\rm u}}$, where $T_{\rm u}$ is the upstream plasma temperature.[25] The upstream Mach number of a shock incident on the WHIM would therefore be higher than that of a shock incident on the ICM if the shocks arrive at the same velocity $v_{\rm u}$ relative to these media.

3.A2.2 The filament regime: low magnetic field strengths

One of the prime reasons for pursuing SCW detections is to gauge the unknown strength of the Universe's largest magnetic fields. Numerical simulations by Vazza et al. (2015, 2017) that reproduce the observed magnetic field strengths in galaxy clusters, predict magnetic field strengths in filaments that depend strongly on the magnetogenesis scenario considered, ranging between $10^{-1} - 10^{2}$ nG. For the purposes of

[25] For example, the speed of sound is 10 times higher in the 10^8 K ICM than in the 10^6 K WHIM, and 100 times higher in the 10^9 K ICM than in the 10^5 K WHIM.

finding a power density–matter density scaling relation, the relevant quantity to compare the filament IGM magnetic field strength B_{IGM} with at cosmological redshift z, is the CMB magnetic field strength $B_{\mathrm{CMB}}(z)$.

As the CMB is well modelled by a blackbody[26], the CMB magnetic field strength $B_{\mathrm{CMB}}(z)$ is derived by equating the electromagnetic energy density u_{EM} of a blackbody of temperature T to the electromagnetic energy density of a magnetic field of magnitude B:

$$\frac{4\sigma}{c}T^4 = u_{\mathrm{EM}} = \frac{1}{2\mu_0}B^2, \tag{3.32}$$

where σ is the Stefan–Boltzmann constant and μ_0 is the vacuum permeability. Upon rearranging, and for $T = T_{\mathrm{CMB}}$ and $B = B_{\mathrm{CMB}}$, we find

$$B_{\mathrm{CMB}}(z) = \sqrt{\frac{8\mu_0\sigma}{c}}T^2_{\mathrm{CMB}}(z)\,. \tag{3.33}$$

Let a be the scale factor and let a_0 be its present-day value. As $T^4_{\mathrm{CMB}} \propto u_{\mathrm{EM}} \propto a^{-4} = a_0^{-4}(1+z)^4$ due to the expansion of the Universe, it follows that $T_{\mathrm{CMB}} \propto 1+z$, and thus

$$\begin{aligned} B_{\mathrm{CMB}}(z) &= \sqrt{\frac{8\mu_0\sigma}{c}}T^2_{\mathrm{CMB}}(0)(1+z)^2 \\ &= B_{\mathrm{CMB}}(0)(1+z)^2\,. \end{aligned} \tag{3.34}$$

Using $T_{\mathrm{CMB}}(0) = 2.725$ K yields $B_{\mathrm{CMB}}(0) = 3.238$ μG.

Thus, under all plausible scenarios of magnetogenesis, $B_{\mathrm{IGM}}(z) \ll B_{\mathrm{CMB}}(0) \leq B_{\mathrm{CMB}}(z)$.

3.A2.3 Magnetic field strength and baryon density

A scaling relation between B and ρ_{BM} follows from considering the conservation of magnetic flux as the Universe expands. The magnetic flux through a surface is the product of the surface area and the magnetic field strength (and the cosine of the angle between the surface normal and the magnetic field). Over time, the surface area increases $\propto a^2$, so that the magnetic field strength must follow $\propto a^{-2}$ if conservation of magnetic flux is to hold. Finally, as $\rho_{\mathrm{BM}} \propto a^{-3}$, one obtains $B \propto \rho_{\mathrm{BM}}^{2/3}$. Fig. 4

[26] In fact, it is the most accurate blackbody ever observed!

of Vazza et al. (2017) compares this scaling relation with simulated magnetic field strengths and baryon densities under various scenarios of magnetogenesis.

3.A2.4 The power density expression simplifies

The Hoeft & Brüggen (2007) power density folded into Eq. 3.2 simplifies appreciably if high Mach numbers and low magnetic field strengths are assumed. In such a regime, $\alpha \approx -1$ and $\Upsilon \approx 1$, while $B_{\mathrm{d}}^2 \ll B_{\mathrm{CMB}}^2$. Thus,

$$\begin{aligned} P_\nu &\propto \rho_{\mathrm{BM,d}} \cdot T_{\mathrm{d}}^{\frac{3}{2}} \cdot B_{\mathrm{d}}^2 \\ &= \frac{\rho_{\mathrm{BM,d}}}{\rho_{\mathrm{BM,u}}} \rho_{\mathrm{BM,u}} \cdot \left(\frac{T_{\mathrm{d}}}{T_{\mathrm{u}}} T_{\mathrm{u}} \right)^{\frac{3}{2}} \cdot \left(\frac{B_{\mathrm{d}}}{B_{\mathrm{u}}} B_{\mathrm{u}} \right)^2 . \end{aligned} \tag{3.35}$$

From the Rankine–Hugoniot jump conditions, one can derive that the compression factor (Fig. 3.1, central panel), as a function of $\mathcal{M}_{\mathrm{u}}$ and γ, is

$$\frac{\rho_{\mathrm{BM,d}}}{\rho_{\mathrm{BM,u}}} (\mathcal{M}_{\mathrm{u}}, \gamma) = \frac{(\gamma+1)\, \mathcal{M}_{\mathrm{u}}^2}{(\gamma-1)\, \mathcal{M}_{\mathrm{u}}^2 + 2} = \left(\frac{B_{\mathrm{d}}}{B_{\mathrm{u}}} (\mathcal{M}_{\mathrm{u}}, \gamma) \right)^{\frac{3}{2}} . \tag{3.36}$$

The same equations dictate that the temperature increase (Fig. 3.1, bottom panel), as a function of $\mathcal{M}_{\mathrm{u}}$ and γ, is

$$\frac{T_{\mathrm{d}}}{T_{\mathrm{u}}} (\mathcal{M}_{\mathrm{u}}, \gamma) = \frac{2\gamma (\gamma-1)\, \mathcal{M}_{\mathrm{u}}^2 + 4\gamma - (\gamma-1)^2 - 2(\gamma-1)\, \mathcal{M}_{\mathrm{u}}^{-2}}{(\gamma+1)^2} . \tag{3.37}$$

Thus, for $\mathcal{M}_{\mathrm{u}} \gg 1$, $\frac{\rho_{\mathrm{BM,d}}}{\rho_{\mathrm{BM,u}}} \propto 1$ and $\frac{T_{\mathrm{d}}}{T_{\mathrm{u}}} \propto \mathcal{M}_{\mathrm{u}}^2$. Returning to the power density scaling relation, we find

$$\begin{aligned} P_\nu &\propto \rho_{\mathrm{BM,u}} \cdot \left(\mathcal{M}_{\mathrm{u}}^2\, T_{\mathrm{u}} \right)^{\frac{3}{2}} \cdot B_{\mathrm{u}}^2 \\ &= \rho_{\mathrm{BM,u}} \cdot v_{\mathrm{u}}^3 \cdot B_{\mathrm{u}}^2 . \end{aligned} \tag{3.38}$$

To arrive at the second line, we use the definition of the upstream Mach number and the upstream sound speed–upstream plasma temperature scaling relation.

3.A2.5 Upstream velocity and total density

We investigate the upstream velocity–total (i.e. dark and baryonic matter) density relation for three simple geometries. We invoke Gauss' law for gravity to find ex-

pressions for the gravitational field, derive the gravitational potential using the fact that the gravitational force is conservative, and equate, for a test particle, the loss in gravitational potential energy to the gain in kinetic energy. We assume that the structures considered have hard edges and are equidense (with total matter density $\rho := \rho_{\mathrm{BM,u}} + \rho_{\mathrm{DM}}$) within.

Upon impact, the velocity of a test particle starting from rest at a distance d, and falling onto an isolated spherical galaxy cluster with radius R, is

$$v_{\mathrm{u}} = \sqrt{\frac{8\pi G}{3}\rho\left(1 - \frac{R}{d}\right)}\, R. \tag{3.39}$$

Upon impact, the velocity of a test particle starting from rest at a distance d, and falling onto an isolated cylindrical filament with radius R, is

$$v_{\mathrm{u}} = \sqrt{4\pi G\rho \ln \frac{d}{R}}\, R. \tag{3.40}$$

Upon impact, the velocity of a test particle starting from rest at a distance d, and falling onto an isolated, thick planar sheet of half-thickness R, is

$$v_{\mathrm{u}} = \sqrt{8\pi G\rho R\,(d - R)}. \tag{3.41}$$

In all three cases, R determines the size of the structure types. Assuming no relation between total density ρ and R, we find $v_{\mathrm{u}} \propto \sqrt{\rho}$, irrespective of the geometry.

3.A2.6 Baryon density and dark matter density

Structure formation theory predicts that after decoupling, and if gas pressure is ignored, $\rho_{\mathrm{BM,u}} \propto \rho_{\mathrm{DM}}$. Including gas pressure, this proportionality is expected to remain valid on large scales only. For filaments in particular, Fig. 6 of Gheller et al. (2016) shows that the baryon fraction in filaments remains close to $f_{\mathrm{cosmic}} := \Omega_{\mathrm{BM},0}\left(\Omega_{\mathrm{BM},0} + \Omega_{\mathrm{DM},0}\right)^{-1}$ $(= 0.167)$ over four orders of magnitude of total baryonic mass, and for redshifts from 1 to 0; this is consistent with $\rho_{\mathrm{BM,u}} \propto \rho_{\mathrm{DM}}$.

Using $B_{\mathrm{u}} \propto \rho_{\mathrm{BM,u}}^{2/3}$, $v_{\mathrm{u}} \propto \sqrt{\rho}$ and $\rho_{\mathrm{BM,u}} \propto \rho_{\mathrm{DM}}$, the final scaling relation becomes

$$j_\nu \propto P_\nu \propto \rho^{\frac{23}{6}}. \tag{3.42}$$

Note that, to obtain a scaling relation for the *total* (rather than single-shock) MEC, one should also consider how the shock number density relates to the total matter density. The total MEC–matter density and the single-shock MEC–matter density scaling relation exponents are only the same when no such relationship exists.

3.A3 Ray tracing in the cosmological setting

Projecting a ray's 4D null geodesic in a pure Friedmann–Lemaître–Robertson–Walker (FLRW) metric onto 3D comoving space results in a straight line. This follows readily from the FLRW metric in hyperspherical coordinates

$$\mathrm{d}s^2 = -c^2\mathrm{d}t^2 + \left(\frac{a(t)}{a_0}\right)^2 \left(\mathrm{d}r^2 + S_k^2\left(r\right)\mathrm{d}\Omega^2\right), \tag{3.43}$$

where $\mathrm{d}s^2$ is the spacetime line element, c is the speed of light *in vacuo*, t is physical time since the Big Bang, and $a\left(t\right)$ is the scale factor (with $a_0 := a\left(t_0\right)$ being its present-day ($t = t_0$) value). Also, r is the radial comoving distance, k is the Universe's Gaussian curvature (with SI units m^{-2}), $S_k\left(r\right) := r\,\mathrm{sinc}(r\sqrt{k})$ is the transverse comoving distance and $\mathrm{d}\Omega^2 := \mathrm{d}\theta^2 + \cos^2\theta\mathrm{d}\varphi^2$. (The sinc function follows the mathematical (unnormalised) convention.) Finally, let the location of present-day Earth be the spatial origin. An initially radial ($\mathrm{d}\Omega^2 = 0$) null ($\mathrm{d}s^2 = 0$) geodesic thus satisfies

$$c\mathrm{d}t = \pm\frac{\mathrm{d}r}{1+z}, \tag{3.44}$$

regardless of k. As time progresses, such rays maintain their direction and only change in r; this justifies considering the path $\mathcal{L}$ of a light ray with direction $\hat{r}$ in comoving space as the set of points $\mathcal{L}\left(\hat{r}\right) := \{r\,\hat{r} \in \mathbb{R}^3 \mid r \in \mathbb{R}_{\geq 0}\}$.

3.A4 Observer's specific intensity

Our aim is to derive an expression for the specific intensity in direction $\hat{r}$ at observing frequency ν_{obs}. We follow Chapter 12 of Peacock (1999), but generalise to arbitrary ΛCDM models (by allowing $\Lambda \neq 0$), and recast the results in terms of the MEC instead of the emissivity.

As in Appendix 3.A3, consider a FLRW metric with arbitrary Gaussian curvature k. A comoving volume element $\mathrm{d}V_{\mathrm{c}}$ and the corresponding proper volume element $\mathrm{d}V_{\mathrm{p}}$ at comoving radial distance r and cosmological redshift $z = z\left(r\right)$ that cover a

solid angle $\mathrm{d}\Omega$ on the sky are given by

$$\mathrm{d}V_{\mathrm{c}} = S_k(r)^2 \ \mathrm{d}\Omega \mathrm{d}r, \tag{3.45}$$

$$\mathrm{d}V_{\mathrm{p}} = S_k(r)^2 \ \mathrm{d}\Omega \mathrm{d}r\,(1+z)^{-3}. \tag{3.46}$$

Recall that we have defined j_ν as the *proper* MEC, and assume that the filament IGM radiates isotropically. The luminosity density $\mathrm{d}L_\nu$ of the volume, seen in direction $\hat{r}$ and at emission frequency $\nu = (1+z)\,\nu_{\mathrm{obs}}$, then equals

$$\mathrm{d}L_\nu(\hat{r},\nu) = 4\pi j_\nu(r\,\hat{r}, z, \nu)\,\mathrm{d}V_{\mathrm{p}}. \tag{3.47}$$

The corresponding observer's flux density $\mathrm{d}F_\nu$ of the volume in direction $\hat{r}$ at observing frequency ν_{obs} is

$$\mathrm{d}F_\nu(\hat{r},\nu_{\mathrm{obs}}) = \frac{\mathrm{d}L_\nu(\hat{r},\nu)}{4\pi S_k(r)^2\,(1+z)}, \tag{3.48}$$

and so the observer's specific intensity $\mathrm{d}I_\nu$ in direction $\hat{r}$ at observing frequency ν_{obs} is

$$\mathrm{d}I_\nu(\hat{r},\nu_{\mathrm{obs}}) := \frac{\mathrm{d}F_\nu(\hat{r},\nu_{\mathrm{obs}})}{\mathrm{d}\Omega} \tag{3.49}$$

$$= \frac{4\pi j_\nu(r\,\hat{r}, z, \nu)\,S_k(r)^2\,\mathrm{d}\Omega\mathrm{d}r}{4\pi S_k(r)^2\,(1+z)\,\mathrm{d}\Omega\,(1+z)^3} = \frac{j_\nu(r\,\hat{r}, z, \nu)\ \mathrm{d}r}{(1+z)^4}. \tag{3.50}$$

We neglect absorption so that the specific intensity of the ray only accumulates as the ray travels through LSS to the observer: the Universe is mostly optically thin for $\nu < 1$ GHz; we assume this holds perfectly. So, by collecting all contributions along the ray's path, we obtain the equivalent of Eq. 12.12 of Peacock (1999):

$$I_\nu(\hat{r},\nu_{\mathrm{obs}}) = \int_0^\infty \frac{j_\nu(r\,\hat{r}, z(r), \nu_{\mathrm{obs}}(1+z(r)))}{(1+z(r))^4}\mathrm{d}r. \tag{3.51}$$

Alternatively, one can express $I_\nu(\hat{r},\nu_{\mathrm{obs}})$ as an integral over z. Because

$$r(z) = \frac{c}{H_0}\int_0^z \frac{\mathrm{d}z'}{E(z')}, \tag{3.52}$$

with the dimensionless Hubble parameter $E(z) := \frac{H(z)}{H_0}$ being $E(z) = \sqrt{\Omega_{\mathrm{R},0}(1+z)^4 + \Omega_{\mathrm{M},0}(1+z)^3 + \Omega_{\mathrm{K},0}(1+z)^2 + \Omega_{\Lambda,0}}$ and today's curvature den-

sity parameter being $\Omega_{\mathrm{K},0} = 1 - \Omega_{\mathrm{R},0} - \Omega_{\mathrm{M},0} - \Omega_{\Lambda,0}$, we have

$$\mathrm{d}r = \frac{c}{H_0} \frac{\mathrm{d}z}{E(z)}. \tag{3.53}$$

Combining Eq. 3.51, 3.52 and 3.53,

$$I_\nu(\hat{r}, \nu_{\mathrm{obs}}) = \frac{c}{H_0} \int_0^\infty \frac{j_\nu(r(z)\,\hat{r}, z, \nu_{\mathrm{obs}}(1+z))}{(1+z)^4 E(z)} \mathrm{d}z. \tag{3.54}$$

(Barring notational differences, the $\Omega_{\Lambda,0} = 0$ (and $\Omega_{\mathrm{R},0} = 0$) limit of this formula is Eq. 12.10 of Peacock (1999).) Finally, we can read off that

$$\frac{\mathrm{d}I_\nu}{\mathrm{d}z}(\hat{r}, \nu_{\mathrm{obs}}, z) = \frac{c}{H_0} \frac{j_\nu(r(z)\,\hat{r}, z, \nu_{\mathrm{obs}}(1+z))}{(1+z)^4 E(z)}. \tag{3.55}$$

3.A5 Volume-filling fractions

cosmic web structure type	cubic unit cells $\frac{w_{\mathrm{f}}}{l} = 10^{-1}$	Forero-Romero et al. (2009) $\lambda_{\mathrm{th}} = 1, R_{\mathrm{eff}} = 0.88\,h^{-1}$ Mpc	VFF ratio	Forero-Romero et al. (2009) $\lambda_{\mathrm{th}} = 1, R_{\mathrm{eff}} = 2.05\,h^{-1}$ Mpc	VFF ratio
voids	72.9%	76%	1.04	82%	1.12
sheets	24.3%	18%	0.74	14%	0.58
filaments	2.7%	5%	1.85	4%	1.48
clusters	0.1%	0.5%	5.00	0.28%	2.80

Table 3.1: Comparison between cosmic web structure type volume-filling fractions (VFFs) predicted by the cubic unit cell model and those obtained by Forero-Romero et al. (2009) from cosmological simulations for eigenvalue threshold $\lambda_{\mathrm{th}} = 1$ and two effective smoothing scales R_{eff}. The VFF ratio columns give the simulation VFFs of the preceding column divided by the cubic unit cell VFFs of the first column.

In this appendix, we compute the volume-filling fractions (VFFs) of the four canonical structure types (clusters, filaments, sheets, and voids) as predicted by the cubic unit cell geometric model developed in Sect. 3.3. A single cubic unit cell features two typical lengthscales: a large scale (l_{f}), and a small scale (w_{f}), which can be interpreted as the typical filament length and width, respectively. The cube obtained by raising the large scale to the third power represents a void. Similarly, the three rectangular cuboids obtained by taking the product of the square of the large scale, and the small scale, resemble three sheets. The three rectangular cuboids obtained by taking the product of the large scale, and the square of the small scale, resemble three filaments. Finally, the cube obtained by raising the small scale to the third power resembles a

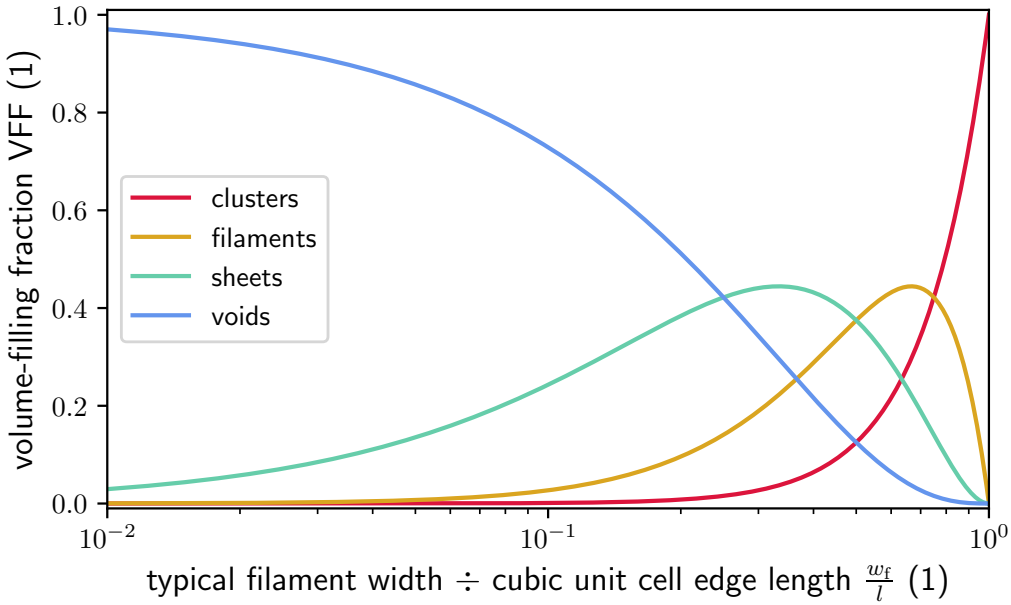


Figure 3.18: Cosmic web structure type volume-filling fractions (VFFs) as predicted by the cubic unit cell model of Sect. 3.3. These VFFs depend on just one parameter: the ratio between the (comoving) filament width and the (comoving) cubic unit cell edge length $\frac{w_f}{l}$.

cluster. The natural volume-filling fractions (VFFs) suggested by this geometry are thus

$$\mathrm{VFF_c}\left(\frac{w_f}{l}\right) = \left(\frac{w_f}{l}\right)^3, \tag{3.56}$$

$$\mathrm{VFF_f}\left(\frac{w_f}{l}\right) = 3\left(\frac{w_f}{l}\right)^2\left(1-\frac{w_f}{l}\right), \tag{3.57}$$

$$\mathrm{VFF_s}\left(\frac{w_f}{l}\right) = 3\frac{w_f}{l}\left(1-\frac{w_f}{l}\right)^2, \tag{3.58}$$

$$\mathrm{VFF_v}\left(\frac{w_f}{l}\right) = \left(1-\frac{w_f}{l}\right)^3. \tag{3.59}$$

See Fig. 3.18. To see that the VFFs sum to 1, we rewrite 1 using the binomial theorem:

$$1 = \left(\frac{w_f}{l}+1-\frac{w_f}{l}\right)^3 = \sum_{n=0}^{3}\binom{3}{n}\left(\frac{w_f}{l}\right)^n\left(1-\frac{w_f}{l}\right)^{3-n}, \tag{3.60}$$

and recognise the VFFs as the four terms in this expansion.

The VFFs obtained from this one-parameter model for $\frac{w_f}{l} = 10^{-1}$, are similar to those retrieved by Forero-Romero et al. (2009) from cosmological simulations for

eigenvalue threshold $\lambda_{\mathrm{th}} = 1$ and effective smoothing scales $R_{\mathrm{eff}} \sim 10^0$ Mpc (see Table 3.1).

3.A6 Notation

This paper adopts the SI system of units (for formulae concerning electromagnetism), and the following symbols. We list dimensionalities in SI base units with the radian 'rad' appended. Current-day quantities are subscripted with a zero: for example a_0 is today's scale factor, while a is the scale factor for arbitrary times. Upstream and downstream quantities are subscripted with a 'u' or 'd', respectively: for example T_{u} is the upstream plasma temperature, whilst T is the general plasma temperature.

symb.	SI base units	description
α	1	integrated synchrotron spectral index
α_{c}	1	typical cluster α
α_{f}	1	typical filament α
$\bar{\alpha}$	1	MEC-weighted mean α
β_{c}	1	typical cluster MEC $(1+z)$ power law exponent
β_{f}	1	typical filament MEC $(1+z)$ power law exponent
γ	1	adiabatic index
θ	rad	declination (J2000)
μ_0	$\mathrm{kg\,m\,s^{-2}\,A^{-2}}$	vacuum permeability
ν	$\mathrm{s^{-1}}$	emission frequency
ν_{obs}	$\mathrm{s^{-1}}$	observing frequency
ν_{ref}	$\mathrm{s^{-1}}$	reference frequency (see $\mathcal{C}$)
ξ_e	1	electron acceleration efficiency
Ξ_ρ	1	CDF of RV $j_\nu \mid \rho$
ρ	$\mathrm{kg\,m^{-3}}$	total matter density
ρ_{BM}	$\mathrm{kg\,m^{-3}}$	baryonic matter density
ρ_{c}	$\mathrm{kg\,m^{-3}}$	critical density
ρ_{DM}	$\mathrm{kg\,m^{-3}}$	dark matter density
σ	$\mathrm{kg\,s^{-3}\,K^{-4}}$	Stefan–Boltzmann constant
φ	rad	right ascension (J2000)
Φ	1	CDF of standard normal RV
Ψ	1	as in Hoeft & Brüggen (2007)

Ω_Λ	1	dark energy density parameter
Ω_{BM}	1	baryonic matter density parameter
Ω_{DM}	1	dark matter density parameter
Ω_{K}	1	curvature density parameter
Ω_{M}	1	matter density parameter
Ω_{R}	1	relativistic particle (i.e. photon and neutrino) density parameter
a	1	scale factor
A	m^2	shock surface area
B	$\mathrm{kg\,s^{-2}\,A^{-1}}$	proper magnetic field strength
B_{CMB}	$\mathrm{kg\,s^{-2}\,A^{-1}}$	CMB magnetic field strength
c	$\mathrm{m\,s^{-1}}$	speed of light *in vacuo*
c_{s}	$\mathrm{m\,s^{-1}}$	speed of sound
$\mathcal{C}$	1	typical cluster-to-filament synchrotron MEC ratio at $\nu = \nu_{\mathrm{ref}}$ and $z = 0$
d	m	initial distance between test particle and equidense cluster centre, filament axis or sheet midplane
d_n	m	comoving distance to the n-th unit cell boundary crossing
$\mathrm{d}\Omega^2$	rad^2	solid angle element
$\mathrm{d}s^2$	m^2	spacetime line element
E	1	dimensionless Hubble parameter
$F_{\bar{Z}}$	1	CDF of $\bar{Z}$
G	$\mathrm{kg^{-1}\,m^3\,s^{-2}}$	Newton's gravitational constant
h	1	Hubble constant divided by 100 $\mathrm{km\,s^{-1}\,Mpc^{-1}}$
H	$\mathrm{s^{-1}}$	Hubble parameter
I_ν	$\mathrm{kg\,s^{-2}\,rad^{-2}}$	specific intensity (observed)
j_ν	$\mathrm{kg\,m^{-1}\,s^{-2}\,rad^{-2}}$	proper (not comoving) MEC
$\mathbf{k}$	$\mathrm{m^{-1}}$	Fourier dual of $\mathbf{r}$
k	$\mathrm{m^{-2}}$	FLRW (Gaussian) curvature of the Universe
l	m	comoving cubic unit cell edge length
l_{f}	m	comoving filament length
l_{SE}	m	SE kernel lengthscale
$\mathcal{L}$	-	light ray path
m_r	1	Petrosian r-band apparent magnitude
M	1	number of sightlines
$\mathfrak{M}$	-	cosmological model parameter tuple

$\mathcal{M}_\mathrm{u}$	1	upstream shock Mach number
n_e	m^{-3}	electron number density
N	1	number of unit cell boundary crossings considered
$p_\mathrm{c-f}$	1	probability to pierce through a cluster, then a filament
p_f	1	probability to pierce through a filament only
$p_\mathrm{s-v}$	1	probability to pierce through a sheet, then a void
P_ν	$\mathrm{kg\,m^2\,s^{-2}}$	proper power density
$\mathcal{P}$	1	percentile random field
$\mathbf{r}$	m	comoving position vector
r	m	radial comoving distance
$\hat{r}$	1	sky direction unit vector
$\hat{r}_i$	1	sky direction unit vector of ray i
R	m	equidense cluster radius, filament radius or sheet half-width
$\mathcal{R}$	-	comoving reconstruction region; subset of $\mathbb{R}^3$
S_k	m	transverse comoving distance
t	s	physical time since the Big Bang
T	K	proper plasma temperature
T_CMB	K	CMB temperature
u_EM	$\mathrm{kg\,m^{-1}\,s^{-2}}$	electromagnetic energy density
v	$\mathrm{m\,s^{-1}}$	shock or test particle velocity
V	m^3	shock effective volume
w_f	m	typical comoving filament width
X_n	1	relative specific intensity contribution of the n-th newly-entered unit cell
X_{nm}	1	relative specific intensity contribution of the n-th newly-entered unit cell for the m-th ray
$\langle y \rangle$	m	shock effective width
z, $z_\mathfrak{M}$	1	cosmological redshift (under parameters $\mathfrak{M}$)
z_max	1	cosmological redshift up to which LSS is considered
z_n	1	cosmological redshift of n-th unit cell boundary crossing
$\bar{z}$	1	specific intensity–weighted mean redshift
$\bar{\bar{z}}$	1	flux density–weighted mean redshift
Z	1	cosmological redshift RV
$\bar{Z}$	1	specific intensity–weighted mean redshift RV
$\mathcal{Z}$	1	standard normal GRF

Table 3.2: Symbols, SI base units, and descriptions of all quantities used in this work. Quantities are ordered alphabetically by their symbols, with Greek symbols preceding Roman ones.

[O]nce powerful Telamon destroyed Troy and the Meropes and the great and terrible warrior Alcyoneus, but not before that giant had laid low, by hurling a rock, twelve chariots and twice twelve horse-taming heroes who were riding in them.

Pindar, Greek poet, *Nemean 4.25–29* (5th century B.C.)

4

The discovery of a radio galaxy of at least 5 Mpc

M. S. S. L. Oei, R. J. van Weeren, M. J. Hardcastle, A. Botteon, T. W. Shimwell, P. Dabhade, A. R. D. J. G. I. B. Gast, H. J. A. Röttgering, M. Brüggen, C. Tasse, W. L. Williams, A. Shulevski — *Astronomy & Astrophysics*, 660, 2, 2022

Abstract

Context Giant radio galaxies (GRGs, or colloquially 'giants') are the Universe's largest structures generated by individual galaxies. They comprise synchrotron-radiating active galactic nucleus ejecta and attain cosmological (megaparsec-scale) lengths. However, the main mechanisms that drive their exceptional growth remain poorly understood.

Aims To deduce the main mechanisms that drive a phenomenon, it is usually instructive to study extreme examples. If there exist host galaxy characteristics that are an important cause for GRG growth, then the hosts of the largest GRGs are likely to possess them. Similarly, if there exist particular large-scale environments that are highly conducive to GRG growth, then the largest GRGs are likely to reside in them. For these reasons, we aim to perform a case study of the largest GRG available.

Methods We reprocessed the LOFAR Two-Metre Sky Survey (LoTSS) DR2 by subtracting compact sources and performing multi-scale CLEAN de-convolutions

at 60″ and 90″ resolution. The resulting images constitute the most sensitive survey yet for radio galaxy lobes, whose diffuse nature and steep synchrotron spectra have allowed them to evade previous detection attempts at higher resolution and shorter wavelengths. We visually searched these images for GRGs.

Results We have discovered Alcyoneus, a low-excitation radio galaxy with a projected proper length $l_{\rm p} = 4.99 \pm 0.04$ Mpc. Both its jets and lobes are detected at very high significance, and the SDSS-based identification of the host, at spectroscopic redshift $z_{\rm spec} = 0.24674 \pm 6 \cdot 10^{-5}$, is unambiguous. The total luminosity density at $\nu = 144$ MHz is $L_\nu = 8 \pm 1 \cdot 10^{25}$ W Hz^{-1}, which is below average, though near median (percentile $45 \pm 3\%$) for GRGs. The host is an elliptical galaxy with a stellar mass $M_\star = 2.4 \pm 0.4 \cdot 10^{11}\, M_\odot$ and a super-massive black hole mass $M_\bullet = 4 \pm 2 \cdot 10^8\, M_\odot$, both of which tend towards the lower end of their respective GRG distributions (percentiles $25 \pm 9\%$ and $23 \pm 11\%$). The host resides in a filament of the Cosmic Web. Through a new Bayesian model for radio galaxy lobes in three dimensions, we estimate the pressures in the megaparsec-cubed-scale northern and southern lobes to be $P_{\rm min,1} = 4.8 \pm 0.3 \cdot 10^{-16}$ Pa and $P_{\rm min,2} = 4.9 \pm 0.6 \cdot 10^{-16}$ Pa, respectively. The corresponding magnetic field strengths are $B_{\rm min,1} = 46 \pm 1$ pT and $B_{\rm min,2} = 46 \pm 3$ pT.

Conclusions We have discovered what is in projection the largest known structure made by a single galaxy — a GRG with a projected proper length $l_{\rm p} = 4.99 \pm 0.04$ Mpc. The true proper length is at least $l_{\rm min} = 5.04 \pm 0.05$ Mpc. Beyond geometry, Alcyoneus and its host are suspiciously ordinary: the total low-frequency luminosity density, stellar mass, and super-massive black hole mass are all lower than, though similar to, those of the medial GRG. Thus, very massive galaxies or central black holes are not necessary to grow large giants, and, if the observed state is representative of the source over its lifetime, neither is high radio power. A low-density environment remains a possible explanation. The source resides in a filament of the Cosmic Web, with which it might have significant thermodynamic interaction. The pressures in the lobes are the lowest hitherto found, and Alcyoneus therefore represents the most promising radio galaxy yet to probe the warm–hot inter-galactic medium.

Key words: galaxies: active – galaxies: individual: Alcyoneus – galaxies: jets – intergalactic medium – radio continuum: galaxies

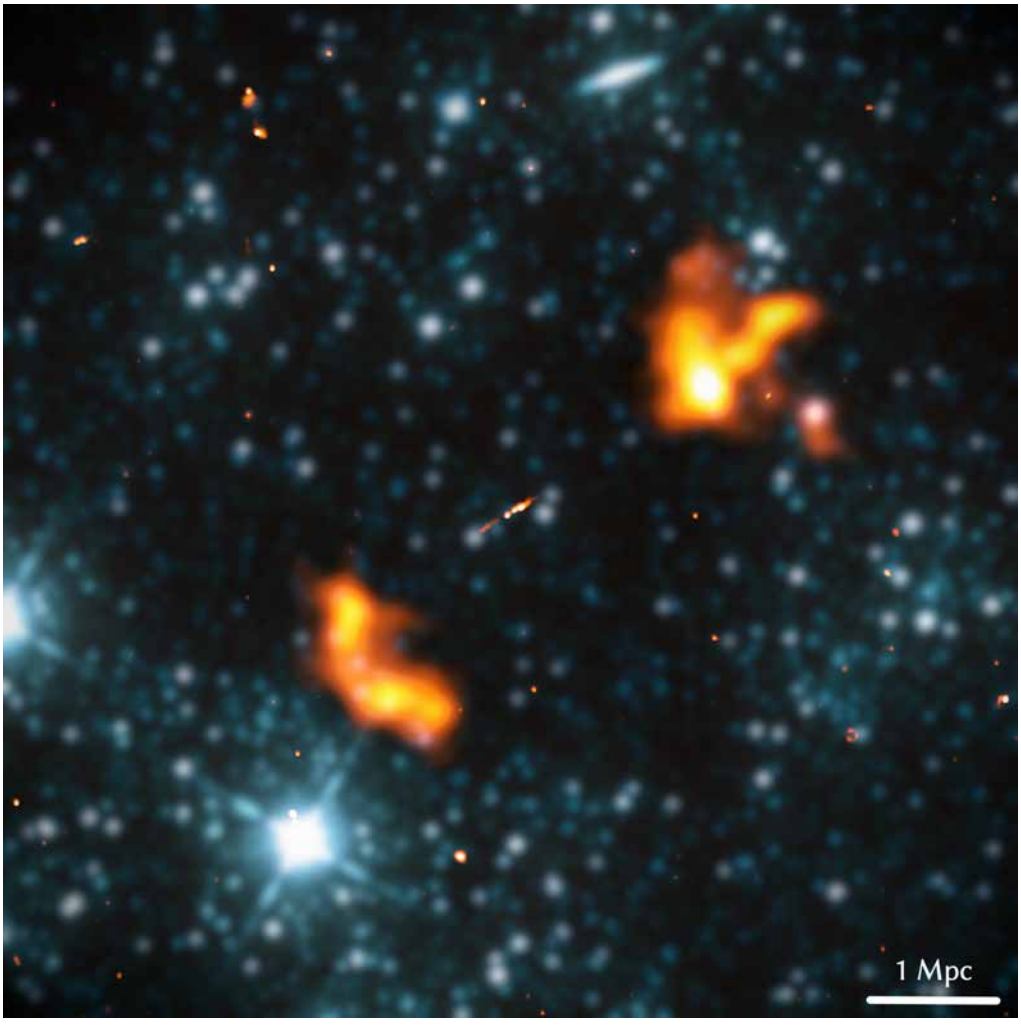


Figure 4.1: Joint radio-infrared view of Alcyoneus, a radio galaxy with a projected proper length of 5.0 Mpc. We show a $2048'' \times 2048''$ solid angle centred around right ascension 123.590372° and declination 52.402795°. We superimpose LoTSS DR2 images at 144 MHz of two different resolutions ($6''$ for the core and jets, and $60''$ for the lobes) (orange), with the WISE image at 3.4 μm (blue). To highlight the radio emission, the infrared emission has been blurred to $0.5'$ resolution.

4.1 Introduction

Most galactic bulges hold a super-massive ($M_\bullet > 10^6\ M_\odot$) Kerr black hole (e.g. Soltan, 1982) that grows by accreting gas, dust, and stars from its surroundings (Kormendy & Ho, 2013). The black hole ejects a fraction of its accretion disk plasma from the host galaxy along two collimated, magnetised jets that are aligned with its rotation axis (e.g. Blandford & Rees, 1974). The relativistic electrons contained herein experi-

ence Lorentz force and generate, through spiral motion, synchrotron radiation that is observed by radio telescopes. The two jets either fade gradually or end in hotspots at the end of diffuse lobes and ultimately enrich the inter-galactic medium with cosmic rays and magnetic fields. The full luminous structure is referred to as a 'radio galaxy' (RG). Members of a rare RG sub-population attain megaparsec-scale proper (and thus also co-moving) lengths (e.g. Willis et al., 1974; Andernach et al., 1992; Ishwara-Chandra & Saikia, 1999; Jamrozy et al., 2008; Machalski, 2011; Kuźmicz et al., 2018; Dabhade et al., 2020a). The definition of giant radio galaxies (GRGs, or colloquially 'giants') accommodates our limited ability to infer an RG's true proper length from observations: an RG is called a GRG if and only if its proper length projected onto the plane of the sky exceeds some threshold, $l_{\mathrm{p,GRG}}$, usually chosen to be 0.7 or 1 Mpc. Because the conversion between angular length and projected proper length depends on cosmological parameters, which remain uncertain, it is not always clear whether a given observed RG satisfies the GRG definition.

Currently, there are about a thousand GRGs known, the majority of which have been found in the northern sky. About one hundred exceed 2 Mpc, and ten exceed 3 Mpc; at 4.9 Mpc, the literature's projectively longest GRG is J1420-0545 (Machalski et al., 2008). As such, GRGs — and the rest of the megaparsec-scale RGs — are the largest single-galaxy–induced phenomena in the Universe. Which physical mechanisms lead some RGs to extend for $\sim 10^2$ times their host galaxy diameter is a key open question. To determine whether there exist particular host galaxy characteristics or large-scale environments that are essential for GRG growth, it is instructive to analyse the largest GRGs, since in these systems it is most likely that all major favourable growth factors are present. We thus aim to perform a case study of the largest GRG available.

As demonstrated by Dabhade et al. (2020b)'s record sample of 225 discoveries, the Low-Frequency Array (LOFAR; van Haarlem et al., 2013) is among the most attractive contemporary instruments for finding new GRGs. This pan-European radio interferometer features a unique combination of short baselines to provide sensitivity to large-scale emission and long baselines to mitigate source confusion.[1] These qualities are indispensable for observational studies of GRGs, which require both extended lobes and compact cores and jets to be identified. Additionally, the metre wavelengths at which the LOFAR operates allow it to detect steep-spectrum lobes far away from host galaxies. Such lobes reveal the full extent of GRGs but are missed by

[1] Source confusion is an instrumental limitation that arises when the resolution of an image is low compared to the sky density of statistically significant sources. It causes angularly adjacent but physically unrelated sources to blend together, making it hard or even impossible to distinguish them (e.g. Condon et al., 2012).

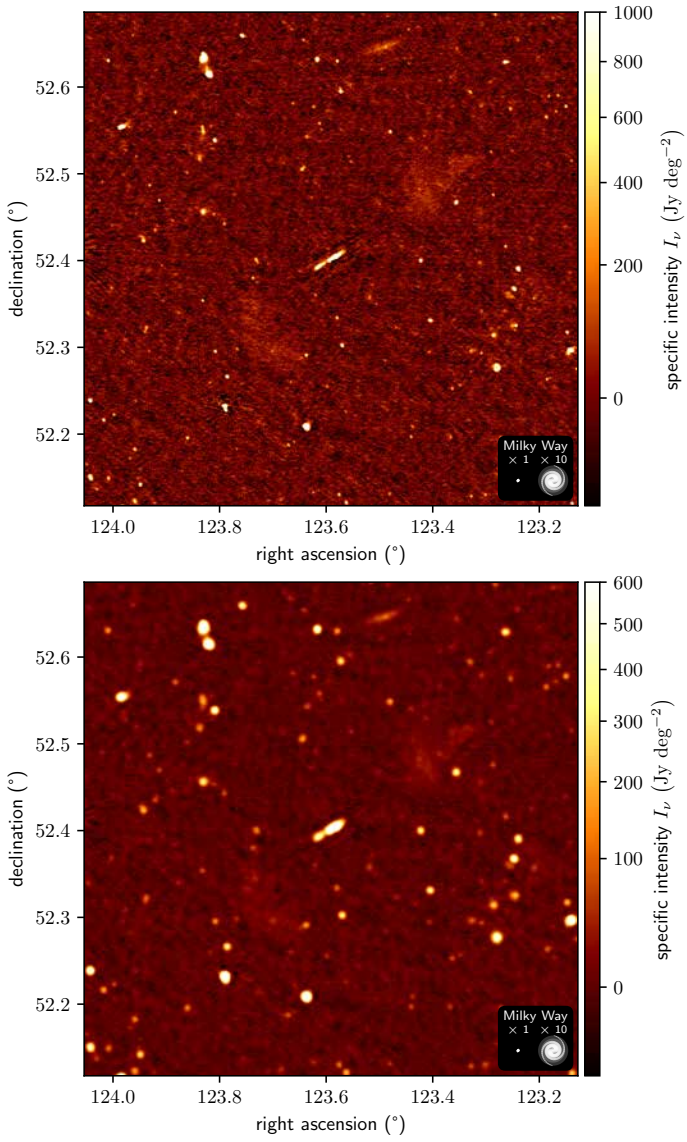


Figure 4.2: LoTSS DR2 images of Alcyoneus, centred around host galaxy J081421.68+522410.0, at central frequency $\nu_c = 144$ MHz and standard resolutions $\theta_{\mathrm{FWHM}} = 6''$ (top) and $\theta_{\mathrm{FWHM}} = 20''$ (bottom). At these resolutions, Alcyoneus's lobes are easily overlooked. For scale, we show the stellar Milky Way disk (diameter: 50 kpc) and a ten-times-inflated version; the spiral galaxy shape follows Ringermacher & Mead (2009).

decimetre observatories. Thus, in Sect. 4.2 we describe a reprocessing of the LOFAR Two-Metre Sky Survey (LoTSS) Data Release 2 (DR2) aimed at revealing hitherto unknown RG lobes — among other goals. An overview of the reprocessed images, which cover thousands of square degrees, and statistics of the lengths and environ-

ments of the GRGs they have revealed are subjects of future publications. In the present article we introduce Alcyoneus[2], a 5 Mpc GRG that these images allowed us to discover. We determine and discuss its properties in Sect. 4.3. Figure 4.1 provides a multi-wavelength, multi-resolution view of this giant. Section 4.4 contains our concluding remarks.

Throughout this work, we assume a concordance cosmological constant–cold dark matter cosmology with parameters $\mathfrak{M}$ from Planck Collaboration et al. (2020): $\mathfrak{M} = (h = 0.6766, \Omega_{\mathrm{BM},0} = 0.0490, \Omega_{\mathrm{M},0} = 0.3111, \Omega_{\Lambda,0} = 0.6889)$, where $H_0 := h \cdot 100\ \mathrm{km\ s^{-1}\ Mpc^{-1}}$. We define the spectral index α such that it relates to flux density F_ν at frequency ν as $F_\nu \propto \nu^\alpha$. Regarding terminology, we strictly distinguish between an RG (a radio-bright structure of relativistic particles and magnetic fields, consisting of a core, jets, hotspots, and lobes) and the host galaxy that generates it.

4.2 Data and methods

The LoTSS, conducted by the LOFAR High-Band Antennae (HBA), is a 120–168 MHz interferometric survey (Shimwell et al., 2017, 2019, 2022) with the ultimate aim to image the full northern sky at resolutions of 6″, 20″, 60″, and 90″. Its central frequency $\nu_c = 144$ MHz. The latest data release — the LoTSS DR2 (Shimwell et al., 2022) — covers 27% of the northern sky, split over two regions of 4178 deg^2 and 1457 deg^2; the largest of these contains the Sloan Digital Sky Survey (SDSS) DR7 (Abazajian et al., 2009) area. By default, the LoTSS DR2 provides imagery at the 6″ and 20″ resolutions. We show these standard products in Fig. 4.2 for the same sky region as in Fig. 4.1. In terms of total source counts, the LoTSS DR2 is the largest radio survey carried out thus far: its catalogue contains $4.4 \cdot 10^6$ sources, most of which are considered 'compact'.

By contrast, the 60″ and 90″ imagery, which we discuss in more detail in Oei et al. (prepa), is intended to reveal extended structures in the low-frequency radio sky, such as GRGs, supernova remnants in the Milky Way, radio halos and shocks in galaxy clusters, and — potentially — accretion shocks or volume-filling emission from filaments of the Cosmic Web. To avoid the source confusion limit at these resolutions, following van Weeren et al. (2021), we used DDFacet (Tasse et al., 2018) to predict visibilities corresponding to the 20″ LoTSS DR2 sky model and subtracted these from

[2] Alcyoneus was the son of Ouranos, the Greek primordial god of the sky. According to Ps.-Apollodorus, he was also one of the greatest of the *Gigantes* (Giants) and a challenger to Heracles during the Gigantomachy — the battle between the Giants and the Olympian gods for supremacy over the cosmos. The poet Pindar described him as 'huge as a mountain', fighting by hurling rocks at his foes.

the data, before imaging at 60″ and 90″ with WSClean IDG (Offringa et al., 2014; van der Tol et al., 2018). We used -0.5 Briggs weighting and multi-scale CLEAN (Offringa & Smirnov, 2017), with `-multiscale-scales 0,4,8,16,32,64`. Importantly, we did not impose an inner (u, v)-cut. We imaged each pointing separately, then combined the partially overlapping images into a mosaic by calculating, for each direction, a beam-weighted average.

Finally, we visually searched the LoTSS DR2 for GRGs, primarily at 6″ and 60″, using the Hierarchical Progressive Survey (HiPS) system in *Aladin Desktop 11.0* (Bonnarel et al., 2000).

4.3 Results and discussion

4.3.1 Radio morphology and interpretation

During our LoTSS DR2 search, we identified a three-component radio structure of total angular length $\varphi = 20.8'$, visible at all (6″, 20″, 60″ and 90″) resolutions. Figure 4.2 provides a sense of our data quality; it shows that the outer components are barely discernible in the LoTSS DR2 at its standard 6″ and 20″ resolutions. Meanwhile, Fig. 4.1 shows the outer components at 60″, and the top panel of Fig. 4.9 shows them at 90″; at these resolutions, they lie firmly above the noise. Compared with the outer structures, the central structure is bright and elongated, with a 155″ major axis and a 20″ minor axis. The outer structures lie along the major axis at similar distances from the central structure, are diffuse and amorphous, and feature specific intensity maxima along this axis.

In the arcminute-scale vicinity of the outer structures, the DESI Legacy Imaging Surveys (Dey et al., 2019) DR9 does not reveal galaxy overdensities or low-redshift spiral galaxies, the ROSAT All-Sky Survey (RASS) (Voges et al., 1999) does not show X-ray brightness above the noise, and there is no Planck Sunyaev–Zeldovich Catalogue 2 (PSZ2) (Planck Collaboration et al., 2016) source nearby. The outer structures therefore cannot be supernova remnants in low-redshift spiral galaxies or radio relics and radio halos in galaxy clusters. Instead, the outer structures presumably represent RG emission. The radio-optical overlays in the top and bottom panels of Fig. 4.3 show that it is improbable that each outer structure is an RG of its own, given the lack of significant 6″ radio emission (solid light green contours) around host galaxy candidates suggested by the morphology of the 60″ radio emission (translucent white contours). For these reasons, we interpret the central (jet-like) structure and the outer (lobe-like) structures as components of the same RG.

Subsequent analysis — presented below — demonstrates that this RG is the largest

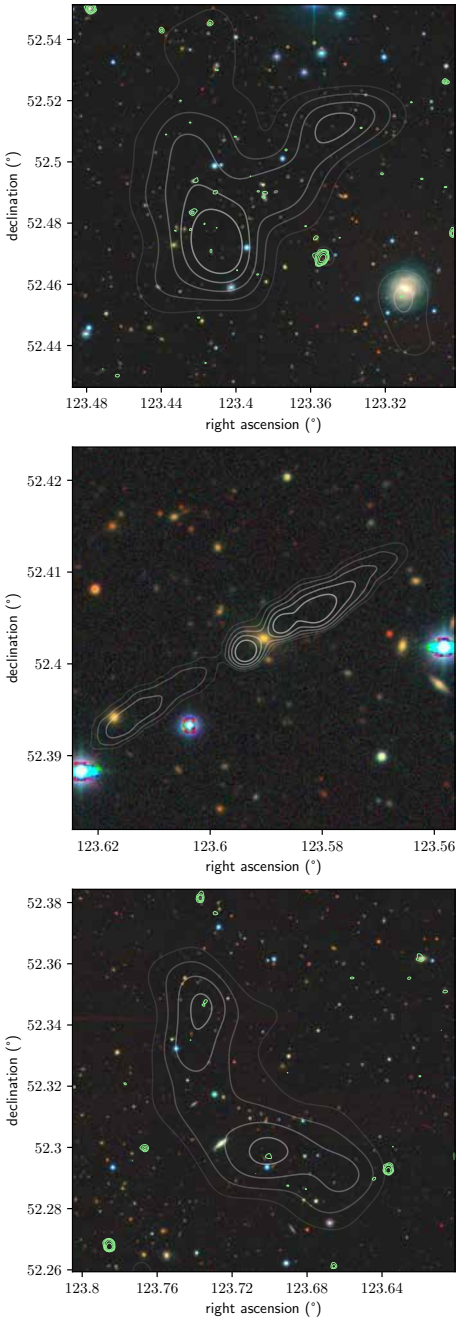


Figure 4.3: Joint radio-optical views showing that the outer structures shown in Fig. 4.1 are best interpreted as a pair of RG lobes fed by central jets. On top of DESI Legacy Imaging Surveys DR9 (g, r, z) imagery, we show the LoTSS DR2 at various resolutions through contours at multiples of σ, where σ is the image noise at the relevant resolution. The top and bottom panels show translucent white $60''$ contours at $3, 5, 7, 9, 11\sigma$ and solid light green $6''$ contours at $4, 7, 10, 20, 40\sigma$. The central panel shows translucent white $6''$ contours at $5, 10, 20, 40, 80\sigma$.

hitherto discovered, with a projected proper length of 5.0 Mpc. We dub this GRG 'Alcyoneus'.

4.3.2 Host galaxy identification

Based on the middle panel of Fig. 4.3 and an SDSS DR12 (Alam et al., 2015) spectrum, we identify a source at a J2000 right ascension of 123.590372°, a declination of 52.402795° and a spectroscopic redshift of $z_{\rm spec} = 0.24674 \pm 6 \cdot 10^{-5}$ as Alcyoneus's host. Like most GRG hosts, this source, with SDSS DR12 name J081421.68+522410.0, is an elliptical galaxy[3] without a quasar. From optical contours, we find that the galaxy's minor axis makes a ∼20° angle with Alcyoneus's jet axis.

In Fig. 4.4 we further explore the connection between J081421.68+522410.0 and Alcyoneus's radio core and jets. From top to bottom, we show the LoTSS DR2 at 6″, the Very Large Array Sky Survey (VLASS) (Lacy et al., 2020) at 2.2″, and the Panoramic Survey Telescope and Rapid Response System (Pan-STARRS) DR1 (Chambers et al., 2016) *i* band. Two facts confirm that the host identification is highly certain. First, for both the LoTSS DR2 at 6″ and the VLASS at 2.2″, the angular separation between J081421.68+522410.0 and the arc connecting Alcyoneus's two innermost jet features is at the sub-arcsecond scale. Moreover, the alleged host galaxy is the brightest Pan-STARRS DR1 *i*-band source within a radius of 45″ of the central VLASS image component.

4.3.3 Radiative- or jet-mode active galactic nucleus

Current understanding (e.g. Heckman & Best, 2014) suggests that the population of active galactic nuclei (AGN) exhibits a dichotomy: AGN seem to be either radiative-mode AGN, which generate high-excitation radio galaxies (HERGs), or jet-mode AGN, which generate low-excitation radio galaxies (LERGs). We wished to determine if Alcyoneus is a HERG or a LERG. The SDSS spectrum of the host features very weak emission lines; indeed, the star formation rate is just $1.6 \cdot 10^{-2}\ M_\odot\ \mathrm{yr}^{-1}$ (Chang et al., 2015). Following the classification rule of Best & Heckman (2012), Best et al. (2014), Pracy et al. (2016), and Williams et al. (2018) based on the strength and equivalent width of the OIII 5007 line, we determined that Alcyoneus is a LERG. Moreover, the Wide-Field Infrared Survey Explorer (WISE) photometry (Cutri & et al., 2012) at 11.6 μm and 22.1 μm is below the instrumental detection limit. Following the classification rule of Gürkan et al. (2014) based on the 22.1 μm luminosity

[3] Based on the SDSS morphology, Kuminski & Shamir (2016) calculate a probability of 89% that the galaxy is an elliptical.

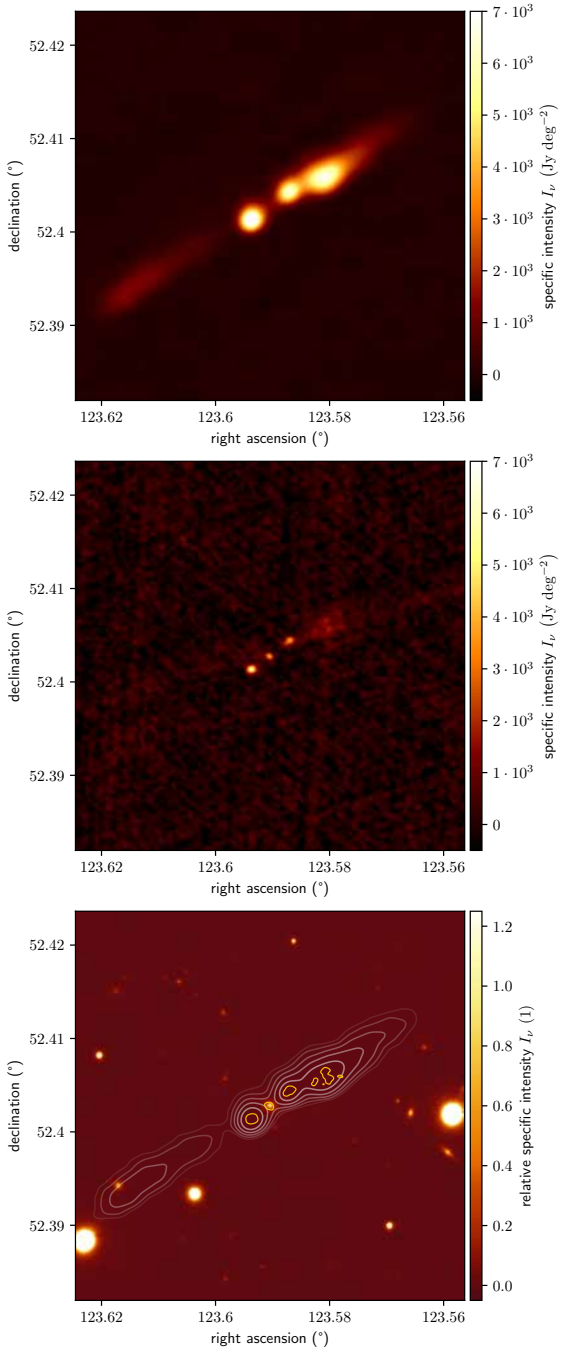


Figure 4.4: Radio and optical specific intensity function details around the SDSS DR12 source J081421.68+522410.0, Alcyoneus's host galaxy. The panels cover a $2.5' \times 2.5'$ region around the host, an elliptical galaxy with spectroscopic redshift $z_{\mathrm{spec}} = 0.24674 \pm 6 \cdot 10^{-5}$. From top to bottom, we show the LoTSS DR2 $6''$, the VLASS $2.2''$, and the Pan-STARRS DR1 i band relative to the host's specific intensity with LoTSS contours (white) as in Fig. 4.3 and a VLASS contour (gold) at 5σ.

density, we affirmed that Alcyoneus is a LERG. Through automated classification, Best & Heckman (2012) came to the same conclusion.

Being a jet-mode AGN, the super-massive black hole (SMBH) in the centre of Alcyoneus's host galaxy presumably accretes at an efficiency below 1% of the Eddington limit and is fuelled mainly by slowly cooling hot gas.

4.3.4 Projected proper length

We calculated Alcyoneus's projected proper length $l_{\rm p}$ through its angular length φ and spectroscopic redshift $z_{\rm spec}$. We formally determined $\varphi = 20.8' \pm 0.15'$ from the compact-source–subtracted 90″ image (top panel of Fig. 4.9) by selecting the largest great-circle distance between all possible pairs of pixels with a specific intensity higher than three sigma-clipped standard deviations above the sigma-clipped median. We find $l_{\rm p} = 4.99 \pm 0.04$ Mpc; this makes Alcyoneus the projectively largest RG known.

Methodology details and a probabilistic comparison between the projected proper lengths of Alcyoneus and J1420-0545 are given in Appendix 4.A1.

4.3.5 Radio luminosity densities and kinetic jet powers

From the LoTSS DR2 6″ image (top panel of Fig. 4.4), we measured that two northern jet local maxima occur at angular distances of $9.2 \pm 0.2''$ and $23.7 \pm 0.2''$ from the host, or at projected proper distances of 36.8 ± 0.8 kpc and 94.8 ± 0.8 kpc. Two southern jet local maxima occur at angular distances of $8.8 \pm 0.2''$ and $62.5 \pm 0.2''$ from the host, or at projected proper distances of 35.2 ± 0.8 kpc and 249.9 ± 0.8 kpc.

At the central observing frequency of $\nu_{\rm c} = 144$ MHz, the northern jet has a flux density $F_\nu = 193 \pm 20$ mJy, the southern jet has $F_\nu = 110 \pm 12$ mJy, whilst the northern lobe has $F_\nu = 63 \pm 7$ mJy and the southern lobe has $F_\nu = 44 \pm 5$ mJy. To minimise contamination from fore- and background galaxies, we determined the lobe flux densities from the compact-source–subtracted 90″ image. The flux density uncertainties are dominated by the 10% flux scale uncertainty inherent to the LoTSS DR2 (Shimwell et al., 2022). The host galaxy flux density is relatively weak, and the corresponding emission has, at $\nu_{\rm c} = 144$ MHz and 6″ resolution, no clear angular separation from the inner jets' emission; we have therefore not determined it.

Due to cosmological redshifting, the conversion between flux density and luminosity density depends on the spectral indices α of Alcyoneus's luminous components. We estimated the spectral indices of the core and jets from the LoTSS DR2 6″ and VLASS 2.2″ images. After convolving the VLASS image with a Gaussian to the common resolution of 6″, we calculated the mean spectral index between the LoTSS central frequency $\nu_{\rm c} = 144$ MHz and the VLASS central frequency $\nu_{\rm c} = 2.99$ GHz.

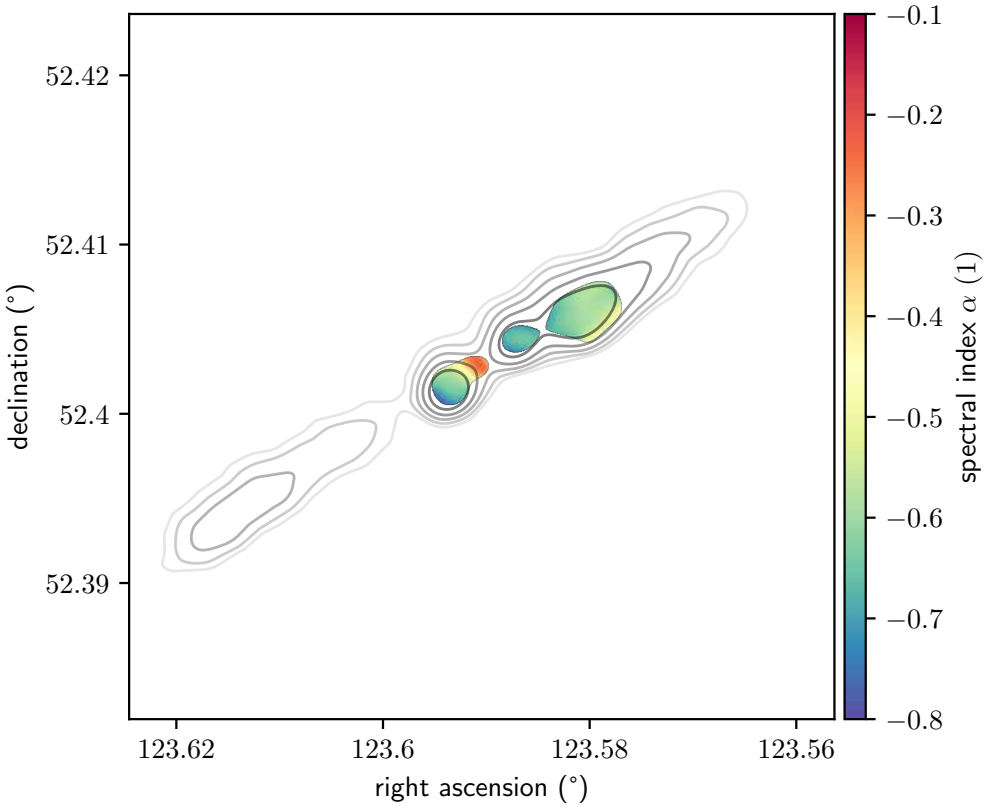


Figure 4.5: LoTSS–VLASS spectral index map, revealing Alcyoneus's flat-spectrum core and steeper-spectrum jets. We show all directions where both the LoTSS and VLASS images have at least 5σ significance. In black, we overlay the same LoTSS contours as in Figs. 4.3 and 4.4. The core spectral index is $\alpha = -0.25 \pm 0.1$, and the combined inner jet spectral index is $\alpha = -0.65 \pm 0.1$.

Using only directions for which both images have a significance of at least 5σ, we deduced a core spectral index $\alpha = -0.25 \pm 0.1$ and a combined inner jet spectral index $\alpha = -0.65 \pm 0.1$. The spectral index uncertainties are dominated by the LoTSS DR2 and VLASS flux scale uncertainties. We show the full spectral index map in Fig. 4.5. We have not determined the spectral index of the lobes, as they are only detected in the LoTSS imagery.

The luminosity densities of the northern and southern jet at rest-frame frequency $\nu = 144\,\mathrm{MHz}$ are $L_\nu = (3.6 \pm 0.4)\cdot 10^{25}\,\mathrm{W\,Hz^{-1}}$ and $L_\nu = (2.0 \pm 0.2)\cdot 10^{25}\,\mathrm{W\,Hz^{-1}}$, respectively. Following Dabhade et al. (2020a), we estimated the kinetic power of the jets from their luminosity densities and the results of the simulation-based analytical model of Hardcastle (2018). We find $Q_{\mathrm{jet},1} = 1.2 \pm 0.1 \cdot 10^{36}\,\mathrm{W}$ and $Q_{\mathrm{jet},2} = 6.6 \pm 0.7 \cdot 10^{35}\,\mathrm{W}$, so the total kinetic jet power is $Q_{\mathrm{jets}} := Q_{\mathrm{jet},1} + Q_{\mathrm{jet},2} = 1.9 \pm 0.2 \cdot 10^{36}\,\mathrm{W}$. Interestingly, this total kinetic jet power is *lower* than the average $Q_{\mathrm{jets}} = 3.7 \cdot 10^{36}\,\mathrm{W}$,

and close to the median $Q_{\text{jets}} = 2.2 \cdot 10^{36}$ W, for low-excitation GRGs in the redshift range $0.18 < z < 0.43$ (Dabhade et al., 2020a).

Because the lobe spectral indices are unknown, we present luminosity densities for several possible values of α in Table 4.1.[4] (Because of electron ageing, α will decrease further away from the core.)

Table 4.1: Luminosity densities L_ν (in 10^{24} W Hz^{-1}) of Alcyoneus's lobes for three potential spectral indices α at rest-frame frequency $\nu = 144$ MHz.

	$\alpha = -0.8$	$\alpha = -1.2$	$\alpha = -1.6$
Northern lobe	12 ± 1	13 ± 1	14 ± 1
Southern lobe	8.3 ± 0.8	9.0 ± 0.9	9.9 ± 1

Assuming $\alpha = -1.2$, Alcyoneus's total luminosity density at $\nu = 144$ MHz is $L_\nu = 7.8 \pm 0.8 \cdot 10^{25}$ W Hz^{-1}. In Fig. 4.6 we compare this estimate to other GRGs' total luminosity density at the same frequency, as found by Dabhade et al. (2020b) through the LoTSS DR1 (Shimwell et al., 2019). Interestingly, Alcyoneus is not particularly luminous: it has a low-frequency luminosity density typical for the currently known GRG population (percentile $45 \pm 3\%$).

4.3.6 True proper length from relativistic beaming

Following Hardcastle et al. (1998a), we simultaneously constrained Alcyoneus's jet speed u and inclination angle θ from the jets' flux density asymmetry: the northern-to-southern jet flux density ratio $J = 1.78 \pm 0.3$.[5] We had assumed that the jets propagate with identical speeds in exactly opposite directions (making line-of-sight angles θ and $\theta + 180°$) and have statistically identical relativistic electron populations; as such, they have a common synchrotron spectral index α. Using $\alpha = -0.65 \pm 0.1$ as before and

$$\beta := \frac{u}{c}; \;\; \beta \cos\theta = \frac{J^{\frac{1}{2-\alpha}} - 1}{J^{\frac{1}{2-\alpha}} + 1}, \tag{4.1}$$

we found $\beta \cos\theta = 0.106 \pm 0.03$. Because $\cos\theta \leq 1$, β is bounded from below by $\beta_{\text{min}} = 0.106 \pm 0.03$.

From detailed modelling of ten Fanaroff–Riley I RGs (which have jet luminosities

[4] The inferred luminosity densities have a cosmology dependence; our results are ~6% higher than for modern high-H_0 cosmologies.

[5] Because J is obtained through the division of two independent normal random variables (RVs) with non-zero mean, J is an RV with an uncorrelated non-central normal ratio distribution.

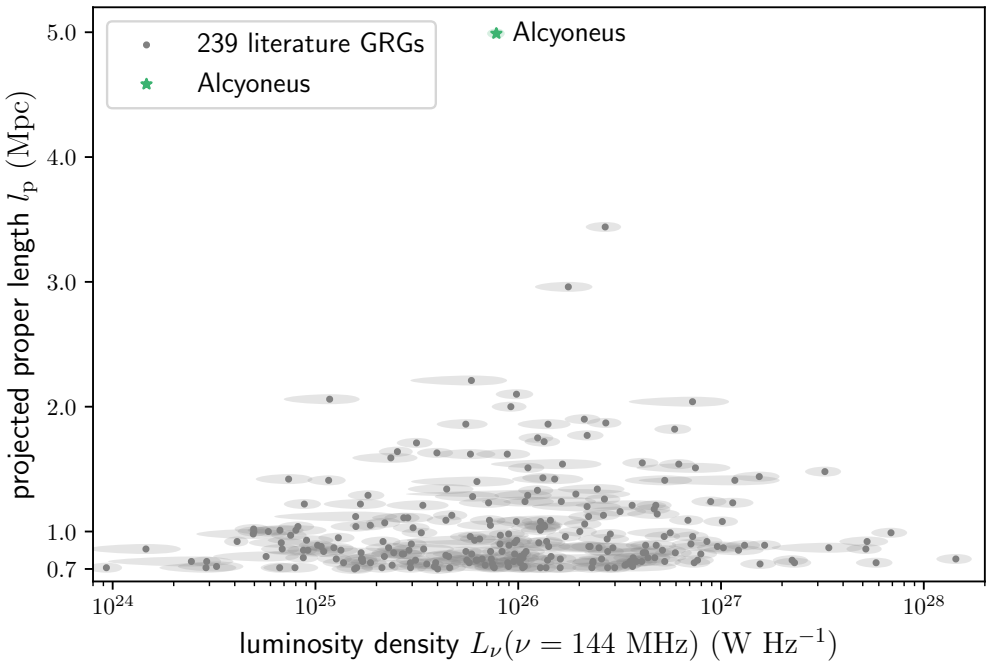


Figure 4.6: Relation between the GRG projected proper length l_p and total luminosity density L_ν at rest-frame frequency $\nu = 144$ MHz. Total luminosity densities include contributions from all available RG components (i.e. the core, jets, hotspots, and lobes). Literature GRGs are from Dabhade et al. (2020b) and are marked with grey disks, and Alcyoneus is marked with a green star. Translucent ellipses indicate -1 to +1 standard deviation uncertainties. Alcyoneus has a typical luminosity density (percentile $45 \pm 3\%$).

comparable to that of Alcyoneus), Laing & Bridle (2014) deduced that initial jet speeds are roughly $\beta = 0.8$, which decrease until roughly $0.6\ r_0$, with r_0 being the recollimation distance. Most of the ten recollimation distances from Laing & Bridle (2014) are between 5 and 15 kpc, with the largest being that of NGC 315: $r_0 = 35$ kpc. Because the local specific intensity maxima in Alcyoneus's jets closest to the host occur at projected proper distances of 36.8 ± 0.8 kpc and 35.2 ± 0.8 kpc, the true proper distances must be even larger. We conclude that the observed jet emission presumably comes from a region farther from the host than r_0, so the initial stage of jet deceleration — in which the jet speed is typically reduced by several tens of percents of c — must already be completed. Thus, $\beta_{\rm max} = 0.8$ is a safe upper bound.

Taking $\beta_{\rm max} = 0.8$, θ is bounded from above by $\theta_{\rm max} = 82.4 \pm 2^\circ$ ($\theta \in [0, 90^\circ]$), or bounded from below by $180^\circ - \theta_{\rm max} = 97.6 \pm 2^\circ$ ($\theta \in [90^\circ, 180^\circ]$).[6] If we model

[6] Taking $\beta_{\rm max} = 1$ instead, θ is bounded from above by $\theta_{\rm max} = 83.9 \pm 2^\circ$ ($\theta \in [0, 90^\circ]$), or bounded from below by $180^\circ - \theta_{\rm max} = 96.1 \pm 2^\circ$ ($\theta \in [90^\circ, 180^\circ]$).

Alcyoneus's geometry as a line segment, and assume no jet reorientation, Alcyoneus's true proper length l and projected proper length $l_{\rm p}$ relate as

$$l = \frac{l_{\rm p}}{\sin\theta};\ \ l \geq l_{\rm min} = \frac{l_{\rm p}}{\sin\theta_{\rm max}}. \tag{4.2}$$

We bounded l from below: $l_{\rm min} = 5.04 \pm 0.05$ Mpc. A triangular prior on β between $\beta_{\rm min}$ and $\beta_{\rm max}$ with the mode at $\beta_{\rm max}$ induces a skewed prior on l; the 90% credible interval is $l \in [5.0\ {\rm Mpc}, 5.5\ {\rm Mpc}]$, with the mean and median being 5.2 Mpc and 5.1 Mpc, respectively. A flat prior on β between $\beta_{\rm min}$ and $\beta_{\rm max}$ also induces a skewed prior on l; the 90% credible interval is $l \in [5.0\ {\rm Mpc}, 7.1\ {\rm Mpc}]$, with the mean and median being 5.6 Mpc and 5.1 Mpc, respectively. The median of l seems particularly well determined, as it is insensitive to variations in the prior on β.

In Appendix 4.A2, we explore the inclination angle conditions under which Alcyoneus has the largest true proper length of all known (> 4 Mpc) GRGs.

4.3.7 Stellar mass and super-massive black hole mass

The question then arises as to whether a galaxy or its central black hole needs to be massive in order to generate a GRG. Alcyoneus's host has a stellar mass $M_\star = 2.4 \pm 0.4 \cdot 10^{11}\ M_\odot$ (Chang et al., 2015). We tested whether or not this is a typical stellar mass among the total known GRG population. We assembled a literature catalogue of 1013 GRGs by merging the compendium of Dabhade et al. (2020a), which is complete up to April 2020, with the GRGs discovered in Galvin et al. (2020), Ishwara-Chandra et al. (2020), Tang et al. (2020), Bassani et al. (2021), Brüggen et al. (2021), Delhaize et al. (2021), Masini et al. (2021), Kuźmicz & Jamrozy (2021), Andernach et al. (2021) and Mahato et al. (2022). We collected stellar masses with uncertainties from Chang et al. (2015), which are based on SDSS and WISE photometry, and from Salim et al. (2018), which are based on Galaxy Evolution Explorer (GALEX), SDSS, and WISE photometry. We gave precedence to the stellar masses by Salim et al. (2018) when both were available. We obtained stellar masses for 151 previously known GRGs. The typical stellar mass range is $10^{11} - 10^{12}\ M_\odot$, the median $M_\star = 3.5 \cdot 10^{11}\ M_\odot$, and the mean $M_\star = 3.8 \cdot 10^{11}\ M_\odot$. Strikingly, the top panel of Fig. 4.7 illustrates that Alcyoneus's host has a fairly low (percentile $25 \pm 9\%$) stellar mass compared with the currently known population of GRG hosts.

For the GRGs in our literature catalogue, we also estimated SMBH masses via the M-sigma relation. We collected SDSS DR12 stellar velocity dispersions with uncertainties (Alam et al., 2015), and applied the M-sigma relation of Eq. 7 in Kormendy & Ho (2013). Alcyoneus's host has a SMBH mass $M_\bullet = 3.9 \pm 1.7 \cdot 10^8\ M_\odot$.

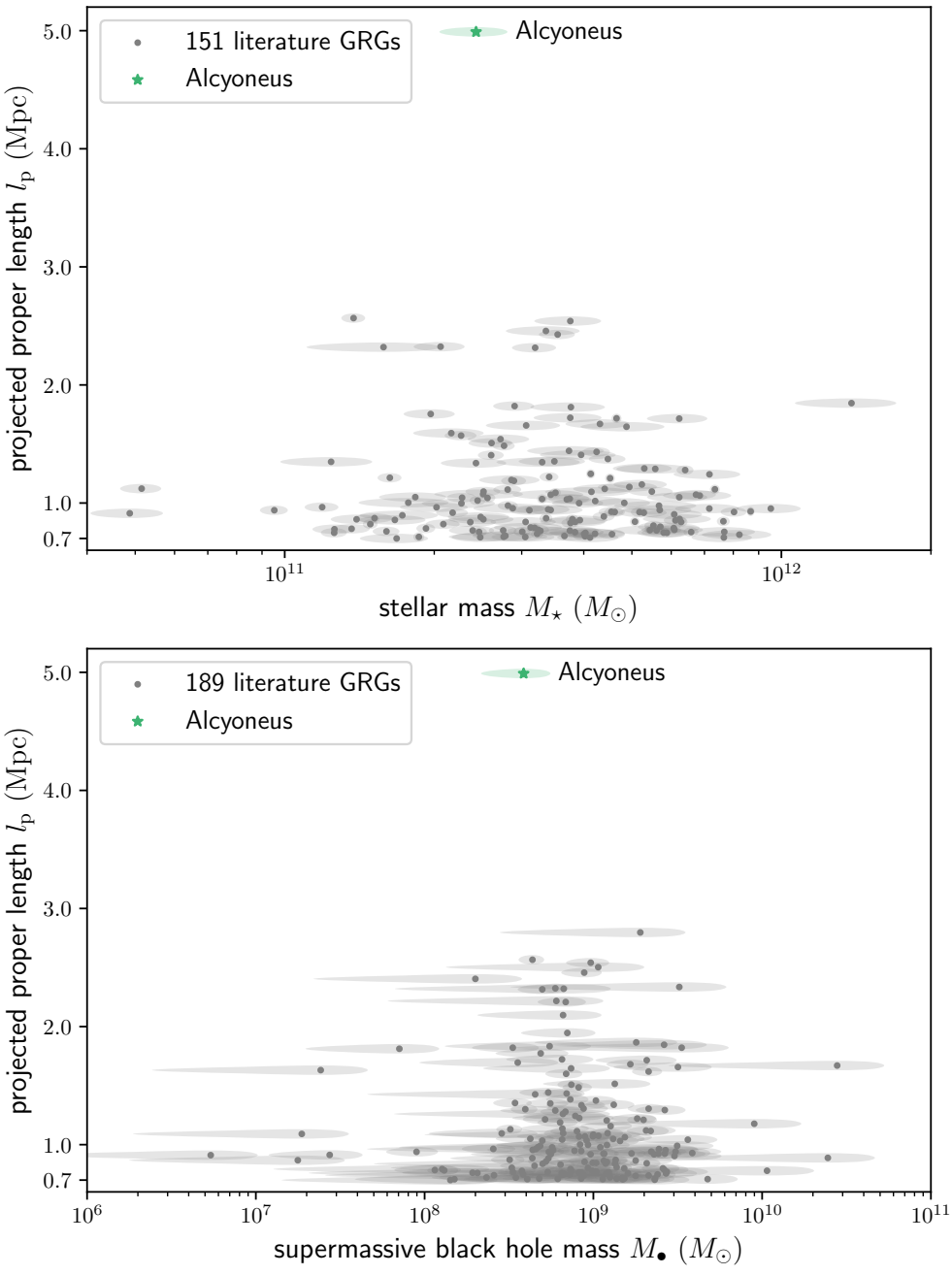


Figure 4.7: Relations between the GRG projected proper length l_p and the host galaxy stellar mass $M_\star$ (top panel) and between the GRG l_p and the host galaxy SMBH mass $M_\bullet$ (bottom panel). Our methods allow these properties to be determined for a small proportion of all literature GRGs only. Literature GRGs are marked with grey disks, and Alcyoneus is marked with a green star. Translucent ellipses indicate -1 to +1 standard deviation uncertainties. Alcyoneus's host has a below-average stellar mass (percentile $25 \pm 9\%$) and SMBH mass (percentile $23 \pm 11\%$).

We obtained SMBH masses for 189 previously known GRGs. The typical SMBH mass range is $10^8 - 10^{10}\ M_\odot$, the median $M_\bullet = 7.9 \cdot 10^8\ M_\odot$ and the mean $M_\bullet = 1.5 \cdot 10^9\ M_\odot$. Strikingly, the bottom panel of Fig. 4.7 illustrates that Alcyoneus's host has a fairly low (percentile $23 \pm 11\%$) SMBH mass compared with the currently known population of GRG hosts.

We note that Alcyoneus is the only GRG with $l_\mathrm{p} > 3$ Mpc whose host's stellar mass is known through Chang et al. (2015) or Salim et al. (2018), and whose host's SMBH mass can be estimated through its SDSS DR12 velocity dispersion. These data allow us to state confidently that exceptionally high stellar or SMBH masses are not necessary to generate 5-megaparsec-scale GRGs.

4.3.8 Surrounding large-scale structure

Several approaches to large-scale structure (LSS) classification, such as the T-web scheme (Hahn et al., 2007), partition the modern Universe into galaxy clusters, filaments, sheets, and voids. In this section, we determine Alcyoneus's most likely environment type using the SDSS DR7 spectroscopic galaxy sample (Abazajian et al., 2009).

In particular, we determined if Alcyoneus's host has fewer, about equal, or more galactic neighbours in SDSS DR7 than a randomly drawn galaxy of similar r-band luminosity density and redshift. Let $r(z)$ be the co-moving radial distance corresponding to cosmological redshift z. We considered a spherical shell with the observer at the centre, inner radius $\max\{r(z = z_\mathrm{spec}) - r_0, 0\}$ and outer radius $r(z = z_\mathrm{spec}) + r_0$. We approximated Alcyoneus's cosmological redshift with z_spec and chose $r_0 = 25$ Mpc. As all galaxies in the spherical shell have a similar distance to the observer (i.e. distances are at most $2r_0$ different), the SDSS DR7 galaxy number density completeness must also be similar throughout the spherical shell.[7] For each enclosed galaxy with an r-band luminosity density between $1 - \delta$ and $1 + \delta$ times that of Alcyoneus's host, we counted the number of SDSS DR7 galaxies $N_{<R}(R)$ within a sphere of co-moving radius R around it — regardless of luminosity density, and excluding itself. Alcyoneus's host has an SDSS r-band apparent magnitude $m_r = 18.20$; the corresponding luminosity density is $L_\nu(\lambda_\mathrm{c} = 623.1\ \mathrm{nm}) = 3.75 \cdot 10^{22}\ \mathrm{W\ Hz^{-1}}$. We chose $\delta = 0.25$; this yielded 9,358 such enclosed galaxies.

In Fig. 4.8, we show the distribution of $N_{<R}(R)$ for $R = 5$ Mpc and $R = 10$ Mpc. We verify that the distributions are insensitive to reasonable changes in r_0 and δ. We note that there is no SDSS DR7 galaxy within a co-moving distance of 5 Mpc from Alcyoneus's host. The nearest such galaxy, J081323.49+524856.1, occurs at a co-

[7] For $r_0 = 25$ Mpc, this is a good approximation, because the shell is cosmologically thin: $2r_0 = 50$ Mpc roughly amounts to the length of a single Cosmic Web filament.

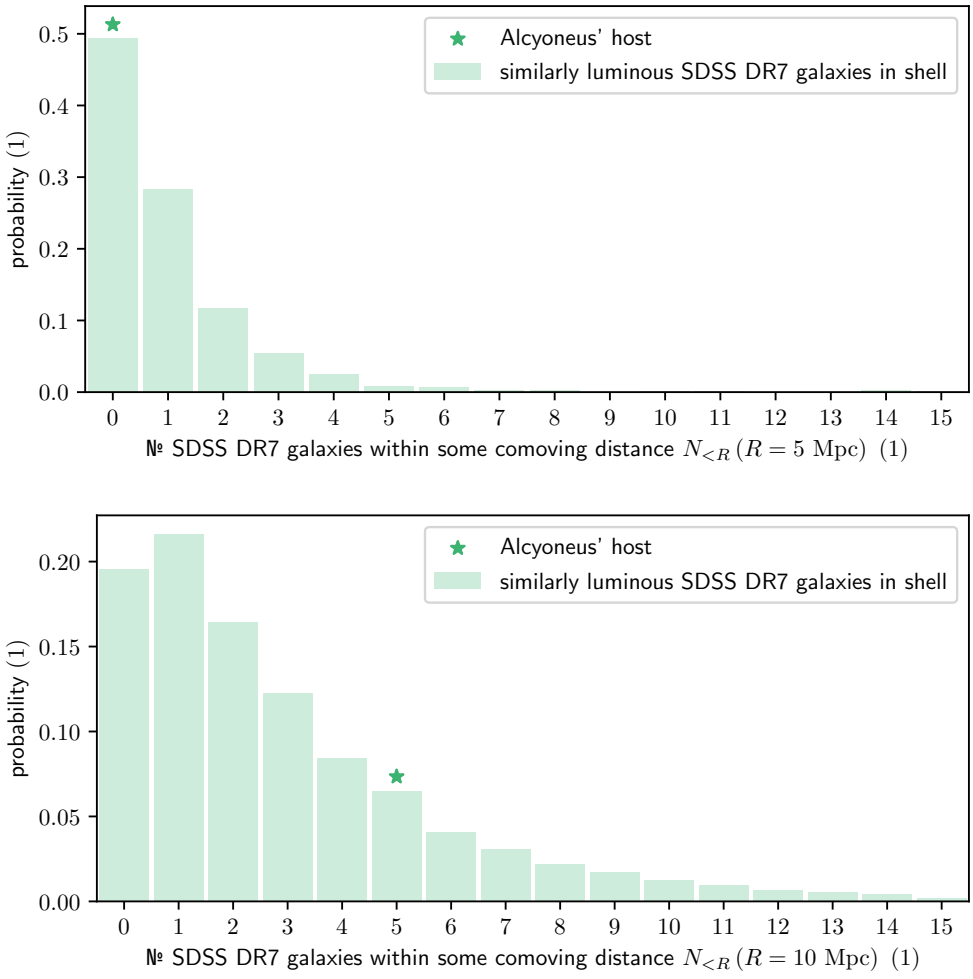


Figure 4.8: Probability mass functions tracing the environmental density distribution of luminous galaxies at Alcyoneus's redshift. Like most galaxies of similar r-band luminosity density and redshift, Alcyoneus's host has no galactic neighbours in SDSS DR7 within 5 Mpc. However, within 10 Mpc, Alcyoneus's host has more neighbours than most similar galaxies. For all 9,358 SDSS DR7 galaxies with an r-band luminosity density between 75% and 125% that of Alcyoneus's host and a co-moving radial distance that differs by at most $r_0 = 25$ Mpc from that of Alcyoneus, we count the number of SDSS DR7 galaxies, $N_{<R}\,(R)$, within a sphere of co-moving radius $R = 5$ Mpc (top panel) and $R = 10$ Mpc (bottom panel). The top panel indicates that Alcyoneus does not inhabit a galaxy cluster; the bottom panel indicates that Alcyoneus does not inhabit a void.

moving distance of 7.9 Mpc: the nearest $\sim$2, 000 Mpc^3 of co-moving space are free of galactic neighbours with $L_\nu\,(\lambda_c) > 5.57 \cdot 10^{22}$ W Hz^{-1}.[8] In the same way as in

[8]This is the luminosity density that corresponds to the SDSS r-band apparent magnitude com-

Sect. 4.3.1, we verified that the DESI Legacy Imaging Surveys DR9, RASS and PSZ2 do not contain evidence for a galaxy cluster in the direction of Alcyoneus's host. The nearest galaxy cluster, according to the SDSS-III cluster catalogue of Wen et al. (2012), instead lies 24′ away at right ascension 123.19926°, declination 52.72468° and photometric redshift $z_{\rm ph} = 0.2488$. It has an $R_{200} = 1.1$ Mpc and, according to the DESI cluster catalogue of Zou et al. (2021), a total mass $M = 2.2 \cdot 10^{14}\, M_\odot$. The co-moving distance between the cluster and Alcyoneus's host is 11 Mpc. All in all, we conclude that Alcyoneus does not reside in a galaxy cluster. Meanwhile, there are five SDSS DR7 galaxies within a co-moving distance of 10 Mpc from Alcyoneus's host: this makes it implausible that Alcyoneus lies in a void. Finally, one could interpret $N_{<R}(R)$ as a proxy for the LSS total matter density around a galaxy. For $R = 10$ Mpc, just 17% of galaxies in the shell with a similar luminosity density as Alcyoneus's host have a higher LSS total matter density. Being on the high end of the density distribution, but lying outside a cluster, Alcyoneus most probably inhabits a filament of the Cosmic Web.

4.3.9 Proper lobe volumes

We determined the proper volumes of Alcyoneus's lobes with a new Bayesian model. The model describes the lobes through a pair of doubly truncated, optically thin cones, each of which has a spatially constant and isotropic monochromatic emission coefficient (MEC; Rybicki & Lightman, 1986). We allowed the 3D orientations and opening angles of the cones to differ, as the lobes can traverse their way through differently pressured parts of the warm–hot inter-galactic medium (WHIM): for example, the medium near the filament axis and the medium near the surrounding voids. By adopting a spatially constant MEC, we neglected electron density and magnetic field inhomogeneities as well as spectral-ageing gradients; by adopting an isotropic MEC, we assumed non-relativistic velocities within the lobe so that beaming effects are negligible. Numerically, we first generated the GRG's 3D MEC field over a cubical voxel grid, and then calculated the corresponding model image through projection, including expansion-related cosmological effects. Before comparison with the observed image, we convolved the model image with a Gaussian kernel to the appropriate resolution. We exploited the approximately Gaussian LoTSS DR2 image noise to formulate the likelihood, and assumed a flat prior distribution over the parameters. Using Metropolis–Hastings (MH) Markov chain Monte Carlo (MCMC), we sam-

pleteness limit $m_r = 17.77$ (Strauss et al., 2002).

pled from the posterior distribution.[9]

In the top panel of Fig. 4.9, we show the LoTSS DR2 compact-source-subtracted 90″ image of Alcyoneus. The central region has been excluded from source subtraction, and hence Alcyoneus's core and jets remain. (However, when we run our MH MCMC on this image, we do mask this central region.) In the middle panel, we show the highest-likelihood (and thus maximum a posteriori (MAP)) model image before convolution. In the bottom panel, we show the same model image convolved to 90″ resolution, with 2σ and 3σ contours of the observed image overlaid. We provide the full parameter set that corresponds with this model in Table 4.2.

The posterior mean, calculated through the MH MCMC samples after burn-in, suggests the following geometry. The northern lobe has an opening angle $\gamma_1 = 10 \pm 1°$, and the cone truncates at an inner distance $d_{\mathrm{i},1} = 2.6 \pm 0.2$ Mpc and at an outer distance $d_{\mathrm{o},1} = 4.0 \pm 0.2$ Mpc from the host galaxy. The southern lobe has a larger opening angle $\gamma_2 = 26 \pm 2°$, but its cone truncates at smaller distances of $d_{\mathrm{i},2} = 1.5 \pm 0.1$ Mpc and $d_{\mathrm{o},2} = 2.0 \pm 0.1$ Mpc from the host galaxy. These parameters fix the proper volumes of Alcyoneus's northern and southern lobes. We find $V_1 = 1.5 \pm 0.2\ \mathrm{Mpc}^3$ and $V_2 = 1.0 \pm 0.2\ \mathrm{Mpc}^3$, respectively (see Eq. 4.20).[10]

Regarding the orientation of the lobes, Fig. 4.1 provides a visual hint that the lobes are subtly non-coaxial. The posterior indicates that the position angles of the northern and southern lobes are $\phi_1 = 307 \pm 1°$ and $\phi_2 = 139 \pm 2°$, respectively. The position angle difference is thus $\Delta\phi = 168 \pm 2°$: although close to $\Delta\phi = 180°$, we can reject the null hypothesis of coaxiality with high significance. Interestingly, the posterior also constrains the angles that the lobe axes make with the plane of the sky: $|\theta_1 - 90°| = 51 \pm 2°$ and $|\theta_2 - 90°| = 18 \pm 7°$. Again, the uncertainties imply that the lobes are probably not coaxial. We stress that these inclination angle results are tentative only. Future model extensions should explore how sensitive they are to the assumed lobe geometry (by testing other shapes than just truncated cones, such as ellipsoids).

One way to validate the model is to compare the observed lobe flux densities of Sect. 4.3.5 to the predicted lobe flux densities. According to the posterior, the MECs

[9] A detailed description of the model parameters, the MH MCMC, and the formulae for the derived quantities are given in Appendix 4.A3.

[10] As a sanity check, we compared our results to those from a less rigorous, though simpler, ellipsoid-based method of estimating volumes. By fitting ellipses to the image in the top panel of Fig. 4.9, one obtains a semi-minor and semi-major axis; the half-diameter along the ellipsoid's third dimension is assumed to be their mean. This method suggests a northern lobe volume $V_1 = 1.4 \pm 0.3\ \mathrm{Mpc}^3$ and a southern lobe volume $V_2 = 1.1 \pm 0.3\ \mathrm{Mpc}^3$. These results agree well with our Bayesian model results. (If the half-diameter along the third dimension is instead treated as an RV with a uniform distribution between the semi-minor axis and the semi-major axis, the estimates remain the same.)

of the northern and southern lobes are $j_{\nu,1} = 17 \pm 2$ Jy deg^{-2} Mpc^{-1} and $j_{\nu,2} = 18 \pm 3$ Jy deg^{-2} Mpc^{-1}. Combining MECs and volumes, we predict northern and southern lobe flux densities $F_{\nu,1}(\nu_c) = 63 \pm 4$ mJy and $F_{\nu,2}(\nu_c) = 45 \pm 5$ mJy (see Eq. 4.21). We find excellent agreement: the relative differences with the observed results are 0% and 2%, respectively.

4.3.10 Lobe pressures and the local WHIM

From Alcyoneus's lobe flux densities and volumes, we can infer lobe pressures and magnetic field strengths. We calculated these through `pysynch`[11] (Hardcastle et al., 1998b), which uses the formulae first proposed by Myers & Spangler (1985) and reexamined by Beck & Krause (2005). Following the notation of Hardcastle et al. (1998b), we assumed that the electron energy distribution is a power law in Lorentz factor γ with $\gamma_{\rm min} = 10$, $\gamma_{\rm max} = 10^4$ and exponent $p = -2$; we also assumed that the kinetic energy density of protons is vanishingly small compared with that of electrons ($\kappa = 0$) and that the plasma filling factor is unity ($\varphi = 1$). Assuming the minimum-energy condition (Burbidge, 1956), we find minimum-energy pressures $P_{\rm min,1} = 4.8 \pm 0.3 \cdot 10^{-16}$ Pa and $P_{\rm min,2} = 4.9 \pm 0.6 \cdot 10^{-16}$ Pa for the northern and southern lobes, respectively. The corresponding minimum-energy magnetic field strengths are $B_{\rm min,1} = 46 \pm 1$ pT and $B_{\rm min,2} = 46 \pm 3$ pT. Assuming the equipartition condition (Pacholczyk, 1970), we find equipartition pressures $P_{\rm eq,1} = 4.9 \pm 0.3 \cdot 10^{-16}$ Pa and $P_{\rm eq,2} = 4.9 \pm 0.6 \cdot 10^{-16}$ Pa for the northern and southern lobes, respectively. The corresponding equipartition magnetic field strengths are $B_{\rm eq,1} = 43 \pm 2$ pT and $B_{\rm eq,2} = 43 \pm 2$ pT. The minimum-energy and equipartition results do not differ significantly.

From pressures and volumes, we estimated the internal energy of the lobes $E = 3PV$. We find $E_{\rm min,1} = 6.2 \pm 0.5 \cdot 10^{52}$ J, $E_{\rm min,2} = 4.3 \pm 0.6 \cdot 10^{52}$ J, $E_{\rm eq,1} = 6.3 \pm 0.5 \cdot 10^{52}$ J and $E_{\rm eq,2} = 4.4 \pm 0.6 \cdot 10^{52}$ J. Next, we could bound the ages of the lobes from below by neglecting synchrotron losses and assuming that the jets have been injecting energy in the lobes continuously at the currently observed kinetic jet powers. Using $\Delta t = EQ_{\rm jet}^{-1}$, we find $\Delta t_{\rm min,1} = 1.7 \pm 0.2$ Gyr, $\Delta t_{\rm min,2} = 2.1 \pm 0.4$ Gyr, and identical results when assuming the equipartition condition. Finally, we could obtain a rough estimate of the average expansion speed of the RG during its lifetime $u = l_{\rm p}(\Delta t)^{-1}$. We find $u = 2.6 \pm 0.3 \cdot 10^3$ km s^{-1}, or about 1% of the speed of light.

Several other authors (Andernach et al., 1992; Lacy et al., 1993; Subrahmanyan et al., 1996; Parma et al., 1996; Mack et al., 1998; Schoenmakers et al., 1998, 2000; Ishwara-Chandra & Saikia, 1999; Lara et al., 2000; Machalski & Jamrozy, 2000; Machal-

[11] The `pysynch` code is publicly available online: `https://github.com/mhardcastle/pysynch`.

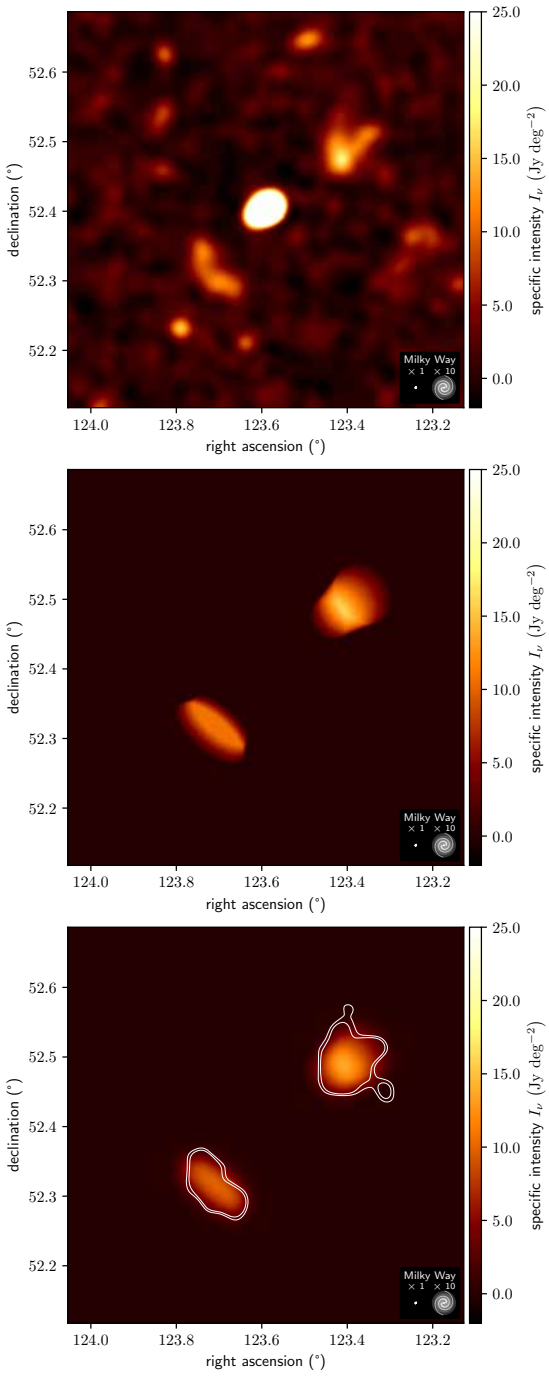


Figure 4.9: Bayesian model overview, showing how Alcyoneus's lobe volumes can be estimated by comparing an observed radio image to modelled radio images. **Top:** LoTSS DR2 compact-source-subtracted 90″ image of Alcyoneus. **Middle:** Highest-likelihood model image. **Bottom:** Same model image convolved to 90″ resolution, with 2σ and 3σ contours of the observed image overlaid.

ski et al., 2001; Saripalli et al., 2002; Jamrozy et al., 2005; Subrahmanyan et al., 2006, 2008; Saikia et al., 2006; Machalski et al., 2006, 2007, 2008; Safouris et al., 2009; Malarecki et al., 2013; Tamhane et al., 2015; Sebastian et al., 2018; Heesen et al., 2018; Cantwell et al., 2020) have estimated the minimum-energy or equipartition pressure of the lobes of GRGs embedded in non-cluster environments (i.e. in voids, sheets or filaments of the Cosmic Web). We compare Alcyoneus to the other 151 GRGs with known lobe pressures in the top panel of Fig. 4.10.[12] Alcyoneus reaffirms the negative correlation between length and lobe pressure (Jamrozy & Machalski, 2002; Machalski & Jamrozy, 2006), and it has the lowest lobe pressures found thus far. Alcyoneus's lobe pressures are in fact so low that they are comparable to the pressure in dense and hot parts of the WHIM: for a baryonic matter (BM) density $\rho_{\rm WHIM} = 10\rho_{\rm c,0}\Omega_{\rm BM,0}$ and $T_{\rm WHIM} = 10^7$ K, $P_{\rm WHIM} = 4 \cdot 10^{-16}$ Pa. Here, $\rho_{\rm c,0}$ is today's critical density, so $\rho_{\rm c,0}\Omega_{\rm BM,0}$ is today's mean baryon density. A more extensive comparison between $P_{\rm min}$ (green line) and $P_{\rm WHIM}$ (red lines) is shown in the bottom panel of Fig. 4.10. For comparison, we also show the lobe pressures of the four other thus-analysed GRGs with $l_{\rm p} > 3$ Mpc (grey lines). These are J1420-0545 of $l_{\rm p} = 4.9$ Mpc (Machalski et al., 2008), 3C 236 of $l_{\rm p} = 4.7$ Mpc (Schoenmakers et al., 2000), J0331-7710 of $l_{\rm p} = 3.4$ Mpc (Malarecki et al., 2013) and B2147+816 of $l_{\rm p} = 3.1$ Mpc (Schoenmakers et al., 2000).

Although proposed as probes of WHIM thermodynamics for decades, the bottom panel of Fig. 4.10 demonstrates that even the largest non-cluster literature GRGs are unlikely to be in pressure equilibrium with their environment. Relying on results from the Overwhelmingly Large Simulations (OWLS) (Schaye et al., 2010), Malarecki et al. (2013) point out that baryon densities $\rho_{\rm BM} > 50\rho_{\rm c,0}\,\Omega_{\rm BM,0}$, which are necessary for pressure equilibrium in these GRGs (see the intersection of grey and red lines in the bottom panel of Fig. 4.10), occur in only 1% of the WHIM's volume. By contrast, Alcyoneus can be in pressure equilibrium with the WHIM at baryon densities $\rho_{\rm BM} \sim 20\rho_{\rm c,0}\,\Omega_{\rm BM,0}$ and, thus, represents the most promising inter-galactic barometer of its kind yet.[13]

A final question is why most, if not all, observed non-cluster GRGs have overpressured lobes. The top panel of Fig. 4.10 suggests that GRGs must grow to several megaparsecs to approach WHIM pressures in their lobes, and such GRGs are rare.

[12] We have included all publications that provide pressures, energy densities or magnetic field strengths. We note that some authors assumed $\gamma_{\rm min} = 1$, we assumed $\gamma_{\rm min} = 10$ and Malarecki et al. (2013) assumed $\gamma_{\rm min} = 10^3$. If possible, angular lengths were updated using the LoTSS DR2 at 6″ and redshift estimates were updated using the SDSS DR12. All projected proper lengths have been recalculated using our Planck Collaboration et al. (2020) cosmology. When authors provided pressures for both lobes, we have taken the average.

[13] At Alcyoneus's redshift, this density amounts to a baryon overdensity of ∼10.

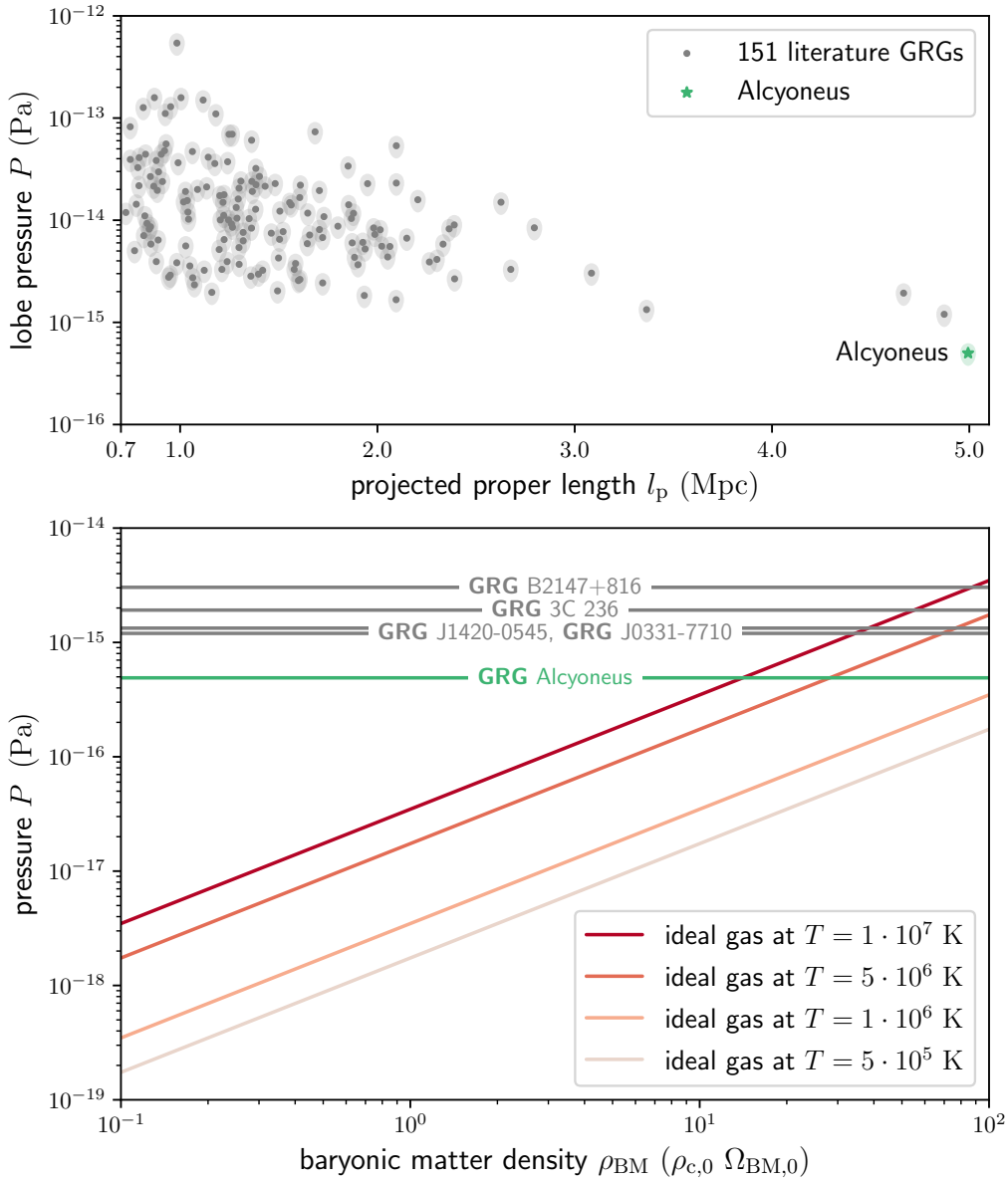


Figure 4.10: Panels showing that, of all GRGs with known lobe pressures, Alcyoneus is the most plausible candidate for pressure equilibrium with the WHIM. In the top panel, we explore the relation between length and lobe pressure for Alcyoneus and 151 literature GRGs. In the bottom panel, we compare the lobe pressure of Alcyoneus (green line) with WHIM pressures (red lines). For reference, we also show the lobe pressures of the four largest similarly analysed GRGs (grey lines).

However, the primary reason is the limited surface brightness sensitivity of all past and current surveys. Alcyoneus's lobes are visible in the LoTSS, but not in the NRAO VLA Sky Survey (NVSS) (Condon et al., 1998) or in the Westerbork Northern Sky

Survey (WENSS) (Rengelink et al., 1997). Their pressures approach that of the bulk of the WHIM within an order of magnitude. Lobes with even lower pressure must be less luminous or more voluminous, and thus will have even lower surface brightness. It is therefore probable that most GRG lobes that are in true pressure equilibrium with the WHIM still lie hidden in the radio sky.

4.4 Conclusion

In this work, we have presented the discovery of a radio galaxy of at least 5 Mpc. We have measured its key physical characteristics, and explored possible reasons for its exceptional growth.

1. We reprocessed the LoTSS DR2, the latest version of the LOFAR's northern sky survey at 144 MHz, by subtracting angularly compact sources and imaging at 60″ and 90″ resolution. The resulting images (Oei et al., prepa) allow us to explore a new sensitivity regime for RG lobes and thus represent promising data to search for unknown GRGs of large angular length. We will present a sample in forthcoming work.

2. We have discovered the first 5 Mpc GRG, which we dub 'Alcyoneus'. The projected proper length is $l_{\rm p} = 4.99 \pm 0.04$ Mpc, while the true proper length is at least $l_{\rm min} = 5.04 \pm 0.05$ Mpc. We confidently associate the $20.8' \pm 0.15'$ radio structure with an elliptical galaxy with a jet-mode AGN detected in the DESI Legacy Imaging Surveys DR9: the SDSS DR12 source J081421.68+522410.0 at J2000 right ascension 123.590372°, declination 52.402795°, and spectroscopic redshift $0.24674 \pm 6 \cdot 10^{-5}$.

3. Alcyoneus has a total luminosity density at $\nu = 144$ MHz of $L_\nu = 8 \pm 1 \cdot 10^{25}$ W Hz^{-1}, which is typical for GRGs (percentile $45 \pm 3\%$). Alcyoneus's host has a fairly low stellar mass and SMBH mass compared with other GRG hosts (percentiles $25 \pm 9\%$ and $23 \pm 11\%$). This implies that — within the GRG population — no strong positive correlation between RG length and (instantaneous) low-frequency radio power, stellar mass, or SMBH mass can exist.

4. A poly-chromatic examination of the surrounding sky — using the RASS, the DESI Legacy Imaging Surveys DR9, the PSZ2, and the LoTSS DR2 — suggests that Alcyoneus does not inhabit a galaxy cluster. According to an SDSS-III cluster catalogue, the nearest cluster occurs at a co-moving distance

of 11 Mpc. A local galaxy number density count suggests that Alcyoneus instead inhabits a filament of the Cosmic Web. A low-density environment therefore remains a possible explanation for Alcyoneus's formidable size.

5. We have developed a new Bayesian model that parametrises in three dimensions a pair of arbitrarily oriented, optically thin, doubly truncated conical RG lobes with a constant MEC. We then generated the corresponding specific intensity function, taking cosmic expansion into account, and compared it to data, assuming Gaussian image noise. We used MH MCMC to optimise the parameters and thus determined northern and southern lobe volumes of 1.5 ± 0.2 Mpc^3 and 1.0 ± 0.2 Mpc^3, respectively. In total, the lobes have an internal energy of $\sim 10^{53}$ J, expelled from the host galaxy over a gigayear-scale period. The lobe pressures are $4.8 \pm 0.3 \cdot 10^{-16}$ Pa and $4.9 \pm 0.6 \cdot 10^{-16}$ Pa, respectively; these are the lowest measured in radio galaxies yet. Nevertheless, the lobe pressures still exceed a large range of plausible WHIM pressures. Most likely, the lobes are still expanding — and Alcyoneus's struggle for supremacy of the cosmos continues.

M.S.S.L. Oei warmly thanks Frits Sweijen for coding the very useful https://github.com/tikk3r/legacystamps.
M.S.S.L. Oei, R.J. van Weeren and A. Botteon acknowledge support from the VIDI research programme with project number 639.042.729, which is financed by the Dutch Research Council (NWO). M. Brüggen acknowledges support from the Deutsche Forschungsgemeinschaft under Germany's Excellence Strategy — EXC 2121 'Quantum Universe' — 390833306. W.L. Williams acknowledges support from the CAS–NWO programme for radio astronomy with project number 629.001.024, which is financed by the Dutch Research Council (NWO).
The LOFAR is the Low-Frequency Array designed and constructed by ASTRON. It has observing, data processing, and data storage facilities in several countries, which are owned by various parties (each with their own funding sources), and which are collectively operated by the ILT Foundation under a joint scientific policy. The ILT resources have benefited from the following recent major funding sources: CNRS–INSU, Observatoire de Paris and Université d'Orléans, France; BMBF, MIWF–NRW, MPG, Germany; Science Foundation Ireland (SFI), Department of Business, Enterprise and Innovation (DBEI), Ireland; NWO, The Netherlands; the Science and Technology Facilities Council, UK; Ministry of Science and Higher Education, Poland; the Istituto Nazionale di Astrofisica (INAF), Italy.
The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc. CIRADA is funded by a grant from the Canada Foundation for Innovation 2017 Innovation Fund (Project 35999), as well as by the Provinces of Ontario, British Columbia, Alberta, Manitoba and Quebec.
Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is http://www.sdss3.org/. SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofisica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.
The Pan-STARRS1 Surveys (PS1) and the PS1 public science archive have been made possible through contributions by the Institute for Astronomy, the University of Hawaii, the Pan-STARRS Project Office, the Max-Planck Society and its participating institutes, the Max Planck Institute for Astronomy, Heidelberg and the Max Planck Institute for Extraterrestrial Physics,

Garching, The Johns Hopkins University, Durham University, the University of Edinburgh, the Queen's University Belfast, the Harvard-Smithsonian Center for Astrophysics, the Las Cumbres Observatory Global Telescope Network Incorporated, the National Central University of Taiwan, the Space Telescope Science Institute, the National Aeronautics and Space Administration under Grant No. NNX08AR22G issued through the Planetary Science Division of the NASA Science Mission Directorate, the National Science Foundation Grant No. AST-1238877, the University of Maryland, Eotvos Lorand University (ELTE), the Los Alamos National Laboratory, and the Gordon and Betty Moore Foundation.
This publication makes use of data products from the Wide-Field Infrared Survey Explorer, which is a joint project of the University of California, Los Angeles, and the Jet Propulsion Laboratory/California Institute of Technology, funded by the National Aeronautics and Space Administration.
The Legacy Surveys consist of three individual and complementary projects: the Dark Energy Camera Legacy Survey (DECaLS; Proposal ID #2014B-0404; PIs: David Schlegel and Arjun Dey), the Beijing–Arizona Sky Survey (BASS; NOAO Prop. ID #2015A-0801; PIs: Zhou Xu and Xiaohui Fan), and the Mayall z-band Legacy Survey (MzLS; Prop. ID #2016A-0453; PI: Arjun Dey). DECaLS, BASS and MzLS together include data obtained, respectively, at the Blanco telescope, Cerro Tololo Inter-American Observatory, NSF's NOIRLab; the Bok telescope, Steward Observatory, University of Arizona; and the Mayall telescope, Kitt Peak National Observatory, NOIRLab. The Legacy Surveys project is honored to be permitted to conduct astronomical research on Iolkam Du'ag (Kitt Peak), a mountain with particular significance to the Tohono O'odham Nation. NOIRLab is operated by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation. This project used data obtained with the Dark Energy Camera (DECam), which was constructed by the Dark Energy Survey (DES) collaboration. Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundacao Carlos Chagas Filho de Amparo, Financiadora de Estudos e Projetos, Fundacao Carlos Chagas Filho de Amparo a Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Cientifico e Tecnologico and the Ministerio da Ciencia, Tecnologia e Inovacao, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey. The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energeticas, Medioambientales y Tecnologicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenössische Technische Hochschule (ETH) Zürich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciencies de l'Espai (IEEC/CSIC), the Institut de Fisica d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig Maximilians Universität München and the associated Excellence Cluster Universe, the University of Michigan, NSF's NOIRLab, the University of Nottingham, the Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, and Texas A&M University. BASS is a key project of the Telescope Access Program (TAP), which has been funded by the National Astronomical Observatories of China, the Chinese Academy of Sciences (the Strategic Priority Research Program "The Emergence of Cosmological Structures" Grant # XDB09000000), and the Special Fund for Astronomy from the Ministry of Finance. The BASS is also supported by the External Cooperation Program of Chinese Academy of Sciences (Grant # 114A11KYSB20160057), and Chinese National Natural Science Foundation (Grant # 11433005). The Legacy Survey team makes use of data products from the Near-Earth Object Wide-Field Infrared Survey Explorer (NEOWISE), which is a project of the Jet Propulsion Laboratory/California Institute of Technology. NEOWISE is funded by the National Aeronautics and Space Administration. The Legacy Surveys imaging of the DESI footprint is supported by the Director, Office of Science, Office of High Energy Physics of the U.S. Department of Energy under Contract No. DE-AC02-05CH1123, by the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility under the same contract; and by the U.S. National Science Foundation, Division of Astronomical Sciences under Contract No. AST-0950945 to NOAO.

4.A1 J1420-0545 comparison

We verified that Alcyoneus is the longest known RG in projection by comparing it with J1420-0545 (Machalski et al., 2008), the literature's record holder.

The angular lengths of Alcyoneus and J1420-0545 are $\varphi = 20.8' \pm 0.15'$ and $\varphi = 17.4' \pm 0.05'$, respectively. For J1420-0545, we adopted the angular length reported by Machalski et al. (2008) because it lies outside the LoTSS DR2 coverage. The spectroscopic redshifts of Alcyoneus and J1420-0545 are $z_{\rm spec} = 0.24674 \pm 6 \cdot 10^{-5}$ and $z_{\rm spec} = 0.3067 \pm 5 \cdot 10^{-4}$, respectively. For both giants, we assumed the peculiar velocity along the line of sight $u_{\rm p}$ to be a Gaussian random variable (RV) with mean 0 and standard deviation 100 km s^{-1}, similar to conditions in low-mass galaxy clusters. (Both giants actually reside in filaments.)

Equations 7.23 describe how to calculate the cosmological redshift RV z via the peculiar velocity redshift RV $z_{\rm p}$:

$$\beta_{\rm p} := \frac{u_{\rm p}}{c};\ z_{\rm p} = \sqrt{\frac{1+\beta_{\rm p}}{1-\beta_{\rm p}}} - 1;\ z = \frac{1+z_{\rm spec}}{1+z_{\rm p}} - 1. \tag{4.3}$$

Here, c is the speed of light *in vacuo*. Finally, we calculated the projected proper length RV $l_{\rm p} = r_\varphi(z, \mathfrak{M}) \cdot \varphi$. Here, r_φ is the angular diameter distance RV, which depends on cosmological model parameters $\mathfrak{M}$. Propagating the uncertainties in angular length φ, spectroscopic redshift $z_{\rm spec}$ and peculiar velocity along the line of sight $u_{\rm p}$ through Monte Carlo simulation, the projected proper lengths of Alcyoneus and J1420-0545 are $l_{\rm p} = 4.99 \pm 0.04$ Mpc and $l_{\rm p} = 4.87 \pm 0.02$ Mpc, respectively.

We show the two projected proper length probability density functions (PDFs) in Fig. 4.11. The probability that Alcyoneus has the largest projected proper length is 99.9%.[14]

4.A2 Inclination angle comparison

In this appendix, we determine under what conditions Alcyoneus is not only the longest GRG in the plane of the sky but also in three dimensions. To this end, we compared Alcyoneus to the five previously known GRGs with projected proper lengths above 4 Mpc, which we dub 'challengers'. A challenger surpasses Alcyoneus in true proper length when

$$l_{\rm c} > l,\ \text{or}\ \frac{l_{\rm p,c}}{\sin\theta_{\rm c}} > \frac{l_{\rm p}}{\sin\theta},\ \text{or}\ \sin\theta_{\rm c} < \frac{l_{\rm p,c}}{l_{\rm p}}\sin\theta, \tag{4.4}$$

[14] This result is insensitive to plausible changes in cosmological parameters; for example, the high-H_0 (i.e. $H_0 > 70$ km s^{-1} Mpc^{-1}) cosmology with $\mathfrak{M} = (h = 0.7020, \Omega_{\rm BM,0} = 0.0455, \Omega_{\rm M,0} = 0.2720, \Omega_{\Lambda,0} = 0.7280)$ yields a probability of 99.8%.

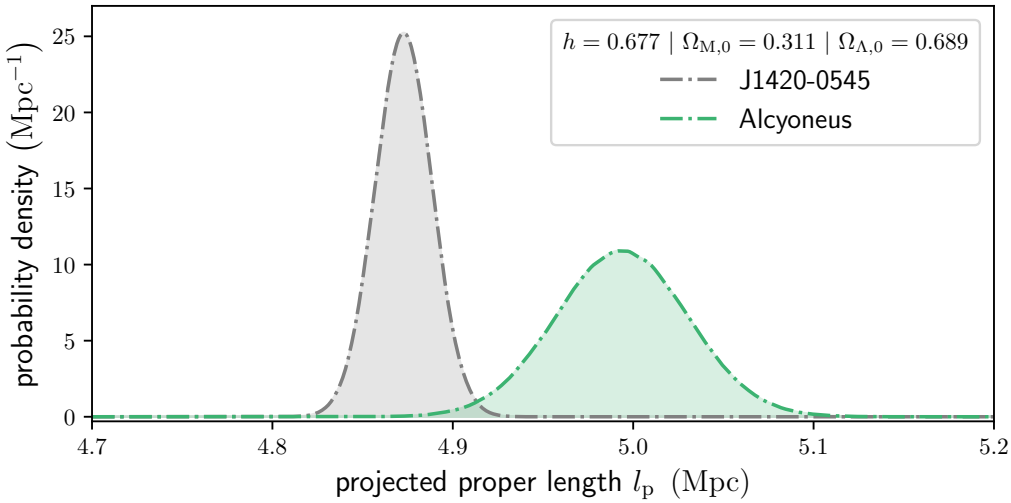


Figure 4.11: PDFs showing that Alcyoneus's projected proper length just exceeds that of J1420-0545. The probability that Alcyoneus (green) has a larger projected proper length than J1420-0545 (grey) (Machalski et al., 2008) is 99.9%. For both GRGs, we take into account uncertainty in angular length and spectroscopic redshift, as well as the possibility of peculiar motion along the line of sight.

where l_{c}, $l_{\mathrm{p,c}}$ and θ_{c} are the challenger's true proper length, projected proper length and inclination angle, respectively. Because the arcsine is a monotonically increasing function, a challenger surpasses Alcyoneus if its inclination angle obeys

$$\theta_{\mathrm{c}} < \theta_{\mathrm{max,c}}(\theta)\,, \text{ where } \theta_{\mathrm{max,c}}(\theta) := \arcsin\left(\frac{l_{\mathrm{p,c}}}{l_{\mathrm{p}}}\sin\theta\right). \tag{4.5}$$

In Fig. 4.12 we show $\theta_{\mathrm{max,c}}(\theta)$ for the five challengers with $l_{\mathrm{p,c}} \in \{4.11\ \mathrm{Mpc},\ 4.35\ \mathrm{Mpc},\ 4.60\ \mathrm{Mpc},\ 4.72\ \mathrm{Mpc},\ 4.87\ \mathrm{Mpc}\}$ (coloured curves). Alcyoneus is least likely to be the longest GRG in 3D when its true proper length equals its projected proper length (i.e. when $\theta = 90°$). The challengers then surpass Alcyoneus in true proper length when their inclination angles are less than 55°, 61°, 67°, 71°, and 77°, respectively. For $\theta < 90°$, the conditions are more stringent.

The third and fourth longest challengers, whose respective SDSS DR12 host names are J100601.73+345410.5 and J093139.03+320400.1, harbour quasars in their host galaxies. If small inclination angles distinguish quasars from non-quasar AGN, as proposed by the unification model (e.g. Hardcastle & Croston, 2020), these two challengers may well be the longest radio galaxies in three dimensions.

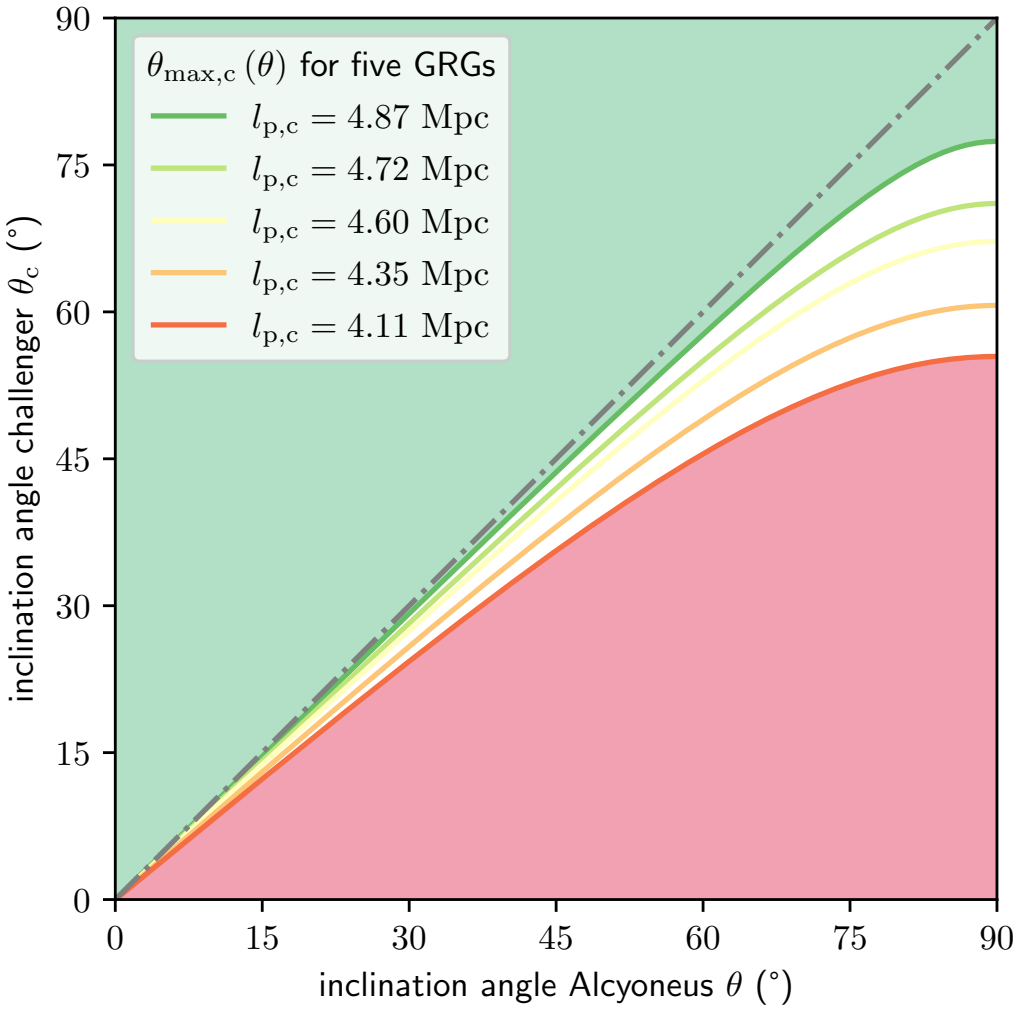


Figure 4.12: Schematic of conditions under which Alcyoneus is not only the longest GRG in the plane of the sky but also in three dimensions. Alcyoneus's inclination angle, θ, is not well determined, and therefore the full range of possibilities is shown on the horizontal axis. To surpass Alcyoneus in true proper length, challengers must have smaller inclination angles than Alcyoneus (i.e. appear below the dotted grey equality line). More specifically, as a function of θ, we show the inclination angle $\theta_{\mathrm{max,c}}$ below which challengers with a projected proper length $l_{\mathrm{p,c}} > 4$ Mpc trump Alcyoneus (coloured curves). The shaded areas of parameter space represent regimes with a particularly straightforward interpretation. One can imagine populating the graph with five points (located along the same vertical line), representing the ground-truth inclination angles of Alcyoneus and its five challengers. If any of these points fall in the red-shaded area, Alcyoneus is not the longest GRG in 3D. If all points fall in the green-shaded area, Alcyoneus is the longest GRG in 3D.

4.A3 Lobe volumes with truncated double cone model

4.A3.1 Synopsis

We built an MH MCMC model, similar in spirit to the model of Boxelaar et al. (2021) for galaxy cluster halos, in order to formalise the determination of RG lobe volumes from a radio image. To this end, we introduced a parametrisation of a pair of 3D RG lobes, and explored the corresponding parameter space via the Metropolis algorithm.[15] For each parameter tuple encountered during exploration, we first calculated the MEC function of the lobes on a uniform 3D grid representing a proper (rather than co-moving) cubical volume. The RG is assumed to be far enough from the observer that the conversion to a 2D image through ray tracing simplifies to summing up the cube's voxels along one dimension, and applying a cosmological attenuation factor. This factor depends on the galaxy's cosmological redshift, which is a hyper-parameter. We blurred the model image to the resolution of the observed image, which is also a hyper-parameter. Next, we calculated the likelihood that the observed image is a noisy version of the proposed model image. We assumed thermal noise to be the dominant type of noise. If one divides the imaged sky region into patches with a solid angle equal to that of the point spread function (PSF), then the noise per patch is approximately an independent Gaussian RV. These RVs have zero mean and share the same variance, which is another hyper-parameter — typically obtained from the observed image. We chose a uniform prior over the full physically realisable part of parameter space. The resulting posterior, which contains both geometric and radiative parameters, allows one to calculate probability distributions for many interesting quantities, such as the RG's lobe volumes and inclination angle. The inferences depend weakly on cosmological parameters $\mathfrak{M}$. Furthermore, their reliability depends significantly on the validity of the model assumptions.

4.A3.2 Model

Geometry

We modelled each lobe in 3D with a truncated right circular cone with apex $\mathcal{O} \in \mathbb{R}^3$, central axis unit vector $\hat{a} \in \mathbb{S}^2$, and opening angle $\gamma \in [0, \frac{\pi}{2}]$, as in Fig. 4.9. The lobes share the same $\mathcal{O}$, which is the RG host location. Each central axis unit vector can be parametrised through a position angle $\phi \in [0, 2\pi)$ and an inclination angle $\theta \in [0, \pi]$. Each cone is truncated twice, through planes that intersect the cone per-

[15] The more general MH variant need not be considered, as we work with a symmetric proposal distribution.

pendicularly to its central axis. Thus, each truncation is parametrised by the distance from the apex to the point where the plane intersects the central axis. The two inner ($d_{\mathrm{i},1}, d_{\mathrm{i},2} \in \mathbb{R}_{\geq 0}$) and two outer ($d_{\mathrm{o},1}, d_{\mathrm{o},2} \in \mathbb{R}_{\geq 0}$) truncation distances are parameters that we allowed to vary independently, with the only constraint that each inner truncation distance could not exceed the corresponding outer truncation distance.

Radiative processes

The radiative formulation of our model is among the simplest possible. The radio emission from the lobes is synchrotron radiation. We approximated the lobes to be perfectly optically thin and so neglected synchrotron self-absorption. The proper MEC is assumed spatially constant throughout a lobe, though possibly different among lobes; this leads to parameters $j_{\nu,1}, j_{\nu,2} \in \mathbb{R}_{\geq 0}$. The relationship between the specific intensity I_ν (in direction $\hat{r}$ at central frequency ν_{c}) and the MEC j_ν (in direction $\hat{r}$ at cosmological redshift z and rest-frame frequency $\nu = \nu_{\mathrm{c}}\,(1+z)$) is

$$I_\nu\left(\hat{r}, \nu_{\mathrm{c}}\right) = \int_0^\infty \frac{j_\nu\left(\hat{r}, z\left(l\right), \nu_{\mathrm{c}}\left(1+z\left(l\right)\right)\right)}{\left(1+z\left(l\right)\right)^3}\,\mathrm{d}l \approx \frac{j_\nu\left(\nu\right)\Delta l\left(\hat{r}\right)}{\left(1+z\right)^3}, \tag{4.6}$$

where l represents proper length. The approximation is valid for a lobe with a spatially constant MEC that is small enough to assume a constant redshift for it. $\Delta l(\hat{r})$ is the proper length of the line of sight through the lobe in direction $\hat{r}$. The inferred MECs $j_{\nu,1}\left(\nu\right)$ and $j_{\nu,2}\left(\nu\right)$ thus correspond to the rest-frame frequency.

4.A3.3 Proposal distribution

In order to explore the posterior distribution on the parameter space, we followed the Metropolis algorithm. The Metropolis algorithm assumes a symmetric proposal distribution.

Radio galaxy axis direction

To propose a new RG axis direction given the current one whilst satisfying the symmetry assumption, we performed a trick. We populated the unit sphere with $N \in \mathbb{N}_{\geq 1}$ points (interpreted as directions) drawn from a uniform distribution. Of these N directions, the proposed axis direction was taken to be the one closest to the current axis direction (in the great-circle distance sense). We note that this approach evidently satisfies the criterion that proposing the new direction given the old one is equally likely as proposing the old direction given the new one. We also note that the

distribution of the angular distance between current and proposed axis directions is determined solely by N.

In the following paragraphs, we first review how to perform uniform sampling of the unit two-sphere. More explicitly than in Scott & Tout (1989), we then derive the distribution of the angular distance between a reference point and the nearest of N uniformly drawn other points. The result is a continuous univariate distribution with a single parameter N and finite support $(0, \pi)$. Finally, we present the mode, median and maximum likelihood estimator of N. As far as we know, these properties are new to the literature.

Uniform sampling of $\mathbb{S}^2$ Let us place a number of points uniformly on the celestial sphere $\mathbb{S}^2$. The spherical coordinates of such points are given by the RVs (Φ, Θ), where Φ denotes position angle and Θ denotes inclination angle. As all position angles are equally likely, the distribution of Φ is uniform: $\Phi \sim \mathbb{U}[0, 2\pi)$. In order to affect a uniform number density, the probability that a point lies within a rectangle of width $\mathrm{d}\phi$ and height $\mathrm{d}\theta$ in the (ϕ, θ) plane equals the ratio of the solid angle of the corresponding sky patch and the sphere's total solid angle:

$$\mathbb{P}(\phi \leq \Phi < \phi + \mathrm{d}\phi, \theta \leq \Theta < \theta + \mathrm{d}\theta) = \frac{\sin\theta \, \mathrm{d}\phi \, \mathrm{d}\theta}{4\pi}. \tag{4.7}$$

The probability that the inclination angle is found somewhere in the interval $[\theta, \theta + \mathrm{d}\theta)$, regardless of the position angle, is therefore

$$\begin{aligned} \mathbb{P}(\theta \leq \Theta < \theta + \mathrm{d}\theta) &= \mathrm{d}F_\Theta(\theta) = f_\Theta(\theta)\mathrm{d}\theta \\ &= \int_0^{2\pi} \frac{\sin\theta \, \mathrm{d}\theta}{4\pi} \mathrm{d}\phi = \frac{1}{2} \sin\theta \, \mathrm{d}\theta, \end{aligned} \tag{4.8}$$

where F_Θ is the cumulative distribution function (CDF) of Θ, and f_Θ the associated PDF. So,

$$f_\Theta(\theta) = \frac{1}{2} \sin\theta; \;\; F_\Theta(\theta) := \int_0^\theta f_\Theta(\theta') \, \mathrm{d}\theta' = \frac{1 - \cos\theta}{2}. \tag{4.9}$$

Nearest-neighbour angular distance distribution Let us pick a reference point and stochastically introduce N other points in above fashion, which we dub its 'neighbours'. We now derive the PDF of the angular distance to the nearest neighbour (NNAD). Let $(\phi_{\mathrm{ref}}, \theta_{\mathrm{ref}})$ be the coordinates of the reference point and let (ϕ, θ) be the coordinates of one of the neighbours. Without loss of generality, due

to spherical symmetry, we can choose to place the reference point in the direction towards the observer: $\theta_{\text{ref}} = 0$. (We note that ϕ_{ref} is meaningless in this case.) The angular distance between two points on $\mathbb{S}^2$ is given by the great-circle distance ξ. For our choice of reference point, we immediately see that $\xi(\phi_{\text{ref}}, \theta_{\text{ref}}, \phi, \theta) = \theta$. Because θ is a realisation of Θ, ξ too can be regarded as a realisation of an RV, which we call Ξ. Evidently, the PDF $f_\Xi(\xi) = f_\Theta(\xi)$ and the CDF $F_\Xi(\xi) = F_\Theta(\xi)$.

Now we consider the generation of N points, whose angular distances to the reference point are the RVs $\{\Xi_i\} := \{\Xi_1, \ldots, \Xi_N\}$. The NNAD RV M is the minimum of this set: $M := \min\{\Xi_i\}$. We next determine the CDF F_M and PDF f_M of M by noting that

$$\begin{aligned} F_M(\mu) := \mathbb{P}(M \leq \mu) &= \mathbb{P}(\text{minimum of } \{\Xi_i\} \leq \mu) \\ &= \mathbb{P}(\text{at least one of the set } \{\Xi_i\} \leq \mu) \\ &= 1 - \mathbb{P}(\text{none of the set } \{\Xi_i\} \leq \mu) \\ &= 1 - \mathbb{P}(\text{all of the set } \{\Xi_i\} > \mu). \end{aligned} \tag{4.10}$$

Because the $\{\Xi_i\}$ are independent and identically distributed,

$$\begin{aligned} F_M(\mu) &= 1 - \prod_{i=1}^{N} \mathbb{P}\left(\Xi_i > \mu\right) \\ &= 1 - \mathbb{P}^N(\Xi > \mu) = 1 - (1 - F_\Xi(\mu))^N. \end{aligned} \tag{4.11}$$

By substitution, the application of a trigonometric identity and differentiation to μ, we obtain the CDF and PDF of M:

$$F_M\left(\mu\right) = 1 - \cos^{2N}\left(\frac{\mu}{2}\right); \; f_M(\mu) = N \sin\left(\frac{\mu}{2}\right) \cos^{2N-1}\left(\frac{\mu}{2}\right). \tag{4.12}$$

In Fig. 4.13, we show this PDF for various values of N.
The mode of M (i.e. the most probable NNAD), μ_{mode}, is the solution to $\frac{\mathrm{d}f_M}{\mathrm{d}\mu}(\mu_{\text{mode}}) = 0$. The median of M, μ_{median}, is the solution to $F_M(\mu_{\text{median}}) = \frac{1}{2}$. Hence,

$$\mu_{\text{mode}} = \arccos\left(1 - \frac{1}{N}\right); \; \mu_{\text{median}} = \arccos\left(2^{1-\frac{1}{N}} - 1\right). \tag{4.13}$$

As common sense dictates, both equal $\frac{\pi}{2}$ for $N = 1$ and tend to 0 as $N \to \infty$. We

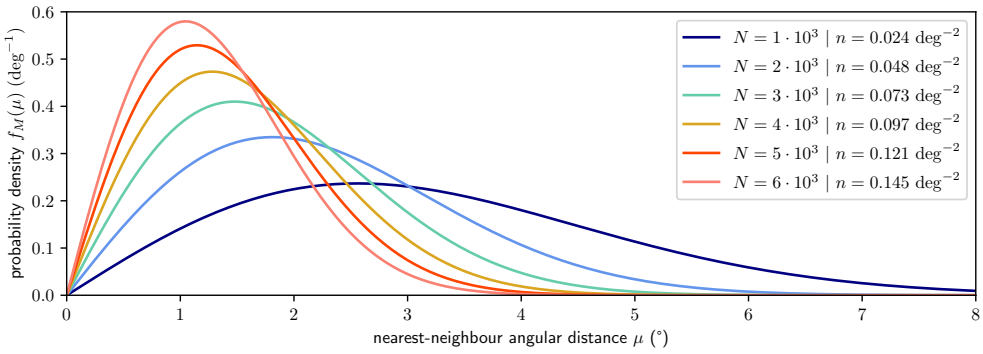


Figure 4.13: PDFs of the NNAD RV M between some fixed point and N other points distributed randomly over the celestial sphere. As the sphere gets more densely packed, the probability of finding a small M increases. For each N, we provide the equivalent mean point number density, n.

find the mean of M through integration by parts:

$$\begin{aligned}\mathbb{E}[M] &:= \int_0^\pi \mu f_M(\mu)\,\mathrm{d}\mu = \int_0^\pi \mu\,\mathrm{d}F_M(\mu) \\ &= \Big[\mu F_M(\mu)\Big]_0^\pi - \int_0^\pi F_M(\mu)\,\mathrm{d}\mu \\ &= \int_0^\pi \cos^{2N}\left(\frac{\mu}{2}\right)\mathrm{d}\mu = 2\int_0^{\frac{\pi}{2}} \cos^{2N}(\mu)\,\mathrm{d}\mu. \end{aligned} \tag{4.14}$$

Again via integration by parts,

$$\mathbb{E}[M] = \pi \prod_{k=1}^{N} \frac{2k-1}{2k} = \frac{\pi}{2^{2N}}\binom{2N}{N}. \tag{4.15}$$

Maximum likelihood estimation A typical application is the estimation of N in the PDF $f_M(\mu \mid N)$ (see Eq. 4.12) using data. We assume we have measured k NNADs, denoted by $\{\mu_1, \ldots, \mu_k\}$. Let the joint PDF or likelihood be

$$\begin{aligned}\mathcal{L}(N) &:= \prod_{i=1}^{k} f_M(\mu_i \mid N) \\ &= \left(\frac{N}{2^N}\right)^k \prod_{i=1}^{k} \sin\mu_i\,(\cos\mu_i + 1)^{N-1}. \end{aligned} \tag{4.16}$$

To find N_{MLE}, we look for the value of N that maximises $\mathcal{L}(N)$. To simplify the algebra, we could however equally well maximise a k-th of the natural logarithm of the likelihood, or the average log-likelihood $\hat{l} := k^{-1} \ln \mathcal{L}(N)$, because the logarithm is a monotonically increasing function:

$$\begin{aligned}\hat{l}(N) &:= \frac{1}{k} \ln \mathcal{L}(N) = \ln N - N \ln 2 \\ &\quad + \frac{1}{k} \sum_{i=1}^{k} \ln \sin \mu_i + (N-1) \ln(\cos \mu_i + 1). \end{aligned} \tag{4.17}$$

We find N_{MLE} by solving $\frac{\mathrm{d}\hat{l}}{\mathrm{d}N}(N_{\mathrm{MLE}}) = 0$. This leads to

$$N_{\mathrm{MLE}} = \left(\ln 2 - \frac{1}{k} \sum_{i=1}^{k} \ln(\cos \mu_i + 1) \right)^{-1}. \tag{4.18}$$

An easy limit to evaluate is the case when $\mu_1, \ldots, \mu_k \to 0$. In such a case, $\cos \mu_i \to 1$, and so $\frac{1}{k} \sum_{i=1}^{k} \ln(\cos \mu_i + 1) \to \ln 2$. Then, $N_{\mathrm{MLE}} \to (0_+)^{-1} \to \infty$. This is expected behaviour: when all measured NNADs approach 0, the number of points distributed on the sphere must be approaching infinity.

Other parameters

The other proposal parameters were each drawn from independent normal distributions centred around the current parameter values. These proposal distributions are evidently symmetric but have support over the full real line, so forbidden parameter values could in principle be proposed. As a remedy, we set the prior probability density of the proposed parameter set to 0 when the proposed opening angle was negative or exceeded $\frac{\pi}{2}$ rad, at least one of the proposed MECs was negative, or when at least one of the proposed inner truncation distances was negative or exceeded the corresponding proposed outer truncation distance. In such cases, the posterior probability density was 0 too, as it is proportional to the prior probability density. Consequently, the Metropolis acceptance probability vanished and the proposal was rejected. We did not enter forbidden regions of parameter space. The condition of detailed balance was still respected: probability densities for transitioning *towards* the forbidden region were 0, just as probability densities for being in the forbidden region.

4.A3.4 Likelihood

Assuming thermal noise to be the dominant noise type, we took the likelihood to be Gaussian. To avoid dimensionality errors, we multiplied the likelihood by a constant before we applied the logarithm:

$$\ln\left(\mathcal{L}\cdot\left(\sigma\sqrt{2\pi}\right)^{N_\mathrm{r}}\right) = -\frac{N_\mathrm{r}}{2\sigma^2 N_\mathrm{p}}\sum_{i=1}^{N_\mathrm{p}}\left(I_{\nu,\mathrm{o}}\left[i\right]-I_{\nu,\mathrm{m}}\left[i\right]\right)^2. \tag{4.19}$$

Here, σ is the image noise, $N_\mathrm{r}\in\mathbb{R}_{\geq 0}$ is the number of resolution elements (i.e. PSF solid angles) in the image, $N_\mathrm{p}\in\mathbb{N}$ is the number of pixels in the image, and $I_{\nu,\mathrm{o}}\left[i\right]$ and $I_{\nu,\mathrm{m}}\left[i\right]$ are the i-th pixel values of the observed and modelled image, respectively. For simplicity, one may multiply the likelihood by a constant factor (or, equivalently, add a constant term to the log-likelihood): the acceptance ratio will remain the same, and the MH MCMC runs correctly.

Table 4.2: Maximum a posteriori probability (MAP) estimates and posterior mean and standard deviation (SD) of the parameters from the Bayesian, doubly truncated, conical RG lobe model of Sect. 4.3.9.

parameter	MAP estimate	posterior mean and SD
ϕ_1	307°	$307\pm1^\circ$
ϕ_2	140°	$139\pm2^\circ$
$\lvert\theta_1-90^\circ\rvert$	54°	$51\pm2^\circ$
$\lvert\theta_2-90^\circ\rvert$	25°	$18\pm7^\circ$
γ_1	9°	$10\pm1^\circ$
γ_2	24°	$26\pm2^\circ$
$d_{\mathrm{i},1}$	2.7 Mpc	2.6 ± 0.2 Mpc
$d_{\mathrm{o},1}$	4.3 Mpc	4.0 ± 0.2 Mpc
$d_{\mathrm{i},2}$	1.6 Mpc	1.5 ± 0.1 Mpc
$d_{\mathrm{o},2}$	2.0 Mpc	2.0 ± 0.1 Mpc
$j_{\nu,1}(\nu)$	17 Jy deg^{-2} Mpc^{-1}	17 ± 2 Jy deg^{-2} Mpc^{-1}
$j_{\nu,2}(\nu)$	22 Jy deg^{-2} Mpc^{-1}	18 ± 3 Jy deg^{-2} Mpc^{-1}

4.A3.5 Results for Alcyoneus

We applied the Bayesian model to the 90″ LoTSS DR2 image of Alcyoneus, shown in the top panel of Fig. 4.9. Thus, the hyper-parameters were $z=0.24674$, $\nu_\mathrm{c}=144$ MHz (so that $\nu=180$ MHz), $\theta_\mathrm{FWHM}=90''$, $N=750$ and $\sigma=\sqrt{2}\cdot 1.16$ Jy deg^{-2}. We set the image noise to $\sqrt{2}$ times the *true* image noise to account

for model incompleteness. This factor follows by assuming that the inability of the model to produce the true lobe morphology yields (Gaussian) errors comparable to the image noise. To speed up inference, we down-sampled the image of 2,048 by 2,048 pixels by a factor of 16 along each dimension. We ran our MH MCMC for 10,000 steps and discarded the first 1,500 steps due to burn-in. Table 4.2 lists the obtained maximum a posteriori probability estimates and posterior mean and standard deviation of the parameters.

Table 4.3: Maximum a posteriori probability (MAP) estimates and posterior mean and standard deviation (SD) of derived quantities from the Bayesian, doubly truncated, conical RG lobe model of Sect. 4.3.9.

derived quantity	MAP estimate	posterior mean and SD
$\Delta\phi$	167°	$168 \pm 2^\circ$
V_1	$1.5\ \mathrm{Mpc}^3$	$1.5 \pm 0.2\ \mathrm{Mpc}^3$
V_2	$0.8\ \mathrm{Mpc}^3$	$1.0 \pm 0.2\ \mathrm{Mpc}^3$
$F_{\nu,1}(\nu_\mathrm{c})$	63 mJy	63 ± 4 mJy
$F_{\nu,2}(\nu_\mathrm{c})$	44 mJy	45 ± 5 mJy
$P_{\mathrm{min},1}$	$4.7 \cdot 10^{-16}$ Pa	$4.8 \pm 0.3 \cdot 10^{-16}$ Pa
$P_{\mathrm{min},2}$	$5.4 \cdot 10^{-16}$ Pa	$5.0 \pm 0.6 \cdot 10^{-16}$ Pa
$P_{\mathrm{eq},1}$	$4.8 \cdot 10^{-16}$ Pa	$4.9 \pm 0.3 \cdot 10^{-16}$ Pa
$P_{\mathrm{eq},2}$	$5.4 \cdot 10^{-16}$ Pa	$5.0 \pm 0.6 \cdot 10^{-16}$ Pa
$B_{\mathrm{min},1}$	45 pT	45 ± 1 pT
$B_{\mathrm{min},2}$	48 pT	46 ± 3 pT
$B_{\mathrm{eq},1}$	42 pT	43 ± 1 pT
$B_{\mathrm{eq},2}$	45 pT	43 ± 3 pT
$E_{\mathrm{min},1}$	$6.3 \cdot 10^{52}$ J	$6.2 \pm 0.4 \cdot 10^{52}$ J
$E_{\mathrm{min},2}$	$3.7 \cdot 10^{52}$ J	$4.4 \pm 0.6 \cdot 10^{52}$ J
$E_{\mathrm{eq},1}$	$6.4 \cdot 10^{52}$ J	$6.3 \pm 0.4 \cdot 10^{52}$ J
$E_{\mathrm{eq},2}$	$3.8 \cdot 10^{52}$ J	$4.4 \pm 0.6 \cdot 10^{52}$ J

The proper volumes V_1 and V_2 are derived quantities:

$$V = \frac{\pi}{3} \tan^2 \gamma \left(d_\mathrm{o}^3 - d_\mathrm{i}^3\right), \tag{4.20}$$

just like the flux densities $F_{\nu,1}(\nu_\mathrm{c})$ and $F_{\nu,2}(\nu_\mathrm{c})$ at central frequency ν_c:

$$F_\nu(\nu_\mathrm{c}) = \frac{j_\nu(\nu)\, V}{(1+z)^3\, r_\varphi^2(z)}. \tag{4.21}$$

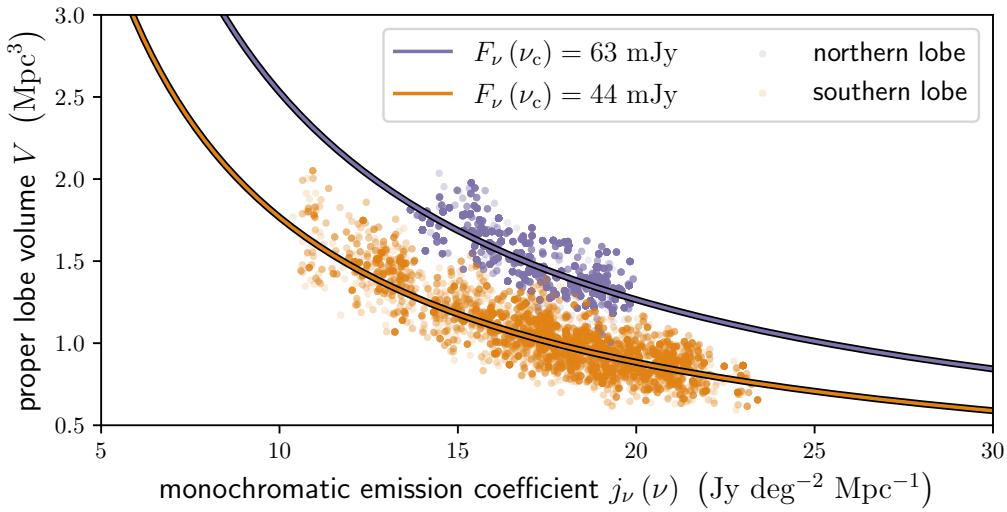


Figure 4.14: Strongly correlated estimates of $j_\nu(\nu)$ and V from our Bayesian model, demonstrating consistency with the observed lobe flux densities. We show MECs $j_\nu(\nu)$ at $\nu = 180$ MHz and proper volumes V of MH MCMC samples for the northern lobe (purple dots) and southern lobe (orange dots). The curves represent all combinations $(j_\nu(\nu), V)$ that correspond to a particular flux density at the LoTSS central wavelength $\nu_c = 144$ MHz. We show the observed northern lobe flux density (purple curve) and the observed southern lobe flux density (orange curve).

Together, V and $F_\nu(\nu_c)$ imply a lobe pressure P and a magnetic field strength B, which are additional derived quantities that we calculate through `pysynch`. Table 5.5 lists the obtained MAP estimates and posterior mean and standard deviation of the derived quantities.

The uncertainties of the parameters and derived quantities reported in Tables 4.2 and 5.5 are not necessarily independent. To demonstrate this, we present MECs and volumes from the MH MCMC samples in Fig. 4.14. MECs and volumes do not vary independently, because their product is proportional to flux density (see Eq. 4.21); only realistic flux densities correspond to high-likelihood model images.

Finally, we explored a simpler variation of the model, in which we forced the lobes to be coaxial. In such a case, the true proper length l and projected proper length l_p are additional derived quantities:

$$l = \frac{d_{o,1} + d_{o,2}}{\cos\gamma}; \;\; l_p = l \sin\theta. \tag{4.22}$$

For Alcyoneus, this simpler model does not provide a good fit to the data.

Cosmology brings us face to face with the deepest mysteries, questions that were once treated only in religion and myth.

Carl E. Sagan, American astronomer, *Cosmos: A Personal Voyage (1980)*

5

An intergalactic medium temperature from a giant radio galaxy

M. S. S. L. Oei, R. J. van Weeren, M. J. Hardcastle, F. Vazza, T. W. Shimwell, F. Leclercq, M. Brüggen, H. J. A. Röttgering — *Monthly Notices of the Royal Astronomical Society*, 518, 240–256, 2023

Abstract

The warm–hot intergalactic medium (warm–hot IGM, or WHIM) pervades the filaments of the Cosmic Web and harbours half of the Universe's baryons. The WHIM's thermodynamic properties are notoriously hard to measure. Here we estimate a galaxy group–WHIM boundary temperature using a new method. In particular, we use a radio image of the giant radio galaxy (giant RG, or GRG) created by NGC 6185, a massive nearby spiral. We analyse this extraordinary object with a Bayesian 3D lobe model and deduce an equipartition pressure $P_{\mathrm{eq}} = 6 \cdot 10^{-16}$ Pa — among the lowest found in RGs yet. Using an X-ray-based statistical conversion for Fanaroff–Riley II RGs, we find a true lobe pressure $P = 1.5^{+1.7}_{-0.4} \cdot 10^{-15}$ Pa. Cosmic Web reconstructions, group catalogues, and MHD simulations furthermore imply an Mpc–scale IGM density $1 + \delta_{\mathrm{IGM}} = 40^{+30}_{-10}$. The buoyantly rising lobes are crushed by the IGM at their inner side, where an approximate balance between IGM and lobe pressure occurs: $P_{\mathrm{IGM}} \approx P$. The ideal gas law then suggests an IGM temperature $T_{\mathrm{IGM}} = 11^{+12}_{-5} \cdot 10^{6}$ K, or $k_{\mathrm{B}} T_{\mathrm{IGM}} = 0.9^{+1.0}_{-0.4}$ keV, at the virial radius — consistent with X-ray-derived tem-

peratures of similarly massive groups. Interestingly, the method is not performing at its limit: in principle, estimates $T_{\rm IGM} \sim 4 \cdot 10^6$ K are already possible — rivalling the lowest X-ray measurements available. The technique's future scope extends from galaxy group outskirts to the WHIM. In conclusion, we demonstrate that observations of GRGs in Cosmic Web filaments are finally sensitive enough to probe the thermodynamics of galaxy groups and beyond.

Key words: methods: statistical – galaxies: active – intergalactic medium – large-scale structure of Universe – radio continuum: galaxies

5.1 Introduction

Although the warm–hot intergalactic medium (WHIM) in the filaments of the Cosmic Web is the main baryon reservoir of the modern Universe, it has proven challenging to determine its physical properties from observations. A handful of techniques have already been successful, ranging from direct X-ray imaging (Eckert et al., 2015), X-ray spectroscopy of blazars in search of O VII absorption by intervening filaments (Nicastro et al., 2018), X-ray image stacking (Tanimura et al., 2020; Vernstrom et al., 2021; Tanimura et al., 2022), microwave image stacking of galaxy pairs targeting the thermal Sunyaev–Zel'dovich effect (Tanimura et al., 2019b; de Graaff et al., 2019), to dispersion measurements of localised fast radio bursts (Macquart et al., 2020). It has long been speculated that giant radio galaxies (GRGs, or colloquially *giants*), of which thousands are now known (Oei et al., 2023a), could serve as yet another probe of the WHIM. (GRGs are radio galaxies whose proper length component in the plane of the sky $l_{\rm p}$ exceeds 0.7 or 1 Mpc, depending on convention.) Both observations and modelling indicate that the pressure in GRG lobes tends to decrease strongly as giants grow (e.g. Oei et al., 2022a) and should, especially when jet feeding halts, approach that of the encompassing intergalactic medium (IGM). Close to pressure equilibrium, the IGM provides a significant resisting force that shapes the dynamics and morphology of the lobes. Therefore, by observing GRGs in filaments of the Cosmic Web, one could indirectly study WHIM thermodynamics (e.g. Subrahmanyan et al., 2008; Malarecki et al., 2013).

In this work, we present record-low pressure measurements of the lobes of NGC 6185, a GRG in the nearby Cosmic Web. The GRG is near enough that Cosmic Web reconstructions, which enable IGM density estimates, are available. This in turn allows us to infer the IGM temperature at the virial radius of NGC 6185's group and thus, for the first time, provide strong constraints on thermodynamics in filaments from radio galaxy observations.

In Section 5.2, we introduce the data used in this work. In Section 5.3, we present methods and results, leading up to our IGM temperature estimate. In Section 5.4, we discuss caveats and potential future extensions of our work, right before Section 5.5's concluding remarks.

We assume a concordance inflationary ΛCDM model with parameters as in Jasche et al. (2015): $h = 0.702$, $\Omega_{\mathrm{M},0} = 0.272$ and $\Omega_{\Lambda,0} = 0.728$; $H_0 =: h \cdot 100$ km s^{-1} Mpc^{-1}. We define spectral indices α such that power-law spectra are of the form $L_\nu \propto \nu^\alpha$. In our terminology, a radio galaxy is distinct from the galaxy that has produced it, and only consists of relativistic plasma, magnetic fields, and radiation.

5.2 Data

5.2.1 NGC 6185 and its giant radio galaxy

In this work, we characterise the Cosmic Web environment of NGC 6185 and its GRG. NGC 6185 is a spiral galaxy at a spectroscopic redshift $z = 0.0343 \pm 0.0002$ (Falco et al., 1999). At a comoving distance of 146 Mpc and a luminosity distance of 151 Mpc, the galaxy lies in the nearby Cosmic Web. It is of Hubble–de Vaucouleurs class SAa (Jansen et al., 2000). Its stellar mass, $M_\star = 3.0^{+1.2}_{-0.9} \cdot 10^{11}\, M_\odot$ (Kannappan et al., 2013), is high for a spiral galaxy, though common for galaxies hosting GRGs (Oei et al., 2022a).[1] Using the stellar velocity dispersion $\sigma_v = 236$ km s^{-1} from Kannappan et al. (2013), for which we assume a 10% error, and the M–sigma relation of Eq. 7 in Kormendy & Ho (2013), we obtain a super-massive black hole (SMBH) mass of $M_\bullet = 6^{+4}_{-2} \cdot 10^8\, M_\odot$. Again, although high for spiral galaxies,[2] such an SMBH mass is common for galaxies hosting GRGs (e.g. Dabhade et al., 2020a; Oei et al., 2022a). In particular, it is similar to the SMBH mass of J2345−0449, the projectively largest known spiral galaxy–hosted GRG ($l_\mathrm{p} = 1.6$ Mpc) before the discovery of the GRG of NGC 6185: $M_\bullet = 10^8$–$10^9\, M_\odot$ (Bagchi et al., 2014). We show a close-up of the galaxy in Fig. 5.1. A major fraction of the gas in the galaxy appears dynamically disrupted and separated from the disk at distances of $\sim 10^1$ kpc.

The GRG of NGC 6185 has been discovered by Oei et al. (2023a) using the Low-Frequency Array (LOFAR; van Haarlem et al., 2013). More specifically, the GRG appeared in Data Release 2 (DR2) of the LOFAR Two-metre Sky Survey (LoTSS; Shimwell et al., 2017, 2022), its Northern Sky imaging survey at central observing

[1] As commented by Kannappan et al. (2013), this stellar mass estimate appears robust against variations in model assumptions. Indeed, using a different prescription, Kannappan et al. (2009) provide an almost identical estimate.

[2] For comparison, the mass of the SMBH in the centre of the Milky Way is $M_\bullet = 4 \cdot 10^6\, M_\odot$; the SMBH in the centre of NGC 6185 is thus roughly 150 times more massive.

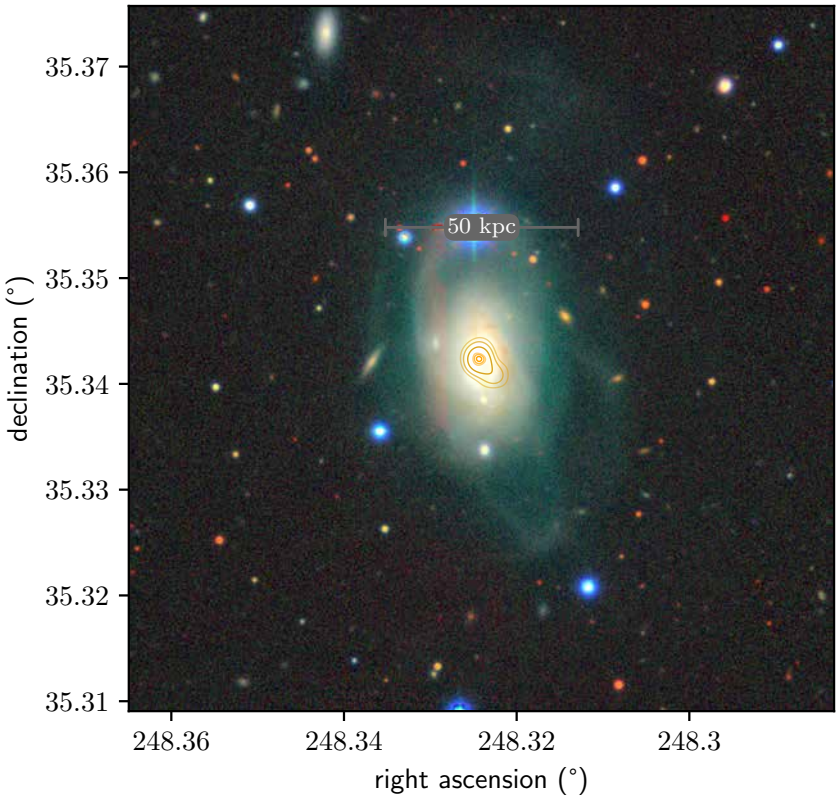


Figure 5.1: Optical close-up of NGC 6185, a spiral galaxy which has generated the GRG shown in Fig. 5.2. On top of the $4' \times 4'$ DESI Legacy Imaging Surveys (Dey et al., 2019) DR9 (g, r, z) image, we show LoTSS DR2 $6''$ contours (yellow) and VLASS $2.2''$ contours (orange) at 50, 100, and 200σ, where $\sigma_{\mathrm{LoTSS}} = 5 \cdot 10^1$ Jy deg^{-2} and $\sigma_{\mathrm{VLASS}} = 2 \cdot 10^2$ Jy deg^{-2}.

frequency $\nu_{\mathrm{obs}} = 144$ MHz and resolutions of $6''$, $20''$, and $60''$. The GRG, shown in Fig. 5.2, consists of a core and two extended lobes of smooth morphology. The GRG is a Fanaroff–Riley II (FRII) radio galaxy. In total, it has a $1.0°$ angular length and a projected proper length $l_{\mathrm{p}} = 2.45 \pm 0.01$ Mpc. At least in an angular sense, NGC 6185 is located symmetrically between the lobes. A chance alignment is improbable, given that galaxies with redshifts as low as NGC 6185's are distributed sparsely over the sky. In fact, the GRG must belong to NGC 6185, as it is the only low-redshift galaxy in the sky patch between the two lobes. If instead it were to belong to a galaxy at even a moderate redshift of $z = 0.1$ or $z = 0.2$, we would find $l_{\mathrm{p}} = 6.6$ Mpc or $l_{\mathrm{p}} = 11.9$ Mpc. However, these projected proper lengths are several megaparsecs larger than that of Alcyoneus, with $l_{\mathrm{p}} = 5.0$ Mpc the projectively longest known GRG (Oei et al., 2022a). GRGs of this extent are rare. Assuming that the GRG projected proper length distribution extends beyond $l_{\mathrm{p}} = 5$ Mpc as a power law

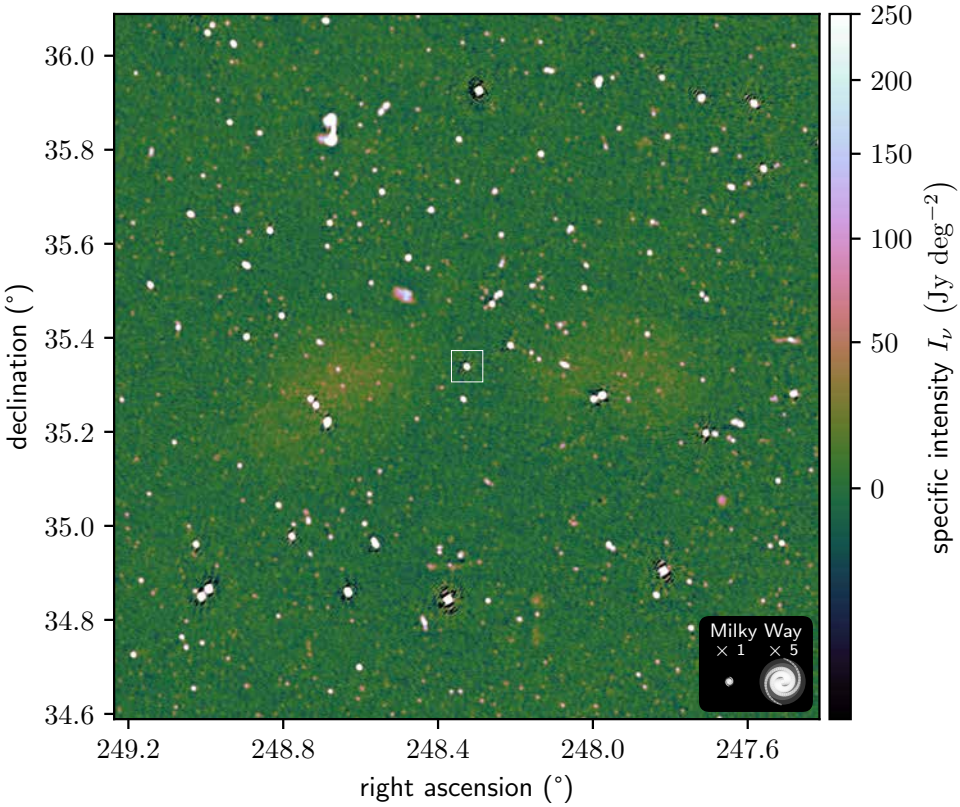


Figure 5.2: Radio view of NGC 6185, the lobes of its GRG, and the surrounding sky, at $\nu_{\rm obs} = 144$ MHz. We show a LoTSS DR2 20″ image spanning 1.5° × 1.5°. The degree-long GRG has a 2.5 Mpc projected proper length and is the largest known specimen with a spiral galaxy host. The white box in the centre marks the region shown in Fig. 5.1.

with exponent $\xi = -3.5$ (Oei et al., 2023a), GRGs with $l_{\rm p} = 6.6$ Mpc and $l_{\rm p} = 11.9$ Mpc would be three and twenty times rarer still. As a final argument, at $\nu_{\rm obs} = 144$ MHz, the centre of NGC 6185 appears radio-bright; the contours of Fig. 5.1 illustrate that the specific intensity rises to hundreds of times the local LoTSS DR2 root mean square noise $\sigma_{\rm LoTSS} = 5 \cdot 10^1$ Jy deg^{-2}. A higher 2.2″ resolution Very Large Array Sky Survey (VLASS; Lacy et al., 2020) image at $\nu_{\rm obs} = 3$ GHz reveals that the majority of this emission is from a region with a diameter of at most 1.5 kpc (which corresponds to the VLASS FWHM at the redshift of NGC 6185) around the galactic centre. This indicates the presence of an active galactic nucleus (AGN) — or, alternatively, a starburst nucleus.[3]

To gain a better understanding of the AGN candidate in NGC 6185, we investigate

[3] A radio image from very-long baseline interferometry (VLBI), which can now be made from LOFAR observations at $\nu_{\rm obs} = 144$ MHz, would resolve this matter.

its radio spectrum. We retrieve flux densities from the LoTSS DR2 (Shimwell et al., 2022), the WENSS (Rengelink et al., 1997), the NVSS (Condon et al., 1998), the Arecibo 2380 MHz Survey of Bright Galaxies (Dressel & Condon, 1978), the VLASS (Gordon et al., 2021), and a VLA follow-up of extragalactic IRAS sources (Condon et al., 1995). We list these literature data in Table 5.1.

Table 5.1: Literature radio flux densities of the AGN candidate in NGC 6185.

frequency ν_{obs} (MHz)	flux density F_ν (mJy)	telescope and literature reference
144	117 ± 12	LOFAR; Shimwell et al. (2022)
325	101	WSRT; Rengelink et al. (1997)
1400	61.4 ± 1.9	VLA; Condon et al. (1998)
2380	41 ± 3	Arecibo; Dressel & Condon (1978)
3000	34.0 ± 0.2	VLA; Gordon et al. (2021)
4860	26	VLA; Condon et al. (1995)

We perform Metropolis–Hastings Markov chain Monte Carlo (MCMC) in order to infer the underlying radio spectrum from the data. We convert flux densities at observing frequencies ν_{obs} to luminosity densities at rest-frame frequencies $\nu = \nu_{\text{obs}}\,(1+z)$, and assume that the AGN's luminosity density in the radio obeys

$$L_\nu\,(\nu) = L_\nu\,(\nu_{\text{ref}}) \cdot \left(\frac{\nu}{\nu_{\text{ref}}}\right)^{\alpha(\nu)} \;;\quad \alpha\,(\nu) = \alpha\,(\nu_{\text{ref}}) + \beta \ln \frac{\nu}{\nu_{\text{ref}}}. \tag{5.1}$$

This model describes a parabola in log–log space. The model's three parameters are $L_\nu\,(\nu_{\text{ref}})$, $\alpha\,(\nu_{\text{ref}})$, and β; ν_{ref} is a constant that determines their meaning. We assume a flat prior over the model parameters, a Gaussian likelihood, and 10% flux density errors when the literature does not provide them. We choose $\nu_{\text{ref}} := 150$ MHz, run the MCMC, and obtain the parameter estimates shown in Table 5.2.

Table 5.2: Maximum a posteriori probability (MAP) and posterior mean and standard deviation (SD) estimates of the parameters from the Bayesian radio spectrum model. We choose $\nu_{\text{ref}} := 150$ MHz.

parameter	MAP	posterior mean and SD
$L_\nu\,(\nu_{\text{ref}})$	$3.0 \cdot 10^{23}$ W Hz^{-1}	$3.0 \pm 0.3 \cdot 10^{23}$ W Hz^{-1}
$\alpha\,(\nu_{\text{ref}})$	0.06	0.07 ± 0.09
β	-0.15	-0.15 ± 0.02

We visualise the data alongside the posterior in Fig. 5.3. Massaro et al. (2014) have

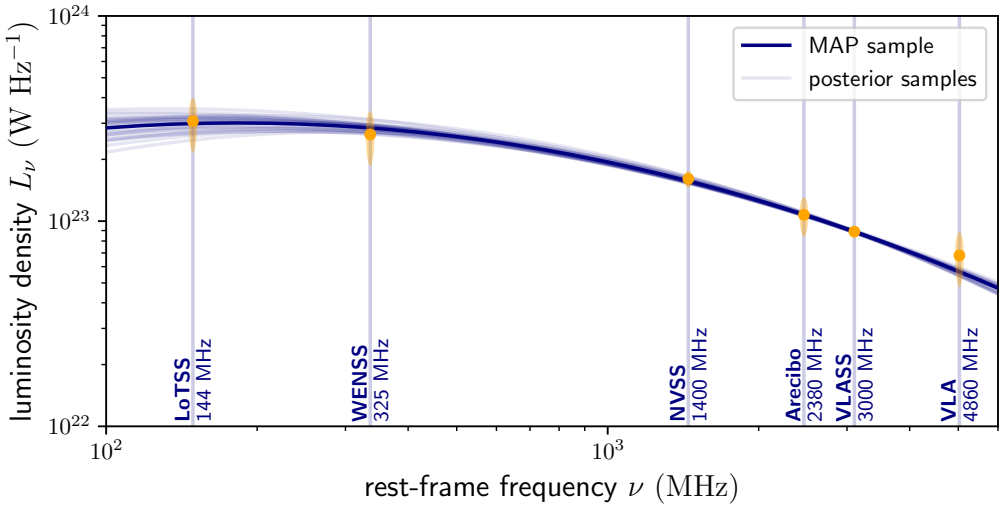


Figure 5.3: Rest-frame radio spectrum of the AGN in NGC 6185. We show measured luminosity densities with 3σ uncertainties (orange) alongside random post–burn-in posterior samples (light blue) and the MAP sample (dark blue). At $\sim 10^2$ MHz frequencies, the spectrum is flat. We denote the surveys used with their *observing* frequencies.

noted the AGN's relatively flat spectrum before — through the WENSS–NVSS spectral index (i.e. between 325 MHz and 1400 MHz), which is $\langle \alpha \rangle = -0.34 \pm 0.04$. Our analysis shows that the spectrum becomes even flatter at lower frequencies, with $\alpha\,(\nu_{\rm ref} = 150\ {\rm MHz}) = 0.07 \pm 0.09$, implying a physically compact emitting structure in which synchrotron self-absorption takes place. This, in turn, strongly suggests that there is a currently active jet. The galactic centre is an ultraluminous X-ray source (ULX): the *Chandra* X-ray Observatory has measured a maximum 0.3–8 keV luminosity $L_{\rm X} = 9.7 \cdot 10^{40}$ erg s^{-1}, at 4σ significance (Wang et al., 2016; Evans et al., 2020).

5.2.2 Cosmic Web late-time total matter density field

Oei et al. (in preparation) have used the Bayesian Origin Reconstruction from Galaxies (BORG; Jasche & Wandelt, 2013) SDSS to measure the large-scale density (Jasche et al., 2015) and dynamical state (Leclercq et al., 2015) of the Cosmic Web around hundreds of GRGs. The BORG SDSS offers a probability distribution, represented by an MCMC, over the possible density fields of the low-redshift ($z < 0.17$) Universe populated by galaxies from the SDSS DR7 Main Galaxy Sample (Abazajian et al., 2009). Each MCMC sample covers the same comoving volume of $(750\ {\rm Mpc}\ h^{-1})^3$ with a 256^3-voxel box. Thus, the side length of a BORG SDSS voxel is $\frac{1}{256} \cdot 750\ {\rm Mpc}\ h^{-1} \approx 4.2$ Mpc. At this resolution, one can consider the baryonic and dark matter den-

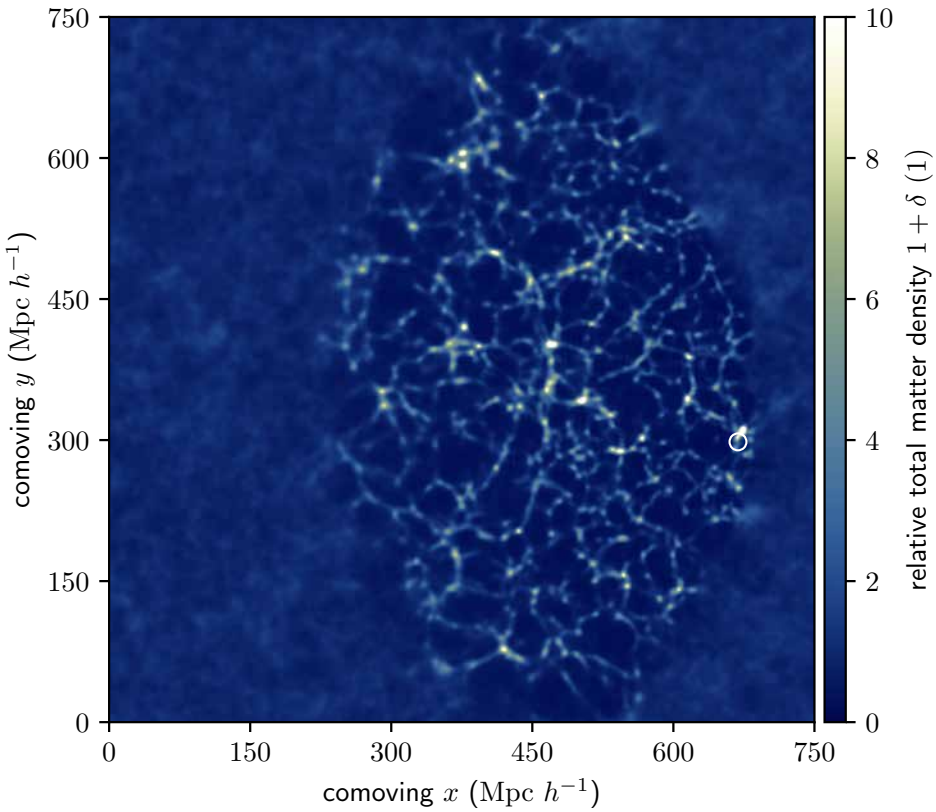


Figure 5.4: Localisation of NGC 6185 within the large-scale structure of the Universe. We show a slice of constant Cartesian comoving z through the late-time BORG SDSS posterior mean total matter density field. The slice covers a square with comoving area 750 Mpc $h^{-1} \cdot 750$ Mpc h^{-1} and is 3 Mpc h^{-1} thick. Outside of the SDSS DR7–constrained volume, the posterior mean tends to the Universe's late-time mean total matter density $\bar{\rho}_0$. The location of the GRG is marked by a white circle.

sity fields as approximately identical; the BORG SDSS does not distinguish between them. Each MCMC sample provides a different total matter density at a given voxel, and so it is the set of all MCMC samples that provides a marginal distribution for the total matter density at the voxel. In Fig. 5.4, we show the mean of these marginal distributions for all voxels in a slice that contains NGC 6185. The slice reveals the location of the galaxy within the Cosmic Web. The total matter density averaged over a $(4.2\ \mathrm{Mpc})^3$ volume is $1+\delta = 2.3 \pm 0.7$. Furthermore, if the Cosmic Web is classified on the basis of its gravitational dynamics in the T-web sense (Hahn et al., 2007), one finds a 99% probability that NGC 6185 resides in a filament. (Galaxy group regions are a part of filaments under the T-web classification at the 4 Mpc–scale.) We show NGC 6185 within a 3D BORG SDSS visualisation in Fig. 5.5. The centre of

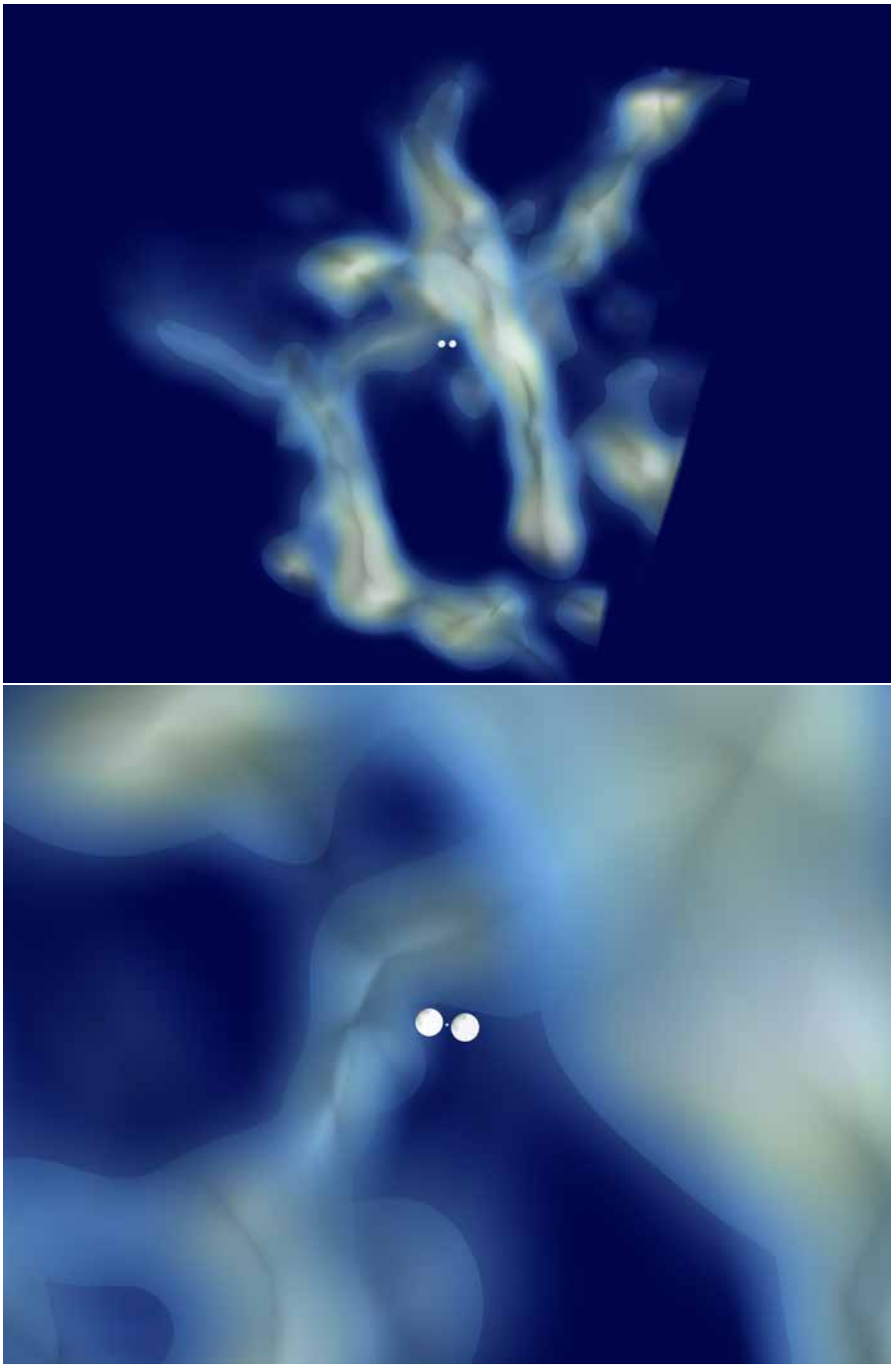


Figure 5.5: Three-dimensional view of NGC 6185 in its Cosmic Web environment. We show a box with 63 Mpc sides (15 BORG SDSS voxels along each side). The lobes of NGC 6185 are to scale. We show an isodensity surface at relative density $1 + \delta = 4$. *Top:* observer's view, with north pointing up and east to the left. At a distance of 15–20 Mpc, the massive galaxy cluster Abell 2199 looms near. *Bottom:* rotated close-up.

the galaxy cluster Abell 2199 occurs at a distance of 15–20 Mpc.

Whereas the lobes of NGC 6185's GRG lie mostly in the WHIM, the host galaxy itself resides in a galaxy group. In particular, the 2MASS galaxy group catalogue (Tully, 2015) suggests that NGC 6185 resides in a group with a virial mass $M = 2.6 \pm 0.5 \cdot 10^{13}\, M_\odot$.[4] Saulder et al. (2016) also report the presence of a group, but place the total mass at $M = 9^{+11}_{-5} \cdot 10^{12}\, M_\odot$. Finally, Tempel et al. (2017) estimate a Navarro–Frenk–White (NFW; Navarro et al., 1996) profile–based mass $M_{200} = 7 \cdot 10^{12}\, M_\odot$, where $R_{200} = 0.4$ Mpc. Assuming that the entire group falls within the same 4 Mpc–scale voxel, one can calculate a lower bound to the voxel's $1 + \delta$, effected by the group mass alone. These lower bounds are $1 + \delta \gtrsim 10.4$, $1 + \delta \gtrsim 3.6$ and $1 + \delta \gtrsim 2.8$, respectively. All lower bounds exceed the BORG SDSS measurement $1 + \delta = 2.3 \pm 0.7$. The mass that has in reality collapsed into the group will not, or barely, have done so in the BORG SDSS, which lacks redshift-space distortion modelling and whose gravity solver and galaxy bias model accuracy are limited.

Given the large discrepancies between these group mass estimates, we perform additional analysis ourselves. For all SDSS DR7–detected galaxies with spectroscopic redshifts, we calculate the proper distance d to NGC 6185 assuming no peculiar motion. In Table 5.3, we list all for which $d < 5$ Mpc; we consider these galaxies to be possible members of a group dominated by NGC 6185.[5] For each galaxy, we collect a stellar mass from Kannappan et al. (2013), Chang et al. (2015), or Mamon et al. (2020). The sum of stellar masses of galaxies within $d < 1$ Mpc, $d < 3$ Mpc, and $d < 5$ Mpc (including NGC 6185 itself) are $M_\star = 3.1 \cdot 10^{11}\, M_\odot$, $M_\star = 6.6 \cdot 10^{11}\, M_\odot$, and $M_\star = 8.4 \cdot 10^{11}\, M_\odot$, respectively. (At this redshift, SDSS DR7 incompleteness is unimportant.) Various studies have quantified the relationship between stellar mass and total mass (Lovisari et al., 2021). Using the IllustrisTNG (e.g. Marinacci et al., 2018; Naiman et al., 2018; Nelson et al., 2018; Springel et al., 2018) relationship of Pillepich et al. (2018) for the stellar masses given above, we find $M_{500} = 1 \cdot 10^{13}\, M_\odot$, $M_{500} = 2 \cdot 10^{13}\, M_\odot$, and $M_{500} = 3 \cdot 10^{13} M_\odot$, respectively.

[4] The value given in the catalogue is different, as it is based on a spurious luminosity distance estimate. We report a K_s luminosity–based virial group mass recalculated through Eq. 7 of Tully (2015), using the correct luminosity distance. We adopt the suggested 20% uncertainty.

[5] Because peculiar motion induces an Mpc-scale error on d, we list all selected galaxies with $d < 5$ Mpc — even though actual groups have radii less than 1 Mpc.

Table 5.3: Properties of SDSS DR7–detected galaxies with spectroscopic redshifts z near NGC 6185. Besides labels and coordinates, we provide the proper distance d to NGC 6185 (assuming no peculiar motion), the probability that the galaxy is a spiral p_s, r-band luminosities L_r, and the stellar mass $M_\star$. We order galaxies on the basis of d, which we compute from coordinates. We take STARLIGHT-based $M_\star$ from Mamon et al. (2020) whenever available, and from Chang et al. (2015) otherwise — except in the case of NGC 6185, for which we follow Kannappan et al. (2013). The other data are from the galaxy group catalogue by Tempel et al. (2017).

rank $d \downarrow$	name SDSS DR16	galaxy group ID	right ascension J2000 (°)	declination J2000 (°)	spectroscopic z heliocentric (1)	d (Mpc)	p_s (%)	L_r ($10^9\,L_\odot$)	$M_\star$ ($10^9\,M_\odot$)
0	NGC 6185	2919	248.32436	35.34235	0.03436 ± 0.00011	0	100	131.5	295.1
1	SDSS J163317.73+352001.5	2919	248.32389	35.33376	0.03454 ± 0.00002	0.7	33	1.9	7.9
2	SDSS J163214.44+351448.7	-	248.06020	35.24688	0.03417 ± 0.00001	1.0	99	3.4	6.2
3	SDSS J163305.64+350600.9	-	248.27352	35.10028	0.03467 ± 0.00001	1.4	98	4.3	2.1
4	SDSS J163528.10+355013.1	-	248.86709	35.83699	0.03414 ± 0.00001	1.9	99	26.3	83.2
5	SDSS J163242.46+352515.2	2919	248.17695	35.42091	0.03481 ± 0.00001	1.9	45	2.0	4.4
6	SDSS J162916.51+352456.5	-	247.31881	35.41570	0.03402 ± 0.00001	2.4	99	7.1	100.0
7	SDSS J163309.59+345534.7	67621	248.28996	34.92631	0.03491 ± 0.00001	2.5	41	24.0	154.9
8	SDSS J163727.41+355604.9	5413	249.36425	35.93472	0.03445 ± 0.00001	2.6	100	1.4	1.1
9	SDSS J163513.80+361318.5	87499	248.80750	36.22182	0.03404 ± 0.00001	2.7	98	2.0	0.9
10	SDSS J163607.24+360900.1	87499	249.03020	36.15004	0.03385 ± 0.00001	3.2	99	3.1	3.8
11	SDSS J163320.66+344825.8	67621	248.33609	34.80717	0.03511 ± 0.00001	3.3	95	5.7	2.7
12	SDSS J162636.40+350242.1	-	246.65167	35.04504	0.03416 ± 0.00002	3.5	96	15.0	57.5
13	SDSS J163322.14+352223.2	2919	248.34227	35.37313	0.03340 ± 0.00002	3.9	99	3.8	12.3
14	SDSS J164041.11+355947.1	-	250.17132	35.99643	0.03419 ± 0.00001	4.1	29	9.4	29.5
15	SDSS J162510.65+351106.7	-	246.29439	35.18521	0.03414 ± 0.00001	4.2	99	2.8	0.8
16	SDSS J163451.06+364506.2	-	248.71278	36.75173	0.03375 ± 0.00001	4.3	98	1.6	0.4
17	SDSS J162441.30+345001.6	-	246.17212	34.83380	0.03400 ± 0.00001	4.7	99	8.7	33.1
18	SDSS J163222.58+343905.0	-	248.09411	34.65141	0.03543 ± 0.00002	4.7	99	1.9	1.3
19	SDSS J163359.47+342308.2	82190	248.49780	34.38562	0.03538 ± 0.00003	4.8	99	6.2	15.5
20	SDSS J163308.49+343759.0	-	248.28539	34.63306	0.03546 ± 0.00001	4.8	99	7.3	22.9
21	SDSS J163307.82+344752.4	67621	248.28259	34.79789	0.03552 ± 0.00001	4.9	99	5.4	1.5

NGC 6185's group environment can also be characterised by counting galaxies. At the redshift of NGC 6185, the cosmic mean proper number density of SDSS DR7–detected galaxies with spectroscopic redshifts is $1.6 \cdot 10^{-2}\,\mathrm{Mpc}^{-3}$. Thus, within spheres of proper radii 2, 3, 4, and 5 Mpc, one expects to find 0.5, 1.8, 4.2, and 8.2 such galaxies, respectively. However, we find 5, 9, 13, and 21 such galaxies (other than NGC 6185) within spheres of said radii centred around NGC 6185. The galaxy number density around NGC 6185 is thus a factor of order unity higher than the cosmic mean at its redshift: $1 + \delta_{\mathrm{gal}} = 3$–10, depending on the averaging scale.

For our final estimate for $1 + \delta$, which we will use throughout the remainder of this work, we treat the BORG SDSS measurement as a background density upon which a group of mass $M = 1 \cdot 10^{13}\,M_{\odot}$ has formed. We adopt a 30% group mass uncertainty; under this assumption, the Saulder et al. (2016) and Tempel et al. (2017) estimates occur within 1 standard deviation. This yields $1 + \delta = 6 \pm 2$.

5.2.3 Cosmological simulation

In order to obtain a statistical conversion relation between total matter density at the 4 Mpc–scale and IGM density at the 1 Mpc–scale, we turn to cosmological simulations. In particular, we use a snapshot of one of the largest uniform-grid magnetohydrodynamics (MHD) simulations to date (Vazza et al., 2019), conducted with the Enzo code (Bryan et al., 2014). The simulation covers a comoving volume of $(100\,\mathrm{Mpc})^3$ with a 2400^3-voxel box. We use the baryonic and dark matter density fields ρ_{BM} and ρ_{DM} at the snapshot for $z = 0.025$, close to the GRG's redshift of $z = 0.034$. Thus, the side length of an Enzo simulation voxel is $\frac{1}{2400} \cdot 100\,\mathrm{Mpc} \approx 42\,\mathrm{kpc}$. Along each dimension, Enzo simulation voxels are 100 times smaller than BORG SDSS voxels. Still, the simulations do not feature galactic physics: chemistry, star formation, radiative cooling, and AGN feedback are all absent. This fact may limit the accuracy of our IGM density determinations on Mpc^3-scale around the simulation's galactic halos; see Section 5.4.4 for a discussion.

5.3 Methods and results

5.3.1 Lobe pressures

We infer the pressure in the lobes of NGC 6185's GRG by fitting a simple Bayesian lobe model to LoTSS DR2 imagery. As Fig. 5.2 shows, the GRG is a degree long, and consequently its lobes directionally coincide with several physically unrelated background sources of substantial radio flux density. To remove contamination from these sources, we predict 6″ LoTSS DR2 sky model visibilities and subtract them

from calibrated data (van Weeren et al., 2021), as we have done with the 20″ LoTSS DR2 sky model in Oei et al. (2022a). In order to avoid subtracting the signal of interest, we verify that the 6″ LoTSS DR2 sky model (in contrast to its 20″ counterpart) does not contain any lobe emission. As before, we then perform multiscale CLEAN deconvolution (Offringa & Smirnov, 2017) with Briggs −0.5 weighting. Using WSClean IDG (Offringa et al., 2014; van der Tol et al., 2018) version 2.10.1, we arrive at an image of 90″ resolution. The source subtraction is not perfect, and as a result some artefacts from unrelated compact sources remain. We remove these by assigning all pixels whose value deviates more than three image noise standard deviations from the local median this latter value. We finally apply to the image a LoTSS DR2 flux density scale correction factor (Hardcastle et al., 2021; Shimwell et al., 2022) of 0.985, based on the Sixth Cambridge Survey of Radio Sources (6C; Hales et al., 1988, 1990) and the NVSS. The final image appears in the top panel of Fig. 5.6.

Next we extend the Bayesian radio galaxy lobe model developed in Oei et al. (2022a) to infer lobe volumes from this image. This model parametrises a pair of lobes in three dimensions with some geometric shape. We choose an appropriate shape simply by inspecting the radio image; for Alcyoneus, we chose truncated cones, whilst for the GRG of NGC 6185, spheroids appear appropriate.[6] (In this particular case, the lobes appear to be well modelled by prolate spheroids. However, the model allows for oblate spheroids too.) In an initial model formulation attempt, we forced each lobe's axis of revolution to pierce through the host galaxy. This constraint encapsulates the idea that the lobes originate from the host. However, we found that the resulting model cannot provide an accurate fit to the data — especially to that of the eastern lobe. We therefore slightly modify the constraint, by still forcing the axes of revolution to pierce through a common point, but by allowing this point to be offset from the currently observed host galaxy position. From a physical perspective, this generalised constraint still captures the fact that the lobes share a common origin, but also allows for the possibility of relative motion of the host galaxy with respect to the lobes during their formation. Such relative motion can cause measurable displacements: a relative speed of $\sim 10^2$ km s^{-1} maintained over a $\sim 10^0$ Gyr period shifts the host galaxy's position by $\sim 10^{-1}$ Mpc. At NGC 6185's distance, this corresponds to an angular shift of $\sim 10^0$ arcmin. Due to the two-dimensional nature of our data, we can only recover the offset in the plane of the sky. We thus describe the offset by means of a two-dimensional vector pointing towards the currently observed host galaxy position, parametrised through a position angle ϕ_0 and a projected proper

[6] Future versions of the model should automatically select an appropriate shape. This can be done by first performing inference for each shape, and then performing model selection — for example through Bayes factors.

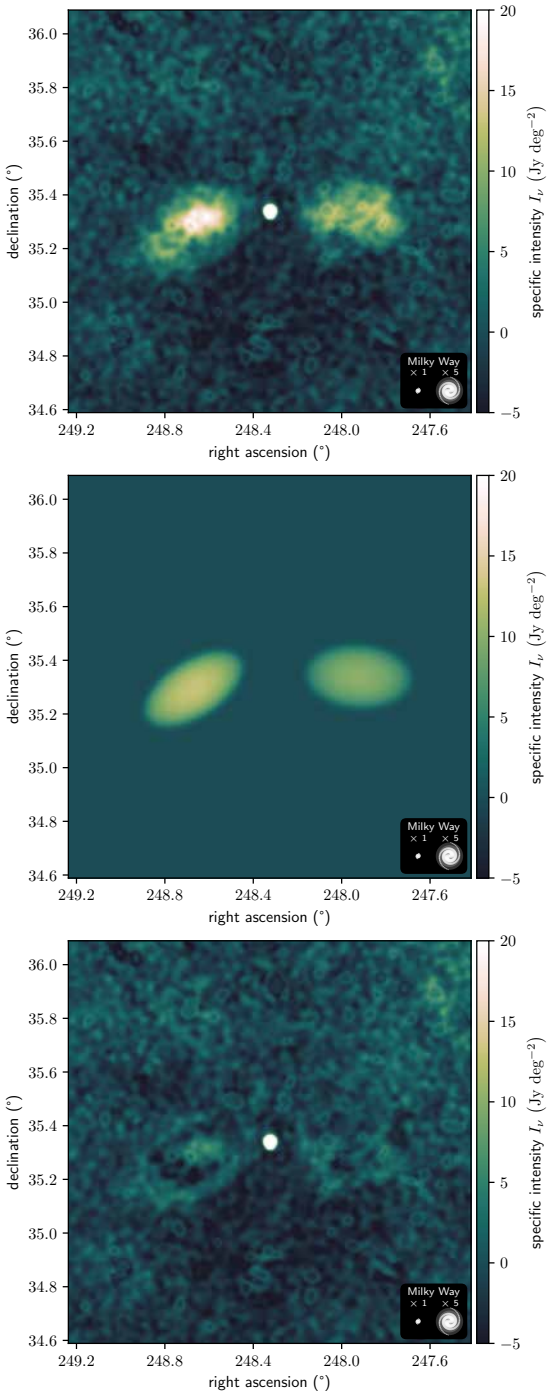


Figure 5.6: Overview of our Bayesian radio galaxy lobe model, which allows inference of physical properties by matching an observed image to modelled images, applied to the lobes of NGC 6185's giant. *Top:* LoTSS DR2 compact source–subtracted 90″ image. *Middle:* MAP model image. *Bottom:* residual image.

length d_0. We parametrise each spheroid not only by the direction of its axis of revolution, which we capture with another position angle ϕ and an inclination angle θ, but also by the distance d of the spheroid's centre from the common origin, the semi-axis a along the axis of revolution, and the semi-axis b perpendicular to this axis.[7] Within each lobe, we assume a constant monochromatic emission coefficient (MEC; Rybicki & Lightman, 1986) j_ν. Given (ϕ_0, d_0), and $(\phi, \theta, d, a, b, j_\nu)$ for each lobe, we generate a MEC field on a voxel grid centered around the host galaxy. The total MEC field is thus fully described by a 14-dimensional parameter vector $\mathbf{p}$. From this 3D MEC field we generate the 2D model radio image by integrating along the line of sight and applying a $(1+z)^{-3}$ cosmological attenuation factor.

To find the posterior, we must calculate the likelihood that the observed image is the modelled image distorted by thermal noise, which we assume to be Gaussian. Importantly, we also make use of a non-flat prior. With a flat prior, there exists a degeneracy between low-MEC lobes with a large extent along the third (i.e. line-of-sight) dimension, and high-MEC lobes with a small extent along this third dimension: such scenarios produce similar images. To break this degeneracy, we make use of the fact that observations indicate that the intrinsic lengths of GRGs are approximately Pareto distributed with tail index $\xi = -3.5 \pm 0.5$ (Oei et al., 2023a). By enforcing a prior on intrinsic length, we elegantly favour smaller 3D configurations that produce a match to the data over larger 3D configurations that accomplish the same. For our modelled radio galaxies, we define the intrinsic length l as the 3D distance between the eastern (E) and western (W) lobe tips. We let $\hat{r}(\phi, \theta)$ denote the unit vector pointing in the direction given by position angle–inclination angle pair (ϕ, θ). Then the prior $\mathcal{P}(\mathbf{p})$ (up to an immaterial constant) becomes

$$\mathcal{P}(\mathbf{p}) \propto l(\mathbf{p})^{\xi}, \tag{5.2}$$

where l is

$$l(\mathbf{p}) := ||(d_{\mathrm{E}} + a_{\mathrm{E}})\hat{r}(\phi_{\mathrm{E}}, \theta_{\mathrm{E}}) - (d_{\mathrm{W}} + a_{\mathrm{W}})\hat{r}(\phi_{\mathrm{W}}, \theta_{\mathrm{W}})||_2. \tag{5.3}$$

We repeat model image generation and likelihood and prior calculation many times for different parameter values. More precisely, we perform Metropolis–Hastings MCMC to explore the posterior distribution; we refer the reader to Oei et al. (2022a) for more details on the algorithm.[8] Numerically, we run 10 independent Markov chains

[7] Both d as well as the semi-axes a and b are proper, not comoving, lengths.

[8] A sensible model extension is to incorporate the additional constraint that d_0 may not be too large. The two velocity components of the host galaxy in the plane of the sky may be approximated through independent Gaussian (e.g. Yahil & Vidal, 1977; Ribeiro et al., 2013) random variables with zero mean

of 10^5 iterations each, where we tune the proposal parameters such that the acceptance rate is 23%, close to the predicted best rate (23.4%) from optimal scaling theory (e.g. Bédard, 2008). From each chain, we discard the first 10^4 samples to avoid burn-in effects, and aggregate the samples of the remaining $9 \cdot 10^5$ iterations.

We illustrate the Bayesian lobe model in Fig. 5.6. From top to bottom, we show the compact source–subtracted 90″ LoTSS DR2 image of NGC 6185's giant, the maximum a posteriori probability (MAP) model image, and the residual image after subtracting the observed image from the modelled one. For the western lobe, the residuals reveal no evidence for model inadequacy; for the eastern lobe, the residuals suggest that our constant-MEC spheroid model is a rough approximation only. We caution that the inferences for the two lobes are therefore not equally reliable. The eastern lobe may not be perfectly spheroidal, or the MEC may be locally enhanced — for example due to an inhomogeneous magnetic field. Detailed observations of nearby radio galaxies such as Fornax A (Maccagni et al., 2020) indeed show that MECs need not be constant within lobes. In Table 5.4, we present MAP and posterior mean and standard deviation (SD) estimates of the model parameters.

Table 5.4: MAP and posterior mean and SD of the parameters from the Bayesian spheroidal RG lobe model. Estimates for the western (W) lobe are more reliable than those for the eastern (E) lobe.

parameter	MAP	posterior mean and SD
d_0	0.40 Mpc	0.37 ± 0.06 Mpc
ϕ_0	257°	$250 \pm 11°$
ϕ_E	122°	$121 \pm 2°$
ϕ_W	266°	$264 \pm 3°$
$\lvert\theta_E - 90°\rvert$	22°	$22 \pm 14°$
$\lvert\theta_W - 90°\rvert$	25°	$15 \pm 10°$
d_E	0.4 Mpc	0.5 ± 0.1 Mpc
d_W	1.3 Mpc	1.2 ± 0.1 Mpc
a_E	0.55 Mpc	0.57 ± 0.06 Mpc
a_W	0.53 Mpc	0.53 ± 0.03 Mpc
b_E	0.27 Mpc	0.26 ± 0.01 Mpc
b_W	0.30 Mpc	0.30 ± 0.01 Mpc
$j_{\nu,E}(\nu)$	25 Jy deg^{-2} Mpc^{-1}	26 ± 3 Jy deg^{-2} Mpc^{-1}
$j_{\nu,W}(\nu)$	18 Jy deg^{-2} Mpc^{-1}	18 ± 1 Jy deg^{-2} Mpc^{-1}

and identical variance σ_v^2. Under this assumption, the total speed in the plane of the sky is Rayleigh distributed. Thus, a time Δt after lobe formation, the proper displacement of the host galaxy in the plane of the sky $d_0 \sim \text{Rayleigh}(\sigma_d)$, where $\sigma_d := \sigma_v \cdot \Delta t$. The prior then becomes $\mathcal{P}(\mathbf{p}) \propto l(\mathbf{p})^{\xi} f_{d_0}(\mathbf{p})$, where f_{d_0} is the PDF of d_0. The drawback of this approach is that one must somewhat arbitrarily choose the hyperparameter σ_d, which is typically unknown; values $\sigma_d \sim 10^{-1}$ Mpc appear justified.

The inferences $d_0 = 0.37 \pm 0.06$ Mpc and $\phi_0 = 250 \pm 11^\circ$ indicate that NGC 6185 may have been moving in southwestern direction with a speed $\sim 10^2$ km s^{-1} maintained over a $\sim 10^0$ Gyr period; however, we stress that this claim is tentative at best. The inferred inclination angles suggest the data are consistent with a moderate deviation from a sky plane geometry — $|\theta_\mathrm{E} - 90^\circ| = 22 \pm 14^\circ$ and $|\theta_\mathrm{W} - 90^\circ| = 15 \pm 10^\circ$ — although the latter is not ruled out given the uncertainties.

We use the post–burn-in samples to calculate derived quantities of interest. One of them is the position angle difference $\Delta\phi := \phi_\mathrm{W} - \phi_\mathrm{E}$, expected to be close to 180° in the most dilute Cosmic Web environments. Others are the proper distances in the plane of the sky between the host galaxy and the inner and outer tips of the (eastern) lobe; these are

$$d_{\mathrm{i,E}} = || - d_0 \hat{r}(\phi_0, 0) + P_\perp (d_\mathrm{E} - a_\mathrm{E}) \hat{r}(\phi_\mathrm{E}, \theta_\mathrm{E}) ||_2; \tag{5.4}$$

$$d_{\mathrm{o,E}} = || - d_0 \hat{r}(\phi_0, 0) + P_\perp (d_\mathrm{E} + a_\mathrm{E}) \hat{r}(\phi_\mathrm{E}, \theta_\mathrm{E}) ||_2, \tag{5.5}$$

with analogous expressions for the western lobe. Here, $P_\perp$ is a 3×3 matrix that projects vectors onto the plane of the sky.[9] Yet another is the intrinsic (3D) proper length l, measured from outer lobe tip to outer lobe tip, as given by Eq. 5.3. As the lobes are spheroidal, their proper volumes $V = \frac{4}{3}\pi a b^2$. The flux densities F_ν at $\nu_\mathrm{obs} = 144$ MHz relate to the parameters and the angular diameter distance to the galaxy as described by Eq. C.16 of Oei et al. (2022a); we also provide the corresponding luminosity densities L_ν at rest-frame frequency $\nu = \nu_\mathrm{obs}(1+z) = 149$ MHz. The minimum energy (Burbidge, 1956) and equipartition (Pacholczyk, 1970) pressure P, magnetic field strength B, and internal energy U of each lobe follow from the galaxy's redshift, the proper lobe volume, and the lobe flux density. As in Ineson et al. (2017), we assume that the electron energy distribution is a power law in Lorentz factor γ between $\gamma_\mathrm{min} = 10$ and $\gamma_\mathrm{max} = 10^5$ with exponent $p = -2.4$. We also assume a proton kinetic energy density vanishingly small compared to that of electrons ($\kappa = 0$), as suggested acceptable for FRII radio galaxies by the results of Ineson et al. (2017), and a maximal plasma filling factor ($\varphi = 1$), in line with the constant-MEC assumption of our model. We perform the calculations with `pysynch` (Hardcastle et al., 1998b), which implements the magnetic field estimation approach

[9] In case of a choice of basis in which the third basis vector is parallel to the line of sight,

$$P_\perp = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{5.6}$$

of Myers & Spangler (1985).[10] In Table 5.5, we present MAP and posterior mean and SD estimates of the derived quantities.

Table 5.5: MAP and posterior mean and SD of derived quantities from the Bayesian spheroidal RG lobe model. Estimates for the western (W) lobe are more reliable than those for the eastern (E) lobe.

derived quantity	MAP	posterior mean and SD
$\Delta\phi$	144°	$143 \pm 3^\circ$
$d_{\mathrm{i,E}}$	0.31 Mpc	0.31 ± 0.02 Mpc
$d_{\mathrm{i,W}}$	0.33 Mpc	0.32 ± 0.02 Mpc
$d_{\mathrm{o,E}}$	1.19 Mpc	1.19 ± 0.02 Mpc
$d_{\mathrm{o,W}}$	1.29 Mpc	1.30 ± 0.02 Mpc
l	2.7 Mpc	2.6 ± 0.1 Mpc
V_{E}	0.17 Mpc^3	0.16 ± 0.02 Mpc^3
V_{W}	0.20 Mpc^3	0.19 ± 0.02 Mpc^3
$F_{\nu,\mathrm{E}}(\nu_{\mathrm{obs}})$	640 mJy	630 ± 63 mJy
$F_{\nu,\mathrm{W}}(\nu_{\mathrm{obs}})$	520 mJy	530 ± 53 mJy
$L_{\nu,\mathrm{E}}(\nu)$	$1.7 \cdot 10^{24}$ W Hz^{-1}	$1.6 \pm 0.2 \cdot 10^{24}$ W Hz^{-1}
$L_{\nu,\mathrm{W}}(\nu)$	$1.4 \cdot 10^{24}$ W Hz^{-1}	$1.4 \pm 0.1 \cdot 10^{24}$ W Hz^{-1}
$P_{\mathrm{min,E}}$	$7.2 \cdot 10^{-16}$ Pa	$7.2 \pm 0.6 \cdot 10^{-16}$ Pa
$P_{\mathrm{min,W}}$	$6.0 \cdot 10^{-16}$ Pa	$6.1 \pm 0.4 \cdot 10^{-16}$ Pa
$P_{\mathrm{eq,E}}$	$7.2 \cdot 10^{-16}$ Pa	$7.2 \pm 0.6 \cdot 10^{-16}$ Pa
$P_{\mathrm{eq,W}}$	$6.0 \cdot 10^{-16}$ Pa	$6.1 \pm 0.4 \cdot 10^{-16}$ Pa
$B_{\mathrm{min,E}}$	50 pT	50 ± 2 pT
$B_{\mathrm{min,W}}$	46 pT	46 ± 2 pT
$B_{\mathrm{eq,E}}$	52 pT	52 ± 2 pT
$B_{\mathrm{eq,W}}$	47 pT	48 ± 2 pT
$U_{\mathrm{min,E}}$	$1.1 \cdot 10^{52}$ J	$1.0 \pm 0.1 \cdot 10^{52}$ J
$U_{\mathrm{min,W}}$	$1.0 \cdot 10^{52}$ J	$1.0 \pm 0.1 \cdot 10^{52}$ J
$U_{\mathrm{eq,E}}$	$1.1 \cdot 10^{52}$ J	$1.0 \pm 0.1 \cdot 10^{52}$ J
$U_{\mathrm{eq,W}}$	$1.0 \cdot 10^{52}$ J	$1.0 \pm 0.1 \cdot 10^{52}$ J

The GRG's total luminosity density at 150 MHz, combining core and lobes, is $L_\nu = 3.3 \pm 0.3 \cdot 10^{24}$ W Hz^{-1}. The lobe pressures, magnetic field strengths, and internal energies inferred from the minimum energy condition are statistically consistent with those inferred from the equipartition condition. Judging from Fig. 5.6, it is likely that our model somewhat underestimates the volume of the eastern lobe. For lobe pressures, we therefore adopt the western lobe estimate in the rest of this work. These

[10] The pysynch code is publicly available at `https://github.com/mhardcastle/pysynch`.

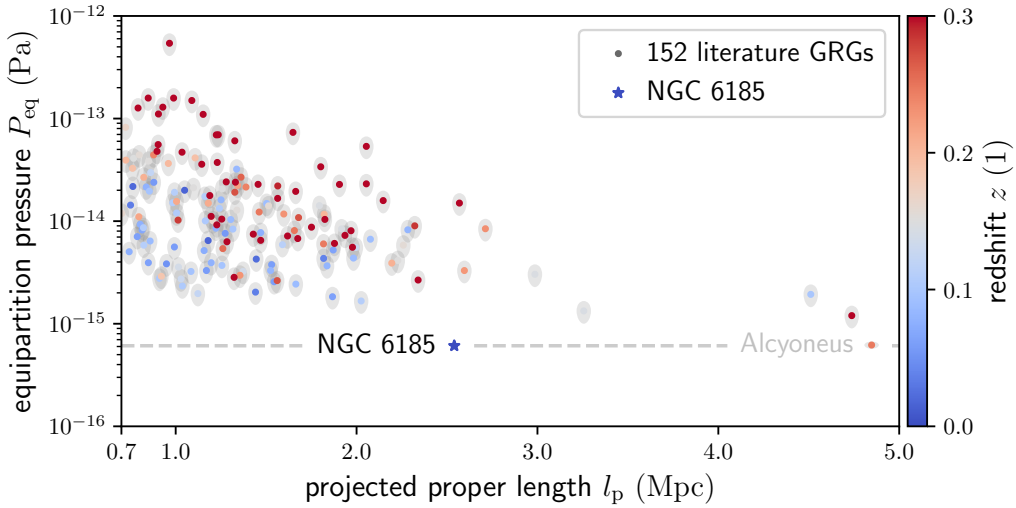


Figure 5.7: Relationship between total size and lobe equipartition pressure for observed giants, with colour denoting redshift. The equipartition pressures in the lobes of NGC 6185's giant are among the lowest measured yet.

pressures, $P_{\rm eq} \sim 6 \cdot 10^{-16}$ Pa, are among the lowest hitherto found in radio galaxy lobes.[11] In Fig. 5.7, we show the relation between projected proper length and lobe equipartition pressure as found by Oei et al. (2022a) for all known GRGs in non-cluster environments, appended with the new datum for NGC 6185's GRG. The fact that low-redshift GRGs in this diagram generally have lower pressures is likely a surface brightness selection effect; due to its $\propto (1+z)^{-3}$ scaling, the surface brightness of a lobe at $z = 0.3$ is already less than half (46%) of the surface brightness of the same lobe at $z = 0$. The record-low equipartition pressures presented here are measurable as a result of a combination of the depth of the LoTSS DR2, the GRG's large projected proper length, and its low redshift.

X-ray observations of inverse Compton scattering between relativistic lobe electrons and cosmic microwave background photons allow for a measurement of the true lobe pressure P. Following this approach, Ineson et al. (2017) have investigated the relation between true and equipartition lobe pressures for a representative sample of FRII radio galaxies. As shown in Fig. 5.8, for almost all studied cases the true pressure is higher than the equipartition pressure by a factor of order unity. Importantly, by plotting the ratio between P and $P_{\rm eq}$ as a function of lobe volume V, we find no clear trend over three orders of magnitude in lobe volume up to $V_{\rm E} \approx V_{\rm W} \approx 0.2$ Mpc^3. It thus appears reasonable to assume that the distribution of

[11] Upon recalculating $P_{\rm eq}$ for Alcyoneus (Oei et al., 2022a) under this work's cosmology and this section's assumptions, one finds $P_{\rm eq} \sim 6 \cdot 10^{-16}$ Pa, too.

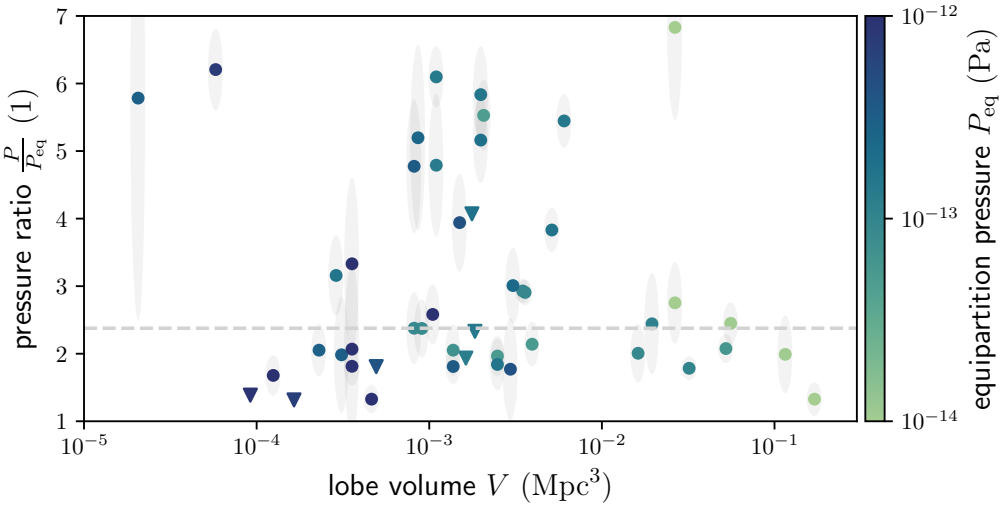


Figure 5.8: The ratio between the true lobe pressure P and the equipartition lobe pressure $P_{\rm eq}$, as a function of lobe volume V. These data, from Ineson et al. (2017), suggest that true pressures are a factor of order unity higher than equipartition pressures, with no clear trend in this factor over several orders of magnitude in V. Triangles symbolise upper bounds. The dashed line marks the median pressure ratio, $\frac{P}{P_{\rm eq}} = 2.4$.

$\frac{P}{P_{\rm eq}}$ is the same for GRGs and non-giant RGs. Furthermore, as expected, $P_{\rm eq}$ and V anticorrelate, but both do not appear to strongly constrain $\frac{P}{P_{\rm eq}}$. To obtain the true lobe pressure for NGC 6185's GRG, we thus resort to a statistical conversion based on the entire shown Ineson et al. (2017) sample. The conversion factor becomes $\frac{P}{P_{\rm eq}} = 2.4^{+2.8}_{-0.6}$. Applying it to NGC 6185's giant, we arrive at a true lobe pressure $P = 1.5^{+1.7}_{-0.4} \cdot 10^{-15}$ Pa.

5.3.2 IGM density

To find the temperature of the IGM surrounding NGC 6185, we must first obtain the density of the local IGM. In Section 5.2.2, we established a 4 Mpc–scale relative total matter density $1 + \delta = 6 \pm 2$ from the BORG SDSS and galaxy group catalogues. However, $1 + \delta$ cannot be considered a direct IGM density estimate for three reasons. Firstly, it combines baryonic and dark matter; secondly, it encompasses — for a significant part — matter that would have collapsed into galaxies and galactic halos in hypothetical higher-resolution reconstructions; and thirdly, it measures density on an inappropriately large scale. By contrast, we are interested in baryons only, and in particular in those occupying the rarefied space outside galaxies and their halos. Moreover, the BORG SDSS provides the density on a 4 Mpc–scale, much larger

than the typical diameter of a filament or cluster. If we could peer into the galaxy's voxel, we would see that a large part of it is void-like. As a result, on this large scale, the relative total matter density of clusters is $1 + \delta \sim 10$ instead of $1 + \delta \sim 10^2$–10^3; meanwhile, in filaments $1 + \delta \sim 1$ instead of $1 + \delta \sim 10$. The large averaging scale of the BORG SDSS density field thus biases high-density environments low and low-density environments high. Clearly, to obtain a reasonable estimate of the IGM density as experienced by the lobes of the GRG, we must use a smaller averaging scale. To obtain a feeling of the dependence of IGM density on the averaging scale, we present a simple analytic analysis in which we compare the density of a fixed piece of large-scale structure on a small averaging scale to the same quantity on a large averaging scale. We consider a Cosmic Web filament, geometrically modelled as a cylinder, whose IGM density around the central axis follows an isothermal β-model.[12] Originally, this model was proposed to describe intra-cluster medium density profiles (e.g. Cavaliere & Fusco-Femiano, 1976, 1978; Arnaud, 2009), but is nowadays also common as a WHIM density profile descriptor (e.g. Gheller & Vazza, 2019; Tuominen et al., 2021). The model is parametrised by a central density, a core radius r_c, and a slope parameter β. We obtain insightful analytic expressions if we consider the volumes over which we average the density to be cylindrical, with central axes that coalesce with the filament's. Let L be the BORG SDSS voxel side length, so that a BORG SDSS voxel has volume L^3. We assign the cylinder representing the large averaging volume a length L and radius R. We choose R such that the area of the cylindrical section perpendicular to the axis equals L^2, the area of a voxel face. Thus, $R := \frac{1}{\sqrt{\pi}} L \approx 2.4$ Mpc. Similarly, we assign the cylinder representing the small averaging volume a length l and radius r; analogously, we set $r := \frac{1}{\sqrt{\pi}} l$. For example, a voxel of side length $l = 1$ Mpc implies $r \approx 0.6$ Mpc. One can show that the ratio of average densities is

$$\frac{\langle \rho \rangle (r)}{\langle \rho \rangle (R)} = \left(\frac{R}{r} \right)^2 \cdot \frac{\left(1 + \left(\frac{r}{r_c} \right)^2 \right)^{1 - \frac{3}{2}\beta} - 1}{\left(1 + \left(\frac{R}{r_c} \right)^2 \right)^{1 - \frac{3}{2}\beta} - 1}, \tag{5.7}$$

[12] Although NGC 6185's small-scale environment might be a galaxy group, its large-scale environment still is a filament.

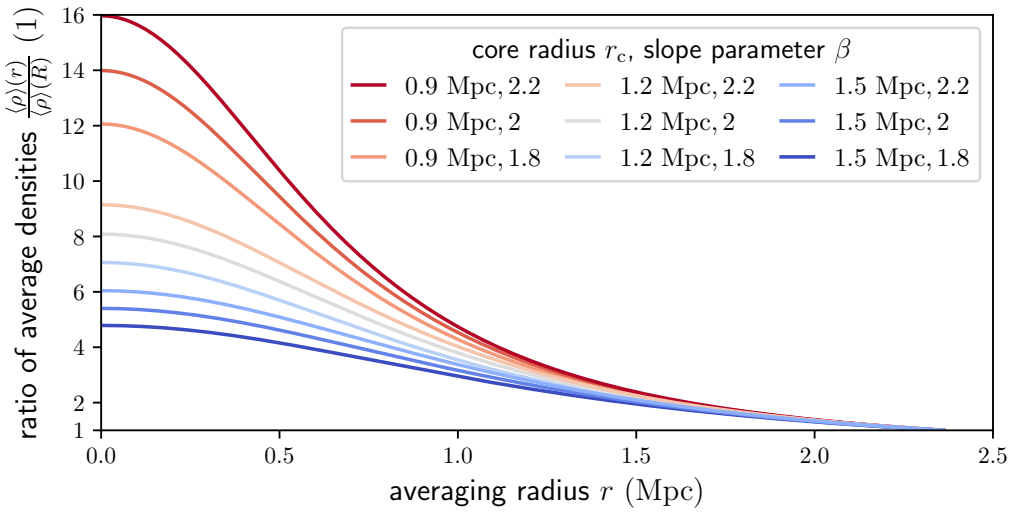


Figure 5.9: Ratio between small-scale density, averaged until radius r, and large-scale density, averaged until radius R, for a cylindrical filament whose WHIM density profile follows the isothermal β-model. We adopt values for the core radius r_c and slope parameter β suggested by Tuominen et al. (2021). The BORG SDSS fixes $R = 2.4$ Mpc. If an isothermal β-model describes the WHIM density profile around the filament spine, then the average WHIM density within radius r can be much larger than within radius R (if $r < R$).

except when $\beta = \frac{2}{3}$; in that case,

$$\frac{\langle\rho\rangle(r)}{\langle\rho\rangle(R)} = \left(\frac{R}{r}\right)^2 \cdot \frac{\ln\left(1 + \left(\frac{r}{r_c}\right)^2\right)}{\ln\left(1 + \left(\frac{R}{r_c}\right)^2\right)}. \tag{5.8}$$

See Appendix 5.A1 for a derivation and interesting limits. As we consider a ratio of average densities, the central density of the isothermal β-model drops out; as a result, for r variable and R fixed just two parameters remain. We visualise the ratio between the average WHIM density within radius r and that within radius R in Fig. 5.9. We use parameter values from the WHIM analysis by Tuominen et al. (2021) of the Evolution and Assembly of Galaxies and their Environments (EAGLE; Schaye et al., 2015; Crain et al., 2015) simulations. For $r = 0.6$ Mpc, which corresponds to a 1 Mpc^3 averaging volume around the filament spine, the WHIM density is 5–10 times higher than for $r = R$.

We invoke cosmological simulation snapshots to find a statistical conversion relation between the total matter density in a (4.2 Mpc)^3 cubical volume around massive galaxies and the IGM density in a 1 Mpc^3 spherical volume around them. Follow-

ing Gheller et al. (2016), we localise galaxies in the simulation in a two-step process. First, we identify all voxels for which $\rho_{\mathrm{DM}} > 1000\,\rho_c(z)$, with $\rho_c(z)$ being the critical density at redshift z. We then group all adjacent voxels together. This leads to 7397 voxel groups in the $(100\ \mathrm{Mpc})^3$ volume, which we interpret as galaxies.

For each galaxy, we obtain a tentative baryonic halo mass by summing up all baryonic mass within a sphere with a diameter of 1 Mpc centred around it. In order to identify simulated galaxies similar to NGC 6185, we seek to convert these halo masses into stellar masses.[13] Studies of the stellar mass–halo mass relation (e.g. Behroozi et al., 2013) show that stellar mass is a strictly increasing function of halo mass.[14] This implies that the same ordering that ranks galaxies by halo mass also ranks them by stellar mass; said differently, a galaxy's halo mass percentile score is the same as its stellar mass percentile score. We use this fact to map halo to stellar masses. To obtain a realistic stellar mass distribution to map to, we select SDSS DR7 galaxies with spectroscopic redshifts in the Local Universe and sort them by stellar mass. We discard the least massive ones until the galaxy number density is similar to that in our Enzo simulation snapshot. For each simulated galaxy, we calculate the halo mass percentile score, assume that its stellar mass percentile score is the same, and determine the corresponding stellar mass from the SDSS DR7 stellar mass distribution thus constructed.

In Fig. 5.10, we show three slices through the baryon density field around galaxies with a relative total matter density similar to that of NGC 6185. These are three different scenarios that could represent the actual baryon density field within NGC 6185's BORG SDSS voxel. We estimate the IGM density near each simulated galaxy not by taking the average, but by taking the median baryon density within the surrounding 1 Mpc^3 volume. In this way, we avoid contamination from the galaxy itself. Even if a galaxy were to measure 0.5 Mpc along each of three dimensions, its total volume would be 0.125 Mpc^3, and so its voxels are likely to occupy the upper 12.5% of baryon density percentile scores only. The median thus comfortably avoids these voxels. We denote the estimated small-scale relative IGM density $1+\delta_{\mathrm{IGM}}$ in Fig. 5.10 with white text and isopycnals. In Fig. 5.12, we demonstrate that the same IGM density estimation rule also works well in lower-density, non-group filament environments, which we envision will be the target of central interest in future applications of our technique.

For each simulated galaxy, we calculate both the large-scale relative total matter density $1 + \delta$ and the small-scale relative IGM density $1 + \delta_{\mathrm{IGM}}$. We aggregate the

[13] The Enzo simulations used do not contain sufficiently rich baryonic physics for stellar masses to be available directly.

[14] Gheller et al. (2016) have shown that applying such a relation to simulated galaxies (see their Eq. 3) leads to a reasonable match with the GAMA survey (Driver et al., 2009) stellar mass distribution.

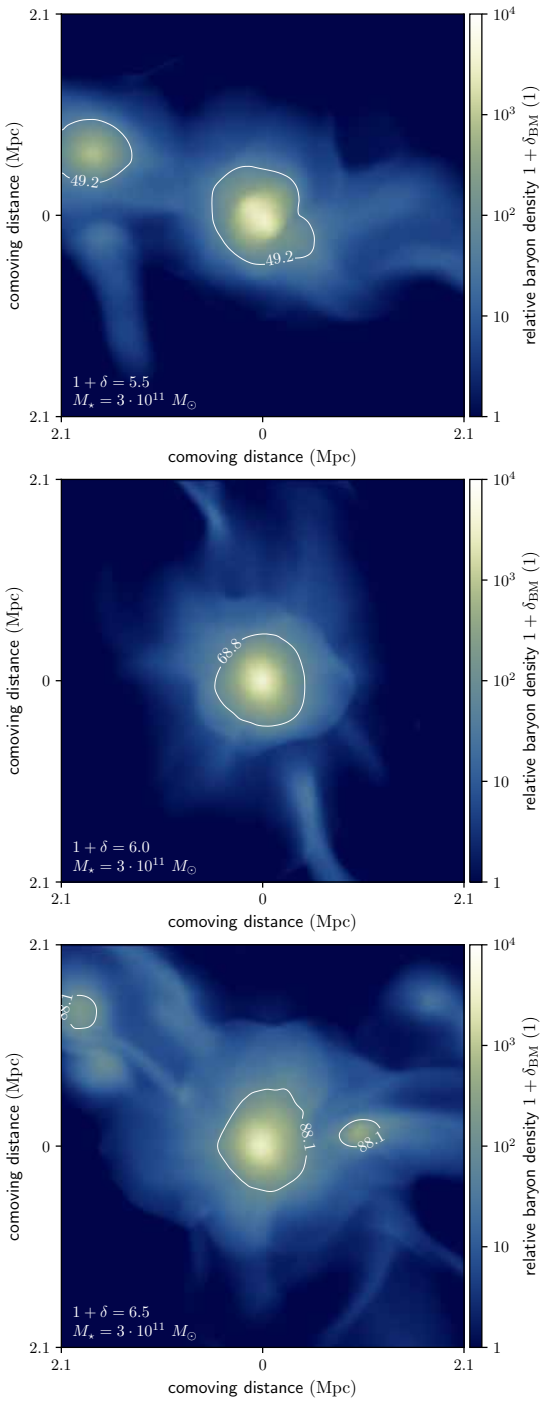


Figure 5.10: Three example possibilities for the baryonic density field within NGC 6185's BORG SDSS voxel, whose 4.2 Mpc–scale group-corrected relative total matter density $1+\delta = 6 \pm 2$. NGC 6185's stellar mass $M_\star = 3{\cdot}10^{11}\,M_\odot$. We show Enzo simulation slices of roughly 42 kpc thick, centred around simulated galaxies of comparable stellar mass. Each contour shows $1+\delta_{\rm IGM}$, the relative baryon density of the IGM estimated within a 1 Mpc^3 volume centred around the galaxy.

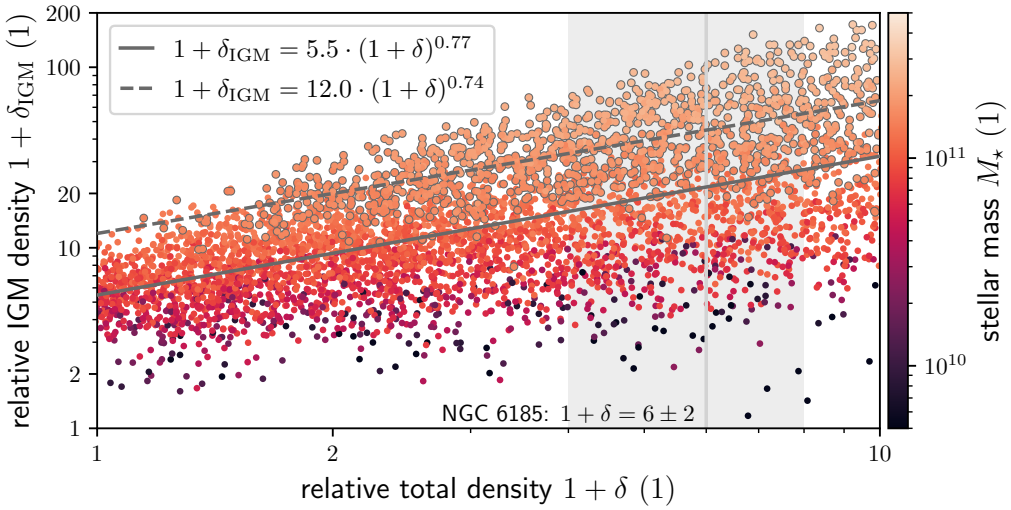


Figure 5.11: Relative total density $1 + \delta$ (including both baryonic and dark matter) versus relative IGM density $1 + \delta_{\rm IGM}$ (including baryonic matter only) around galaxies in a low-redshift snapshot of the Vazza et al. (2019) Enzo simulations. The former density is an average over a $(4.2\ \mathrm{Mpc})^3$ cubical volume, whilst the latter density is the median within a $1\ \mathrm{Mpc}^3$ spherical volume; in both cases, the galaxy lies at the centre. By colouring galaxies by their stellar mass $M_\star$, we reveal that a much more precise determination of $1 + \delta_{\rm IGM}$ can be achieved by conditioning on both $1 + \delta$ and $M_\star$. In solid dark grey, we show a best-fit power-law relation between $1 + \delta$ and $1 + \delta_{\rm IGM}$ for all simulated galaxies; in dotted dark grey, we show the analogous relation for those in the stellar mass range $1.5\text{–}6 \cdot 10^{11}\ M_\odot$ only. For these galaxies, the discrepancy with the stellar mass of NGC 6185 is at most a factor 2.

results in Fig. 5.11. At any given $1+\delta$ in the 1–10 range, there is an order of magnitude variation in the corresponding $1 + \delta_{\rm IGM}$. We fit a power-law relation to the data by squared error minimisation in log–log space and obtain $1 + \delta_{\rm IGM} = 5.5 \cdot (1 + \delta)^{0.77}$. If we restrict the fit to galaxies with a stellar mass at most a factor 2 different from that of NGC 6185 (i.e. $M_\star = 1.5\text{–}6 \cdot 10^{11}\ M_\odot$), we obtain $1 + \delta_{\rm IGM} = 12.0 \cdot (1 + \delta)^{0.74}$. However, to convert $1+\delta$ into $1+\delta_{\rm IGM}$ for NGC 6185, we must also take into account the variability in the relation. Furthermore, we must propagate the uncertainty in $1 + \delta$, which is most accurately done by sampling from the marginal distribution for NGC 6185's voxel using the full BORG SDSS MCMC. However at present, for simplicity, we just assume that $1 + \delta$ is lognormally distributed (Jasche & Wandelt, 2013). For NGC 6185, our statistical conversion relation then implies $1 + \delta_{\rm IGM} = 40^{+30}_{-10}$.

5.3.3 IGM temperature

We have estimated the pressure in the lobes of NGC 6185's giant, alongside the IGM density in the megaparsec-cubed–scale vicinity of the galaxy. Together, these quanti-

ties allow us to estimate the IGM temperature at the inner side of the lobes.

Because the lobes have a high internal sound speed compared to their environment as long as they are significantly underdense, the lobes do not maintain an internal pressure gradient. By contrast, their environment *does* feature a pressure gradient, so that the net force acting on the boundary of the lobe differs from point to point. As a result, the lobes cannot stay put, but rise buoyantly in the direction opposite to that of the local gravitational field. While the lobe is expanding at its outer tip, where it is overpressured with respect to its environment, the lobe is crushed at its inner tip, where an approximate pressure balance with the environment occurs: $P \approx P_{\mathrm{IGM}}$. This condition is key to infer the temperature of the IGM at the inner lobe tips. From Section 5.3.1, we have measured that the projected proper distances from NGC 6185 to the inner lobe tips are $d_{\mathrm{i,E}} = 0.31 \pm 0.02$ Mpc and $d_{\mathrm{i,W}} = 0.32 \pm 0.02$ Mpc. Taking into account the possibility of an extension along the line-of-sight dimension, the inner lobe tips occur at a distance $\gtrsim 0.3$ Mpc from the host galaxy. Because NGC 6185 is the most luminous and most massive galaxy of its group, we assume that it lies close to the group centre. We recall that — according to Tempel et al. (2017) — $R_{200} = 0.4$ Mpc, which these authors also identify with the group's virial radius. Thus, we consider P_{IGM} from the pressure balance condition to roughly correspond to the group's virial radius. We employ the ideal gas law to infer an IGM temperature from the IGM pressure and the IGM density. We find a temperature at the group's virial radius of $T_{\mathrm{IGM}} = 11^{+12}_{-5} \cdot 10^6$ K, or $k_{\mathrm{B}} T_{\mathrm{IGM}} = 0.9^{+1.0}_{-0.4}$ keV.

5.4 Discussion

5.4.1 Comparison to X-ray measurements of group temperatures

As mentioned in Section 5.2.2, the literature mass estimates for NGC 6185's group range between 0.7–2.6 $\cdot 10^{13}\, M_\odot$; our own estimate is $M = 1 \cdot 10^{13}\, M_\odot$. How does our inferred IGM temperature compare to X-ray observations of similar groups? A *Chandra* study of the IGM in the similarly spiral-rich group HCG 16 (O'Sullivan et al., 2014) has revealed low temperatures $T_{\mathrm{IGM}} = 3$–$4 \cdot 10^6$ K; however, the estimated group mass, $M_{500} = 4 \cdot 10^{12}\, M_\odot$, is also lower. The three lowest-mass objects in the *Chandra* sample of nearby groups by Sun et al. (2009) have total masses $M_{500} = 1.5 \cdot 10^{13}\, M_\odot$, $M_{500} = 2.0 \cdot 10^{13}\, M_\odot$, and $M_{500} = 3.2 \cdot 10^{13}\, M_\odot$. Their temperatures are $T_{500} = 9 \cdot 10^6$ K, $T_{500} = 11 \cdot 10^6$ K, and $T_{500} = 12 \cdot 10^6$ K, or $k_{\mathrm{B}} T_{500} = 0.8$ keV, $k_{\mathrm{B}} T_{500} = 1.0$ keV, and $k_{\mathrm{B}} T_{500} = 1.1$ keV, respectively. Moreover, using the scaling relation of Lovisari et al. (2015), a group mass $M_{500} = 1 \cdot 10^{13}\, M_\odot$ corresponds to $T_{\mathrm{IGM}} = 7 \cdot 10^6$ K, or $k_{\mathrm{B}} T_{\mathrm{IGM}} = 0.6$ keV. These examples show

that our IGM temperature estimate is broadly consistent with X-ray–derived temperatures of similarly massive groups. We recommend follow-up X-ray observations of the NGC 6185 group in order to directly measure T_{IGM} at the virial radius. A comparison between the X-ray- and GRG-inferred IGM temperature would not only put this particular result to the test, but would also provide a feeling of the general potential of our methodology.

5.4.2 IGM pressure balance and lobe smoothness

A key building block of our methodology is the condition of pressure balance at the inner lobe tips, as presented in Section 5.3.3. Here we argue that the observed smooth lobe shapes provide evidence that this condition indeed occurs.

As can be seen from Figs. 5.2 and 5.6, the lobes of the GRG of NGC 6185 appear to be of smooth morphology. Thanks to our proximity to the galaxy and the resolutions of the LoTSS, the lobes are highly resolved. Thus, the apparent smoothness cannot be due to a large physical scale per angular resolution element, but must instead be a feature intrinsic to the lobes. In comparison to other GRGs, the degree of smoothness is atypical: the double spheroid model of Section 5.3.1 would not provide a good fit to the MEC field of most other known GRG lobes. We hypothesise that the smoothness of the lobes of NGC 6185 might be the result of a surface tension effect. Wherever observations indicate a sharp boundary between a lobe and the surrounding IGM, the magnetic field of the lobe at the boundary must run parallel to it; if it would not, the plasma in the lobe would not remain confined and would instead start streaming into the IGM along the magnetic field lines. In turn, this would lead to a blurring of the boundary between lobe and IGM. A mixing shell with a 100 kpc thickness could form within a period of a few hundred kiloyears, which is short compared to an RG lifetime.[15] In the case of NGC 6185, a mixing shell with such thickness would be observable as a surface brightness gradient spanning a few arcminutes; however, a comparison between the top and middle panel of Fig. 5.6 suggests that the observed image is consistent with our sharp-boundary spheroid model. It is therefore reasonable to assume that the magnetic fields of the lobes of NGC 6185's GRG run approximately parallel to their surface. Weak shocks at lobe boundaries during the expansion phase also compress the magnetic fields of the IGM, boosting the magnetic field component parallel to the boundary by a factor of order unity (Guidetti et al., 2011). However, in typical cases, the magnetic field of the IGM is much weaker than that of the lobes, so that the former is not expected to play a major role in confining

[15] For any particular RG, it is therefore unlikely to observe it in a state without mixing shell if it were able to form one.

the lobe plasma.

In the last stage of lobe evolution, uncompensated adiabatic losses rapidly reduce the lobe's pressure. The pressure contribution from relativistic electrons, which may or may not dominate the lobe pressure at this evolutionary stage (e.g. O'Sullivan et al., 2013; Croston et al., 2018), will fall even more rapidly because of radiative losses. The IGM will start to compress the lobe. It does so at the inner side of the lobe only; at the outer side, the lobe rises buoyantly towards lower densities and pressures, and so remains locally overpressured. If the inner lobe would contract in a roughly shape-preserving way while its volume and surface area are reduced, its magnetic field lines would have to change direction in space more rapidly — said differently, the magnetic field curvature κ would increase. However, magnetic field lines resist being curved: bent field lines can be thought of as elastic bands under tension (Yang et al., 2019b) that exert an additional pressure $\propto \kappa B^2$ on the local plasma. A lower-energy configuration is achieved when the lines straighten out and the potential energy associated with the bent field lines is released. Therefore, inner lobe compression leads to a suppression of local lobe features, and over time, the inner lobe shape tends towards a featureless semi-ellipsoid. Thus, the smooth shapes of the lobes of NGC 6185's GRG are consistent with the inner lobe pressure balance scenario bound to occur at the end of the giant's life.

5.4.3 Evidence for late-stage radio galaxy evolution

Implausibly high age if RG assumed active

The total luminosity density of the GRG — $L_\nu\,(\nu = 150\ \mathrm{MHz}) = 3.3 \cdot 10^{24}\ \mathrm{W\ Hz^{-1}}$ — implies an implausibly high age if the RG were still active. The right panel of Fig. 12 of Hardcastle (2018), which describes results from a simulation-based analytical model, suggests that the two-sided jet power Q and the luminosity density $L_\nu\,(\nu = 150\ \mathrm{MHz})$ of active radio galaxies in environments of mass $M_{500} \sim 10^{13}\ M_\odot$ and at redshifts $z < 0.5$ approximately obey the proportionality

$$Q = L_\nu\,(\nu = 150\ \mathrm{MHz}) \cdot 10^{11}\ \mathrm{Hz}. \tag{5.9}$$

For the GRG of NGC 6185, we thus find $Q = 3.3 \cdot 10^{35}$ W.[16] At the epoch of observation, the combined internal energy of the lobes is $U = 2.0 \cdot 10^{52}$ J; see Table 5.5. The total energy carried by the jets over time, which can be divided by the jet power to esti-

[16] However, we note that a significant fraction of Hardcastle (2018)'s simulated active radio galaxies are outliers to this relation, especially as M_{500} approaches $10^{13}\ M_\odot$. This is not due to the crude assumption of a simple proportionality, as its predictions are quantitatively similar to those from the

mate the giant's age, also includes work W done on the external medium and energy lost through radiation. The simulations of Hardcastle & Krause (2013) show that, at least in clusters, the work done on the external medium is comparable to the internal energy of the lobes. The sum $U + W \approx 2U$ bounds the total energy carried by the jets from below and thus yields a lower bound to the giant's age: $\Delta t \gtrsim 4$ Gyr. Section 3.7 of Hardcastle (2018) predicts that the lobes' combined radiative losses consistently amount to $\sim$10% of Q, leading to a final estimate $\Delta t = 4.3$ Gyr. This age is excessively high, and in possible tension with the age of the NGC 6185 group itself. The abundance of spiral galaxies in the vicinity of NGC 6185 provides a qualitative argument for the youth of the group. For all group member candidates from Table 5.3, we collect from Mamon et al. (2020) STARLIGHT-based ages $\Delta t_\star$ within which half of the stellar mass has been formed. The arithmetic average age $\langle \Delta t_\star \rangle = 6$ Gyr. Interestingly, 7 galaxies out of the 19 for which this data is available have a $\Delta t_\star < 4$ Gyr. This suggests that major group formation activity has taken place in the last few gigayears; the dynamical disruption of NGC 6185 shown in Fig. 5.1 provides further evidence of this. Unfortunately, no $\Delta t_\star$ estimate is available for NGC 6185 itself.

If instead the GRG were *not* active, but a remnant, Eq. 5.9 could underestimate the true jet power by one or two orders of magnitude (Hardcastle, 2018). A significantly higher jet power would lower the GRG's estimated age into a plausible range ($\Delta t \in 10^2$–10^3 Myr) and relieve any potential tension with the group age. Besides, a jet power $Q \in 10^{36}$–10^{38} W would be much more common for currently known GRGs than one in the range 10^{35}–10^{36} W (Dabhade et al., 2020a); however, the former range is biased high because of selection effects. High jet powers are possible for spiral galaxy–hosted giants: J2345−0449 has a $Q \gtrsim 1.7 \cdot 10^{37}$ W (Bagchi et al., 2014).

Absence of extragalactic jets and hotspots

The LoTSS DR2 and VLASS images of NGC 6185 do not provide evidence for jet-mediated energy injection into the lobes at the epoch of observation. This implies that either jet feeding must have ceased entirely, or that it is still ongoing, but then through extragalactic jets faint enough to evade detection. As discussed in Section 5.2.1, the VLASS 2.2″ image strongly suggests the presence of current nuclear jets, but these are too weak to have generated the observed lobes. Under the lobe

power-law relation of Ineson et al. (2017), who found from combining radio and X-ray observations

$$Q = \left(\frac{L_\nu \left(\nu = 150\ \mathrm{MHz}\right)}{10^{25}\ \mathrm{W\,Hz^{-1}}} \right)^{0.9} \cdot 1.1 \cdot 10^{36}\ \mathrm{W}. \tag{5.10}$$

model of Section 5.3.1, the projected proper distances between the host galaxy and the outer lobe tips — where potential hotspots should reside — are $d_o \sim 1.3$ Mpc. Clearly, if no pockets of jet energy occur along the entire path from galaxy to outer lobe tip, such a hotspot will be devoid of jet-mediated energy injection for at least some time coming. Assuming $u = 0.1c$ for the average jet speed on the 10^0 Mpc–scale, no LoTSS DR2 or VLASS–detectable jet-mediated energy injection into potential hotspots will occur for at least ∼40 Myr. Of course, this period of future energy injection deprivation bounds the total period of energy injection deprivation from below. From observations of double–double radio galaxies (DDRGs; e.g. Mahatma et al., 2019), it appears that actively growing RGs regularly show jet activity hiatus of ∼1 Myr duration; however, hiatus of ∼10 Myr are much less common. A period of energy injection deprivation of the length calculated above therefore appears to be more consistent with a dying, rather than with a sputtering, RG scenario. Spiral galaxies usually do not have lobes, and we speculate that their formation in the current case has been the result of a rare merger event, now largely foregone, which triggered SMBH activity. The galaxy's disturbed appearance, shown in Fig. 5.1, supports this scenario.

5.4.4 Future prospects

Robustness and tightness of IGM density inference

One of the main sources of uncertainty in the IGM temperature inferred in this work comes from the statistical determination of IGM density $1 + \delta_{\rm IGM}$ from total density $1 + \delta$ and stellar mass $M_\star$; see Fig. 5.11. Future work should test the robustness of this relation by rederiving it from another cosmological simulation, such as from EAGLE or IllustrisTNG. Besides, it appears worthwhile to explore how sensitively the inferred IGM density depends on its exact definition. In this work, we have understood the IGM density around a galaxy to be the median baryonic matter density within a 1 Mpc^3 spherical volume; however, reasonable alternatives certainly appear possible.

If indeed proven robust, it makes sense to investigate whether the determination of $1 + \delta_{\rm IGM}$ can be further tightened by conditioning on additional information available for both simulated galaxies and the observed galaxy of interest. This, however, requires cosmological simulations with sufficiently rich galactic physics; again, EAGLE and IllustrisTNG appear to be attractive contemporary simulation suites. Additional information to condition on could be the galaxy' s morphological type (for NGC 6185: SAa spiral), or the number density of other — sufficiently massive — galaxies in some Mpc-scale vicinity (for NGC 6185: see Table 5.3).

Simulations with rich galactic physics also naturally generate galaxies' stellar masses. Using these stellar masses would eliminate the uncertainty that now arises from the mapping of baryonic halo masses to stellar masses as described in Section 5.3.2. Currently, to determine $1+\delta_{\mathrm{IGM}}$, we make use of simulated galaxies within a rather broad range of stellar mass around NGC 6185's $M_\star = 3.0^{+1.2}_{-0.9} \cdot 10^{11}\, M_\odot$: $M_\star \in 1.5$–$6 \cdot 10^{11}\, M_\odot$. Once we are more confident that the stellar masses from the simulation are reliable, we could reduce this range to just the range required by the uncertainty in the observed galaxy's stellar mass. This, in turn, would reduce the uncertainties of both the IGM density and temperature estimates.

More accurate large-scale structure reconstructions

Our methodology depends on a determination of the large-scale total density $1+\delta$, which we correct for group presence in the case of NGC 6185. What if the BORG SDSS measurement $1+\delta = 2.3 \pm 0.7$ is inaccurate? An improved BORG run, the BORG 2M++ (Jasche & Lavaux, 2019), is already available (though not publicly). This BORG data set has all-sky coverage, a higher spatial resolution, and more accurate selection functions, bias modelling, and gravitational dynamics — albeit at the cost of probing a more limited redshift range. Fortunately, at $z = 0.03$, NGC 6185 falls within the BORG 2M++ volume. Future work should establish whether the total density derived from the BORG 2M++ (or similar data sets) leads to an inferred IGM density consistent with the estimate derived here. In this way, one could build further confidence that our methodology is robust.

Comparison to simulated IGM temperatures

Finally, simulations with rich galactic physics also include supernova and AGN feedback on the IGM. Gheller & Vazza (2019) have shown that past AGN activity can significantly boost the temperature of the IGM surrounding the host galaxy. As a result, only such simulations can give a realistic idea of the IGM temperature variation around galaxies with recent AGN activity.[17] By constructing a simulation relation analogous to that of Fig. 5.11, but now between T_{IGM} on the one hand and $1+\delta$ and $M_\star$ on the other, it will be possible to evaluate whether or not our estimate $T_{\mathrm{IGM}} = 11^{+12}_{-5} \cdot 10^6$ K tightens the distribution already suggested by $1+\delta$ and $M_\star$. This, in turn, would quantify the information gain achieved by our technique. One can also turn this around: discrepancies between inferred and simulated IGM tem-

[17] Because we do not use a simulation with AGN feedback in the present work, we have for now omitted a comparison with simulated IGM temperatures. The results would be unreliable.

peratures could offer a way to calibrate and test (sub-grid) AGN feedback models, which are still largely uncertain.

Potential with present-day data

The technique put forward in the current work does not exploit features unique to NGC 6185's giant, and thus may be applied to other targets in the future. Notably, we have not yet reached the limits of the technique as set by today's data quality. From the top panel of Fig. 5.6, it is apparent that we could have visually recognised, re-imaged and analysed the GRG of NGC 6185 correctly if its lobes had been significantly fainter. In Appendix 5.A3, we quantify this detection limit, by simulating Gaussian noise, adding it to the original image, and rescaling the result. We demonstrate that the lobes could not have been 4 times fainter — see Fig. 5.13 — but that analysis with little loss of fidelity is possible if they had been 3 times fainter; see Fig. 5.14. This latter analogon would have an equipartition pressure $P_{\rm eq} = 3{\cdot}10^{-16}$ Pa and — if it would occur in the same environment as NGC 6185 — an IGM temperature $T_{\rm IGM} = 4^{+3}_{-2} \cdot 10^6$ K, or $k_{\rm B} T_{\rm IGM} = 0.3^{+0.3}_{-0.1}$ keV. Such estimates would be on par with the lowest X-ray group temperature measurements available today (Lovisari et al., 2021).

The next step is to apply the method not to a single case, but to a sample. Some promising targets, for which preliminary LoTSS DR2 analysis suggests that $P \in 10^{-16}$–10^{-15} Pa, include the southern lobe of the LEDA 2048533 GRG ($z = 0.06$) and the southern remnant lobe of the LEDA 2103724 GRG ($z = 0.16$) (Oei et al., 2023a). It would be of particular interest to study targets in groups whose temperatures or temperature profiles are known from X-ray observations.

5.5 Conclusion

In this work, we have demonstrated a new IGM temperature estimation technique, which probes the edges of galaxy groups and holds promise to extend into the WHIM. We combine a radio galaxy image, stellar mass and redshift information, large-scale structure reconstructions, and cosmological simulations.

1. We demonstrate our methodology using NGC 6185, a spiral galaxy in a nearby ($z = 0.0343 \pm 0.0002$) filament of the Cosmic Web. It is the most luminous and massive ($M_\star = 3 \cdot 10^{11}\, M_\odot$) member of a galaxy group ($M \sim 10^{13}\, M_\odot$), and has generated the projectively longest ($l_{\rm p} = 2.45 \pm 0.01$ Mpc) known GRG of all spiral galaxies. Spiral galaxy–hosted GRGs are exceedingly rare.

At $\nu = 150$ MHz, this aged FRII radio galaxy has a luminosity density $L_\nu = 3.3 \pm 0.3 \cdot 10^{24}$ W Hz^{-1}.

2. We apply a Bayesian parametric 3D lobe model to a LoTSS radio image. We assume spheroidal lobe shapes with axes of revolution that pierce through a common point near the host galaxy, optically thin volume-filling lobe plasmata of constant monochromatic emission coefficient, and Gaussian image noise. We infer lobe volumes $V \sim 0.2$ Mpc^3 and equipartition lobe pressures $P_{\rm eq} \sim 6 \cdot 10^{-16}$ Pa — amongst the lowest hitherto found. Using an X-ray based statistical conversion, we find a true lobe pressure $P = 1.5^{+1.7}_{-0.4} \cdot 10^{-15}$ Pa.

3. From the BORG SDSS, an SDSS-derived Monte Carlo Markov chain of possible Local Universe density fields, we measure that the total density averaged over a $(4.2\ \text{Mpc})^3$ volume around NGC 6185 is $1 + \delta = 2.3 \pm 0.7$. For the particular case of NGC 6185, we perform a group correction, yielding $1 + \delta = 6 \pm 2$. Next, from Enzo cosmological simulations, we determine the relationship between $1 + \delta$ and the typical IGM density $1 + \delta_{\rm IGM}$ in a 1 Mpc^3 volume around galaxies with a stellar mass similar to NGC 6185's. Applying this relationship to NGC 6185, we find $1 + \delta_{\rm IGM} = 40^{+30}_{-10}$.

4. Radio galaxy lobes, being significantly underdense with respect to their environment, rise buoyantly in the direction opposite to the local gravitational field. Especially for aged lobes, the IGM crushes the lobes around the inner lobe tips, causing an approximate local balance between IGM and lobe pressure: $P_{\rm IGM} \approx P$. From this physical effect and the ideal gas law, we deduce that $T_{\rm IGM} = 11^{+12}_{-5} \cdot 10^6$ K, or $k_{\rm B} T_{\rm IGM} = 0.9^{+1.0}_{-0.4}$ keV. This temperature corresponds to a distance $\gtrsim 0.3$ Mpc from NGC 6185; as such, we probe the thermodynamics at the virial radius of the group.

5. Interestingly, our case study does not yet fully demonstrate the potential of our technique. Given the noise levels of currently available LoTSS survey data, it is possible to perform the above analysis on a three times fainter analogon to NGC 6185's GRG. This would allow one to probe temperatures on par with the lowest X-ray group temperatures available today.

6. Although we have currently applied our method to a radio galaxy whose host lies in a group, no step requires this to be the case. Quite to the contrary, it is likely that BORG-like large-scale structure reconstructions are more accurate for non-group filament environments. Our method thus holds promise to extend beyond the outskirts of galaxy groups — and into the WHIM. We

envision our estimate to be the first of many temperature constraints in Cosmic Web filaments from radio galaxy observations. At the moment of writing, there are hundreds of other LOFAR-imaged giants that reside within the volume probed by the BORG SDSS, and future radio surveys and large-scale structure reconstructions will expand in sky coverage and depth. Our methodology, which bypasses expensive X-ray observations, might thus be employed on a large scale in the future. This, in turn, could lead to a three-dimensional map of WHIM temperatures and temperature bounds at concrete locations within the nearby Cosmic Web.

M.S.S.L. Oei and R.J. van Weeren acknowledge support from the VIDI research programme with project number 639.042.729, which is financed by the Dutch Research Council (NWO). F. Vazza acknowledges support from the ERC STG MAGCOW (714196) from the H2020. The LOFAR is the Low-Frequency Array designed and constructed by ASTRON. It has observing, data processing, and data storage facilities in several countries, which are owned by various parties (each with their own funding sources), and which are collectively operated by the ILT Foundation under a joint scientific policy. The ILT resources have benefited from the following recent major funding sources: CNRS–INSU, Observatoire de Paris and Université d'Orléans, France; BMBF, MIWF–NRW, MPG, Germany; Science Foundation Ireland (SFI), Department of Business, Enterprise and Innovation (DBEI), Ireland; NWO, The Netherlands; the Science and Technology Facilities Council, UK; Ministry of Science and Higher Education, Poland; the Istituto Nazionale di Astrofisica (INAF), Italy. The cosmological simulations used in this work were produced with the Enzo code (`enzo-project.org`) and run on the Piz-Daint supercluster at LSCS (Lugano) under project 's701' with F. Vazza as P.I.. This work made use of the `legacystamps` package (`https://github.com/tikk3r/legacystamps`). The Legacy Surveys consist of three individual and complementary projects: the Dark Energy Camera Legacy Survey (DECaLS; Proposal ID #2014B-0404; PIs: David Schlegel and Arjun Dey), the Beijing–Arizona Sky Survey (BASS; NOAO Prop. ID #2015A-0801; PIs: Zhou Xu and Xiaohui Fan), and the Mayall z-band Legacy Survey (MzLS; Prop. ID #2016A-0453; PI: Arjun Dey). DECaLS, BASS and MzLS together include data obtained, respectively, at the Blanco telescope, Cerro Tololo Inter-American Observatory, NSF's NOIRLab; the Bok telescope, Steward Observatory, University of Arizona; and the Mayall telescope, Kitt Peak National Observatory, NOIRLab. The Legacy Surveys project is honored to be permitted to conduct astronomical research on Iolkam Du'ag (Kitt Peak), a mountain with particular significance to the Tohono O'odham Nation. NOIRLab is operated by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation. This project used data obtained with the Dark Energy Camera (DECam), which was constructed by the Dark Energy Survey (DES) collaboration. Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundacao Carlos Chagas Filho de Amparo, Financiadora de Estudos e Projetos, Fundacao Carlos Chagas Filho de Amparo a Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Cientifico e Tecnologico and the Ministerio da Ciencia, Tecnologia e Inovacao, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey. The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energeticas, Medioambientales y Tecnologicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenössische Technische Hochschule (ETH) Zürich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciencies de l'Espai (IEEC/CSIC), the Institut de Fisica d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig Maximilians Universität München and the associated Excellence Cluster Universe, the University of Michigan, NSF's NOIRLab, the University of Nottingham, the Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, and Texas A&M University. BASS is a key project of the Telescope Access Program (TAP), which has been funded by the National Astronomical Observatories of China, the Chinese Academy of Sciences (the Strategic Priority Research Program "The Emergence of Cosmological Structures" Grant # XDB09000000), and the Special Fund for Astronomy from the Ministry of Finance. The BASS

is also supported by the External Cooperation Program of Chinese Academy of Sciences (Grant # 114A11KYSB20160057), and Chinese National Natural Science Foundation (Grant # 11433005). The Legacy Survey team makes use of data products from the Near-Earth Object Wide-field Infrared Survey Explorer (NEOWISE), which is a project of the Jet Propulsion Laboratory/California Institute of Technology. NEOWISE is funded by the National Aeronautics and Space Administration. The Legacy Surveys imaging of the DESI footprint is supported by the Director, Office of Science, Office of High Energy Physics of the U.S. Department of Energy under Contract No. DE-AC02-05CH1123, by the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility under the same contract; and by the U.S. National Science Foundation, Division of Astronomical Sciences under Contract No. AST-0950945 to NOAO.

Data Availability

The LoTSS DR2 is publicly available at `https://lofar-surveys.org/dr2_release.html`. VLASS Quick Look images are publicly available at `https://science.nrao.edu/science/surveys/vlass`. The BORG SDSS data release is publicly available at `https://github.com/florent-leclercq/borg_sdss_data_release`. BORG SDSS total matter densities $1 + \delta$ for GRGs are tabulated in Appendix A of Oei et al. (in preparation).

5.A1 Average density ratio for WHIM with isothermal β-profile

Here we derive the average density ratio formulae of Eqs. 5.7 and 5.8, alongside the average density ratio asymptote for fixed parameters.
We consider a filament whose WHIM density profile obeys the isothermal β-model:

$$\rho(r) := \rho(0) \left(1 + \left(\frac{r}{r_c}\right)^2\right)^{-\frac{3}{2}\beta}, \tag{5.11}$$

with $\rho(0)$, r_c, and β free parameters. A cylinder of radius r and length l has an enclosed mass $m(r)$ and average density $\langle\rho\rangle(r)$ given by

$$m(r) = \int_0^r \rho(r') \cdot l \cdot 2\pi r' \, \mathrm{d}r'; \quad \langle\rho\rangle(r) = \frac{m(r)}{\pi r^2 l}, \tag{5.12}$$

so that

$$\langle\rho\rangle(r) = \frac{2\rho(0)}{r^2} \int_0^r \left(1 + \left(\frac{r'}{r_c}\right)^2\right)^{-\frac{3}{2}\beta} r' \, \mathrm{d}r'. \tag{5.13}$$

The integration result depends on β. For $\beta \neq \frac{2}{3}$, we find

$$\langle\rho\rangle(r) = \frac{\rho(0)}{1-\frac{3}{2}\beta}\frac{r_c^2}{r^2}\left(\left(1+\left(\frac{r}{r_c}\right)^2\right)^{1-\frac{3}{2}\beta} - 1\right), \tag{5.14}$$

whilst for $\beta = \frac{2}{3}$, we find

$$\langle\rho\rangle(r) = \rho(0)\frac{r_c^2}{r^2}\ln\left(1+\left(\frac{r}{r_c}\right)^2\right). \tag{5.15}$$

Upon dividing $\langle\rho\rangle(r)$ by $\langle\rho\rangle(R)$, the shared factors $\rho(0)\, r_c^2$ and $(1-\frac{3}{2}\beta)^{-1}$ cancel, and we arrive at Eqs. 5.7 and 5.8 for the average density ratio.
From Eq. 5.11, we see that for $r \ll r_c$ the density is approximately constant — at level $\rho(0)$. For this reason, one expects that

$$\lim_{\frac{r}{r_c}\to 0}\langle\rho\rangle(r) = \rho(0), \tag{5.16}$$

a fact that can be formally verified from Eqs. 5.14 and 5.15 using L'Hôpital's rule. Because $\beta > 0$, $\rho(r)$ attains its maximum at $r = 0$. For $\beta \neq \frac{2}{3}$, the average density ratio asymptote therefore is

$$\lim_{\frac{r}{r_c}\to 0}\frac{\langle\rho\rangle(r)}{\langle\rho\rangle(R)} = \frac{R^2}{r_c^2}\frac{1-\frac{3}{2}\beta}{\left(1+\left(\frac{R}{r_c}\right)^2\right)^{1-\frac{3}{2}\beta} - 1}, \tag{5.17}$$

whilst for $\beta = \frac{2}{3}$, we find

$$\lim_{\frac{r}{r_c}\to 0}\frac{\langle\rho\rangle(r)}{\langle\rho\rangle(R)} = \frac{R^2}{r_c^2}\ln^{-1}\left(1+\left(\frac{R}{r_c}\right)^2\right). \tag{5.18}$$

5.A2 Large-to-small-scale density conversion: the low-density regime

In this section, we provide a visual impression of the practical effect of the 1 Mpc–scale IGM density definition introduced in Section 5.3.2. This definition is key in the conversion from large-scale density $1+\delta$ to small-scale density $1+\delta_{\rm IGM}$.
As an example, we consider an NGC 6185 analogon with a 4.2 Mpc–scale relative

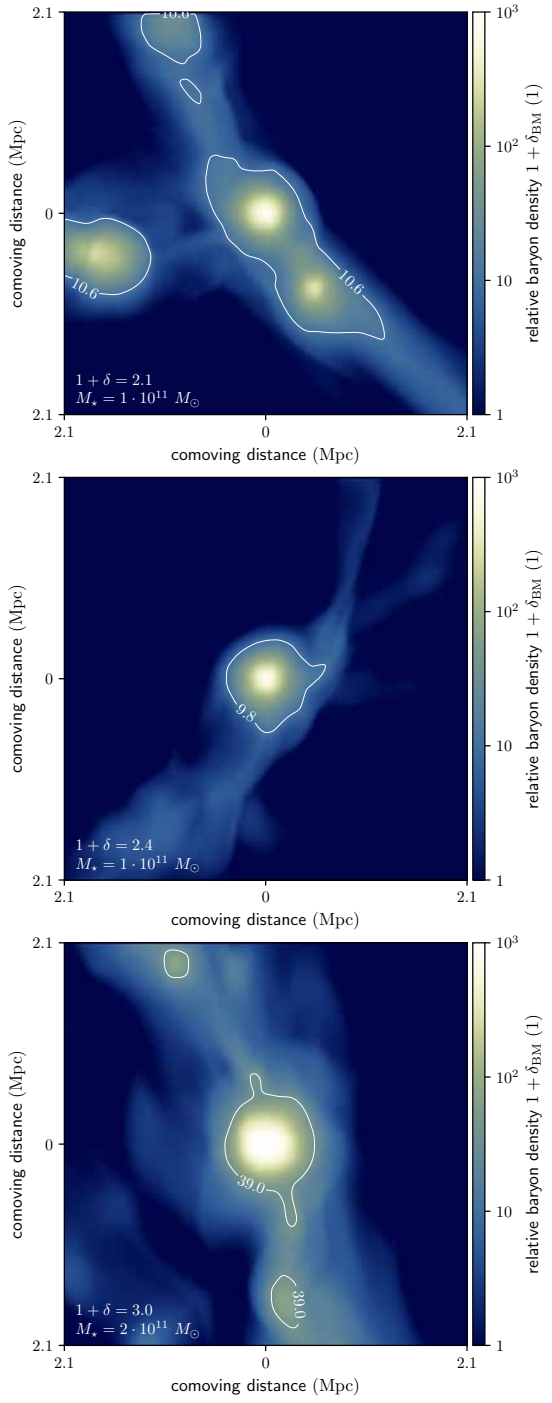


Figure 5.12: As in Fig. 5.10, but for an NGC 6185 analogon that does not occur in a galaxy group, but in a lower-density Cosmic Web filament environment. Such a galaxy might have a 4.2 Mpc–scale relative total matter density $1 + \delta = 2.3 \pm 0.7$. As is clear from visual comparison to Fig. 5.10, Enzo simulation cutouts consistent with this scenario feature both less dense 1 Mpc–scale environments as well as less massive galaxies. Note the more restricted colour bar scale here.

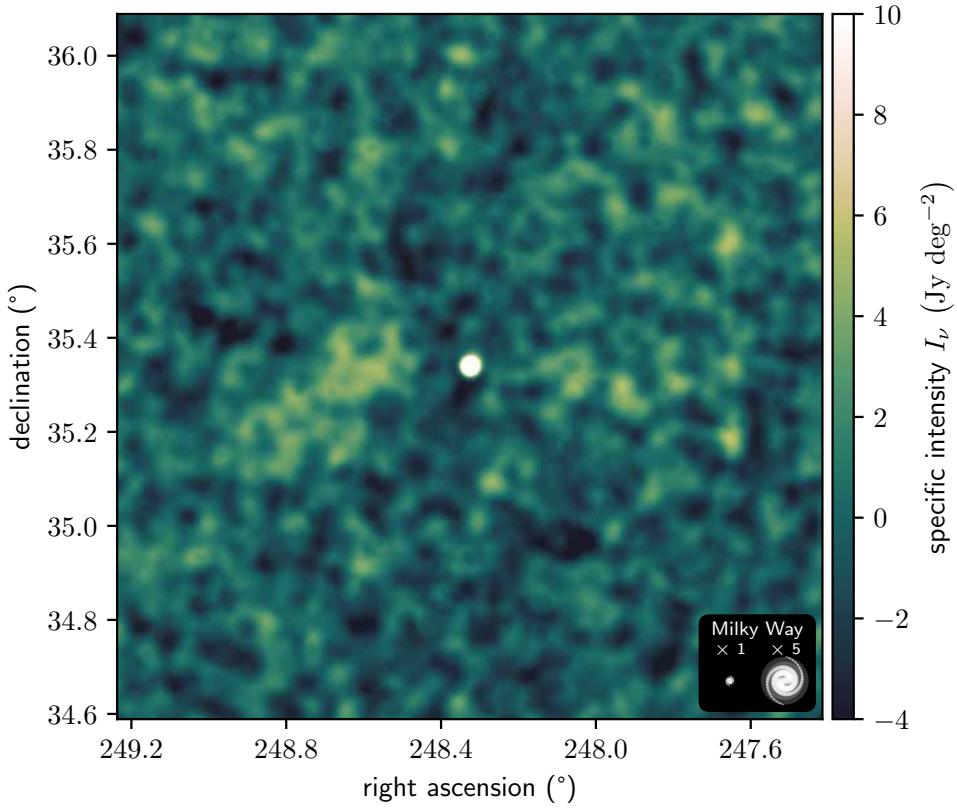


Figure 5.13: A simulated 90″ image at $\nu_{\rm obs} = 144$ MHz of a 4 times fainter analogon of NGC 6185's GRG. Given the LoTSS DR2 90″ noise level $\sigma \sim 2$ Jy deg^{-2}, such a GRG is unlikely to be recognised, correctly deconvolved, and analysed.

total matter density $1 + \delta = 2.3 \pm 0.7$. In Fig. 5.12, we show three Enzo simulation cutouts consistent with this scenario. As can be seen from the top and, to a lesser degree, the bottom panel of Fig. 5.12, the thus-defined IGM density surrounding the galaxy may represent the IGM density of a broader part of the filament. By comparing Fig. 5.12 to Fig. 5.10, it appears that the 1 Mpc–scale IGM density definition introduced in Section 5.3.2 not only produces reasonable density estimates in group-like environments, but also in lower-density filament environments. This is important, because future work may attempt to constrain IGM thermodynamics outside galaxy groups with the methodology presented in this work.

5.A3 Lower WHIM temperature constraints from current-day radio data

In this section, we demonstrate that our methodology, combined with current-day radio survey data, such as the publicly available LoTSS DR2, in principle allow for more stringent WHIM temperature constraints than those derived in the current work — that is, given the availability of a suitable target.

In Fig. 5.13, we show that given typical LoTSS DR2 noise levels, analoga of NGC 6185's GRG but with surface brightnesses that are 4 times lower, are unsuitable targets. However, in Fig. 5.14, we show that analoga that are 2 or 3 times fainter can be accurately analysed. For each analogon, we repeat the Metropolis–Hastings MCMC procedure described in Section 5.3.1. A comparison between the middle panels of Figs. 5.6 and 5.14 reveals that lobe geometry inference remains stable far into the low signal-to-noise regime. We find that the 3 times fainter analogon yields an IGM temperature estimate of $T_{\mathrm{IGM}} = 4^{+3}_{-2} \cdot 10^6$ K, or $k_{\mathrm{B}} T_{\mathrm{IGM}} = 0.3^{+0.3}_{-0.1}$ keV.

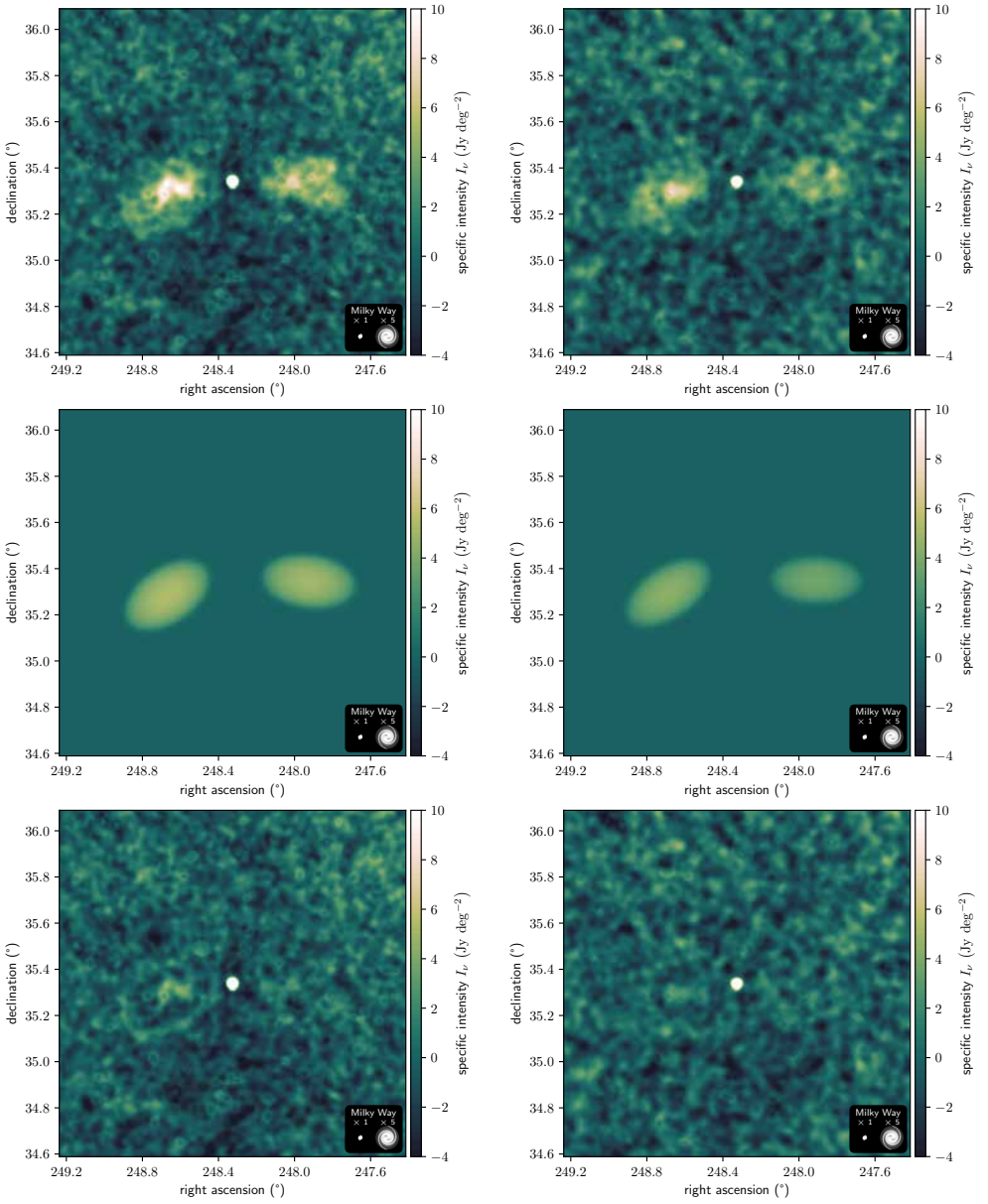


Figure 5.14: Bayesian radio galaxy lobe model fits in the sense of Section 5.3.1 applied to simulated analoga of NGC 6185's GRG, but with lobes that are 2 (*left column*) and 3 (*right column*) times fainter. This is a visual demonstration of the fact that (minimum energy and equipartition) lobe pressures can be robustly extracted from analoga of NGC 6185's GRG whose surface brightnesses are a factor of order unity lower. The figure is similar to Fig. 5.6, but has a different colour bar scaling. Alongside 90″ images at $\nu_{\mathrm{obs}} = 144$ MHz of the 2 and 3 times fainter analoga of NGC 6185's GRG (*top row*), we show 90″ MAP estimates from the Bayesian radio galaxy lobe model (*middle row*) and residual images (*bottom row*) obtained by subtracting the middle row images from the top row images.

Mathematics, rightly viewed, possesses not only truth, but supreme beauty — a beauty cold and austere, like that of sculpture, without appeal to any part of our weaker nature, without the gorgeous trappings of painting or music, yet sublimely pure, and capable of a stern perfection such as only the greatest art can show.

Bertrand A. W. Russell, British mathematician and philosopher, *The Study of Mathematics (1907)*

6

Measuring the giant radio galaxy length distribution with the LoTSS

M. S. S. L. Oei, R. J. van Weeren, A. R. D. J. G. I. B. Gast, A. Botteon, M. J. Hardcastle, P. Dabhade, T. W. Shimwell, H. J. A. Röttgering, A. Drabent — *Astronomy & Astrophysics*, 672, 163, 2023

Abstract

Context Many massive galaxies launch jets from the accretion disk of their central black hole, but only $\sim 10^3$ instances are known in which the associated outflows form giant radio galaxies (GRGs, or giants): luminous structures of megaparsec extent that consist of atomic nuclei, relativistic electrons, and magnetic fields. Large samples are imperative to understanding the enigmatic growth of giants, and recent systematic searches in homogeneous surveys constitute a promising development. For the first time, it is possible to perform meaningful precision statistics with GRG lengths, but a framework to do so is missing.

Aims We measured the intrinsic GRG length distribution by combining a novel statistical framework with a LOFAR Two-metre Sky Survey (LoTSS) sample of freshly discovered giants. In turn, this allowed us to answer an array of questions on giants. For example, we can now assess how rare a 5 Mpc giant is compared with one of 1 Mpc, and how much larger — given a projected length — the corresponding intrinsic length is expected to be. Notably, we can now also infer the GRG number density in the Local Universe.

Methods We assumed the intrinsic GRG length distribution to be Paretian (i.e. of power-law form) with tail index ξ, and predicted the observed distribution by modelling projection and selection effects. To infer ξ, we also systematically searched the LoTSS for hitherto unknown giants and compiled the largest catalogue of giants to date.

Results We show that if intrinsic GRG lengths are Pareto distributed with index ξ, then projected GRG lengths are also Pareto distributed with index ξ. Selection effects induce curvature in the observed projected GRG length distribution: angular length selection flattens it towards the lower end, while surface brightness selection steepens it towards the higher end. We explicitly derived a GRG's posterior over intrinsic lengths given its projected length, laying bare the ξ dependence. We also discovered 2060 giants within LoTSS DR2 pipeline products; our sample more than doubles the known population. Spectacular discoveries include the largest, second-largest, and fourth-largest GRG known ($l_{\rm p} = 5.1$ Mpc, $l_{\rm p} = 5.0$ Mpc, and $l_{\rm p} = 4.8$ Mpc), the largest GRG known hosted by a spiral galaxy ($l_{\rm p} = 2.5$ Mpc), and the largest secure GRG known beyond redshift 1 ($l_{\rm p} = 3.9$ Mpc). We increase the number of known giants whose angular length exceeds that of the Moon from 10 to 23; among the discoveries is the angularly largest known radio galaxy in the Northern Sky, which is also the angularly largest known GRG ($\varphi = 2^\circ$). Combining theory and data, we determined that intrinsic GRG lengths are well described by a Pareto distribution, and measured the index $\xi = -3.5 \pm 0.5$. This implies that, given its projected length, a GRG's intrinsic length is expected to be just 15% larger. Finally, we determined the comoving number density of giants in the Local Universe to be $n_{\rm GRG} = 5 \pm 2\ (100\ {\rm Mpc})^{-3}$.

Conclusions We developed a practical mathematical framework that elucidates the statistics of giant radio galaxy lengths. Through a LoTSS search, we also discovered 2060 new giants. By combining both advances, we determined that intrinsic GRG lengths are well described by a Pareto distribution with index $\xi = -3.5 \pm 0.5$, and that giants are truly rare in a cosmological sense: most clusters and filaments of the Cosmic Web are not currently home to a giant. Thus, our work yields new observational constraints for analytical models and simulations featuring radio galaxy growth.

Key words: Galaxies: active – jets – kinematics and dynamics – Radio continuum: galaxies

6.1 Introduction

When gas, dust, and stars accrete onto a supermassive black hole (SMBH) in the centre of a galaxy, collimated jets arise along the Kerr rotation axis that blast some of the infalling material into the intergalactic medium (IGM) (e.g. Blandford & Rees, 1974). In this process, the ejecta dissolve into a relativistic plasma that drags along a magnetic field and glows in synchrotron light. The resulting luminous structure is called a radio galaxy (RG); the central black hole that has generated it an active galactic nucleus (AGN).

It is increasingly clear that RGs and their AGN play an important role in galaxy evolution and cosmology. By heating the interstellar medium (ISM) or even expelling it from their host galaxies through galactic superwinds, AGN quench star formation (e.g. Di Matteo et al., 2005). Beckmann et al. (2017) have shown that AGN-induced star formation quenching is most pronounced in massive galaxies. There is also compelling evidence that the accompanying RGs provide the energy necessary to stop (e.g. McNamara & Nulsen, 2012; Yang et al., 2019a) bremsstrahlung-mediated cooling flows (Fabian et al., 1984) in clusters of galaxies. In the absence of cooling flows, the intra-cluster medium (ICM) remains dilute and hot, and galaxies in the centres of clusters are denied infalling gas that could otherwise reignite star formation. Cosmological simulations that include this RG feedback indeed resolve (e.g. Croton et al., 2006) the overprediction of baryonic masses and luminosities of central cluster galaxies that early simulations found. Finally, RGs may be responsible for magnetising the IGM that pervades the filaments of the Cosmic Web (e.g. Vazza et al., 2017).

Despite the emerging picture that RGs trace quenched star formation, inhibit cooling flows, and magnetise filaments, our knowledge of them is far from complete. Concerning geometry, a major unknown is the exact connection between the morphology of RGs and the pressure field of the ambient IGM, especially in filaments and cluster outskirts. Another question is whether small and large RGs come from the same initial population, or whether their growth is driven by distinct physical processes. Finally, we do not know how large can RGs become, and, more generally, how many RGs there are of each length.

To test RG growth models that answer these and other questions, it is imperative to study the subpopulation of most spatially extreme RGs: the giant radio galaxies (GRGs). The defining feature of giants is that their proper lengths — when projected onto the celestial sphere — exceed some threshold $l_{\mathrm{p,GRG}}$, which is canonically chosen as 0.7 Mpc or 1 Mpc. If $l_{\mathrm{p,GRG}} = 0.7$ Mpc, then the preceding literature describes a total of 1281 giants.

In recent years, several studies have successfully searched for giants in systematic,

wide-area surveys such as the NRAO VLA Sky Survey (NVSS; Condon et al., 1998) and the LOFAR Two-metre Sky Survey (LoTSS; Shimwell et al., 2017). A combination of manual (i.e. visual) and automated searches (Solovyov & Verkhodanov, 2011; Solovyov & Verkhodanov, 2014; Amirkhanyan, 2016; Proctor, 2016; Dabhade et al., 2017, 2020a) in the NVSS yielded 313 new giants (24% of the aforementioned literature population). Meanwhile, Dabhade et al. (2020b) discovered 225 new giants (17% of this same population) in the LoTSS DR1 (Shimwell et al., 2019), whose survey footprint is 80 times smaller than NVSS's. Such searches have the advantage of introducing almost homogeneous selection effects throughout the survey footprint, which can potentially be modelled and thus corrected for during any subsequent statistical inference.

In this work this idea comes to fruition, by conducting a precision analysis of the intrinsic giant radio galaxy length distribution. To do so, we require two ingredients. First, in Sect. 6.2, we develop a statistical framework that allows one to answer probabilistic questions regarding both large samples of giants and individual specimina. Then, in Sect. 6.3, we describe our LoTSS DR2 (Shimwell et al., 2022) GRG search campaign and the trove of previously unknown giants it has yielded; moreover, we describe the assemblage of the most complete GRG catalogue to date. In Sect. 6.4, combining theory and data, we infer the tail index parameter that describes the intrinsic GRG length distribution, which constrains future models and simulations aimed at understanding RG growth. In Sect. 6.5, we discuss caveats of the current work and give recommendations for future extensions, before we present conclusions in Sect. 6.6.

We assume a concordance inflationary Λ cold dark matter cosmology with parameters $\mathfrak{M}$ from Planck Collaboration et al. (2020); that is to say $\mathfrak{M} = (h = 0.6766, \Omega_{\mathrm{BM},0} = 0.0490, \Omega_{\mathrm{M},0} = 0.3111, \Omega_{\Lambda,0} = 0.6889)$, where $H_0 := h \cdot 100\ \mathrm{km\ s^{-1}\ Mpc^{-1}}$. We define the spectral index α such that it relates to flux density F_ν at frequency ν as $F_\nu \propto \nu^\alpha$, and define giants using threshold $l_{\mathrm{p,GRG}} := 0.7$ Mpc. Regarding the terminology, we use 'angular length' where others use 'largest angular size' (LAS), and 'projected proper length' where others use 'largest linear size' (LLS).[1]

[1] An object's size can refer to either its one-, two-, or three-dimensional extent. We therefore consider 'angular size' to be ambiguous terminology; we propose that 'angular length' better captures one-dimensionality. Naturally, an object's 'length' is understood to be its total one-dimensional extent, so that the qualifier 'largest' seems superfluous. We further remark that 'length' is synonymous with, but more succinct than, 'linear size'. In cosmology, one must distinguish between proper and comoving lengths, especially when objects are not gravitationally bound — like GRG lobes. In this work, we consider two types of proper lengths: intrinsic proper lengths and projected proper lengths.

6.2 Theory

To measure the intrinsic GRG length distribution, we must first establish a suitable statistical framework. In this section, we provide a summary of the theory developed in Appendix 6.A1. Following Occam's razor, we construct a model with minimal assumptions that provides new insight into the GRG phenomenon and the detection biases inherent to systematic search campaigns.

6.2.1 Intrinsic proper length

Firstly, we assume that giants and non-giant RGs share a common length distribution.[2] In particular, because power laws are abundant in Nature, we assume that the intrinsic proper length random variable (RV) L has a Pareto distribution with tail index $\xi < -1$ and support from $l_{\min} > 0$ onwards. If an RV is Pareto distributed, then the relative occurrence of two possible outcomes equals their ratio raised to a power: the tail index ξ. In astrophysics, Pareto distributions describe the kinetic energies of freshly accelerated electrons in large-scale structure and supernova shocks (e.g. Kirk & Schneider, 1987), the initial masses of main-sequence stars (e.g. Kroupa, 2001), and the luminosities of gamma-ray bursts (e.g. Bloom et al., 2001), to name a few examples. Previous works (e.g. Andernach et al., 2021) have already hinted at the approximate validity of a Pareto distribution description for GRG lengths. By comparing our final model to observations, as discussed in Section 6.4 and visualised in Fig. 6.14, we demonstrate that this assumption is indeed a powerful approximation in the current case.

The probability density function (PDF) $f_L : \mathbb{R} \to \mathbb{R}_{\geq 0}$ thus becomes

$$f_L(l) = \begin{cases} 0 & l < l_{\min}, \\ \frac{-(\xi+1)}{l_{\min}} \left(\frac{l}{l_{\min}}\right)^{\xi} & l \geq l_{\min}. \end{cases} \tag{6.1}$$

We refer the reader to Appendix 6.A1.1 for a derivation of this expression and a demonstration of its connection to the literature's most common parametrisation.

6.2.2 Projected proper length

From the distribution of intrinsic lengths and the assumption of random radio galaxy orientations, we now derive the distribution of projected lengths. This distribution

[2] Luckily, this assumption turns out to be irrelevant in the forthcoming GRG-only expressions, which are this work's focus.

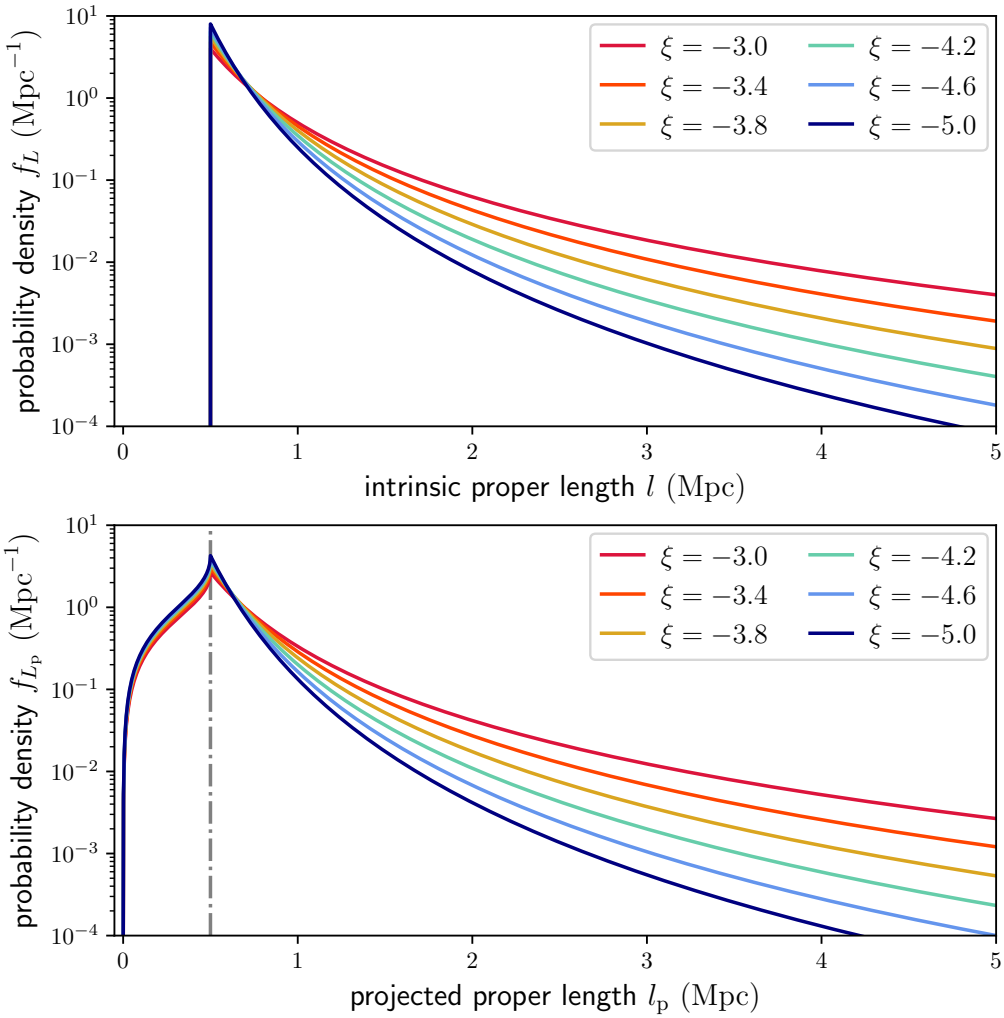


Figure 6.1: PDFs of radio galaxy intrinsic proper lengths L and projected proper lengths L_p. If the intrinsic lengths L are Pareto distributed above some cut-off l_{min}, then their projections on the sky L_p are also Pareto distributed above this cut-off. The tail indices are the same. We show the PDFs f_L *(top)* and f_{L_p} *(bottom)* for $l_{min} = 0.5$ Mpc, $l_{max} = \infty$ (see Appendix 6.A1) and various tail indices ξ. The support of f_L starts at l_{min}, which is marked by the vertical grey line in the bottom panel.

is more easily compared to observations, which usually lack inclination angle information.

Distribution for RGs

To model length and orientation, we consider a vector $\mathbf{L} \in \mathbb{R}^3$ (of length $L := ||\mathbf{L}||_2$) for each RG. In accordance with the IAU Solar System convention for positive poles,

the unit vector $\hat{L} \in \mathbb{S}^2$ marks the direction from which the central Kerr black hole is seen rotating in anticlockwise direction.[3] We define the inclination angle Θ as the angle between $\mathbf{L}$ and a vector parallel to the line of sight pointing towards the observer.[4] Observations that allow one to measure the orientation of the RG axis in 3D are time-intensive, and so usually only the RG length projected onto the plane of the sky is known.

Geometrically, we model RGs as line segments — as if they were 'thin sticks', with vanishing volumes — so that the projected proper length RV L_p relates to L and Θ through

$$L_\mathrm{p} = L \sin \Theta. \tag{6.2}$$

We assume that $\hat{L}$ is drawn from a uniform distribution on $\mathbb{S}^2$, so that the PDF $f_\Theta : [0, \pi] \to \mathbb{R}_{\geq 0}$ becomes $f_\Theta(\theta) = \frac{1}{2} \sin \theta$. Since L and Θ are independent, we find in Appendix 6.A1.2 that the PDF of L_p, $f_{L_\mathrm{p}} : \mathbb{R} \to \mathbb{R}_{\geq 0}$, is

$$f_{L_\mathrm{p}}(l_\mathrm{p}) = \begin{cases} 0 & \text{if } l_\mathrm{p} \leq 0, \\ \frac{-(\xi+1)}{l_\mathrm{min}} \frac{l_\mathrm{p}}{l_\mathrm{min}} I\left(\xi - 1, \frac{l_\mathrm{p}}{l_\mathrm{min}}\right) & \text{if } 0 < l_\mathrm{p} < l_\mathrm{min}, \\ \frac{(\xi+1)^2}{l_\mathrm{min}} \frac{\sqrt{\pi}}{4} \left(\frac{l_\mathrm{p}}{l_\mathrm{min}}\right)^{\xi} \frac{\Gamma\left(-\frac{\xi}{2} - \frac{1}{2}\right)}{\Gamma\left(-\frac{\xi}{2} + 1\right)} & \text{if } l_\mathrm{p} \geq l_\mathrm{min}, \end{cases} \tag{6.3}$$

where

$$I(a, b) := \int_1^\infty \frac{\eta^a \, \mathrm{d}\eta}{\sqrt{\eta^2 - b^2}} \text{ for } a < 0, |b| < 1. \tag{6.4}$$

We note that for $l_\mathrm{p} \geq l_\mathrm{min}$, the projected proper length has a Pareto distribution with the same tail index as the intrinsic proper length distribution. We compare f_L and f_{L_p} in Fig. 6.1.

Distribution for giants

For giants specifically (i.e. RGs such that $L_\mathrm{p} > l_\mathrm{p,GRG}$, where $l_\mathrm{p,GRG}$ is some constant threshold; in this work, $l_\mathrm{p,GRG} := 0.7\,\mathrm{Mpc}$), the projected proper length distribution

[3] $\mathbb{S}^2$ is the unit two-sphere.

[4] When $\Theta = 0$, the positive pole's jet points towards us, and the black hole is seen rotating in anticlockwise direction; when $\Theta = 90°$, the jets lie in the plane of the sky; when $\Theta = 180°$, the positive pole's jet points away from us, and the black hole rotates in clockwise direction.

becomes a Pareto distribution with tail index ξ again:

$$f_{L_\mathrm{p} \mid L_\mathrm{p} > l_\mathrm{p,GRG}}\left(l_\mathrm{p}\right) = \begin{cases} 0 & \text{if } l_\mathrm{p} < l_\mathrm{p,GRG} \\ \frac{-(\xi+1)}{l_\mathrm{p,GRG}} \left(\frac{l_\mathrm{p}}{l_\mathrm{p,GRG}}\right)^{\xi} & \text{if } l_\mathrm{p} \geq l_\mathrm{p,GRG}. \end{cases} \tag{6.5}$$

In other words, for giants, projection retains the Paretianity of lengths. A measurement of the tail index of the projected length distribution is immediately also a measurement of the tail index of the intrinsic length distribution.

The survival function, which gives the probability that a GRG has a projected proper length exceeding l_p, takes on a particularly simple form:

$$\mathbb{P}\left(L_\mathrm{p} > l_\mathrm{p} \mid L_\mathrm{p} > l_\mathrm{p,GRG}\right) = \left(\frac{l_\mathrm{p}}{l_\mathrm{p,GRG}}\right)^{\xi+1}. \tag{6.6}$$

The mean projected proper length of giants is the expectation value of $L_\mathrm{p} \mid L_\mathrm{p} > l_\mathrm{p,GRG}$:

$$\mathbb{E}[L_\mathrm{p} \mid L_\mathrm{p} > l_\mathrm{p,GRG}] = l_\mathrm{p,GRG} \frac{\xi+1}{\xi+2}, \tag{6.7}$$

which is only defined when $\xi < -2$. For example, when $\xi = -4$, $\mathbb{E}[L_\mathrm{p} \mid L_\mathrm{p} > l_\mathrm{p,GRG}] = \frac{3}{2} l_\mathrm{p,GRG}$, which becomes 1.05 Mpc for $l_\mathrm{p,GRG} := 0.7$ Mpc and 1.5 Mpc for $l_\mathrm{p,GRG} := 1$ Mpc.

Appendix 6.A1.2 provides a derivation for all three expressions.

6.2.3 Deprojection factor

The deprojection factor, $D := \frac{L}{L_\mathrm{p}} = \sin^{-1}\Theta$, quantifies how much larger intrinsic lengths are compared with projected lengths. The PDF of D, $f_D : \mathbb{R} \to \mathbb{R}_{\geq 0}$, is

$$f_D\left(d\right) = \begin{cases} 0 & \text{if } d \leq 1; \\ \frac{1}{d^2\sqrt{d^2-1}} & \text{if } d > 1. \end{cases} \tag{6.8}$$

The mean deprojection factor $\mathbb{E}[D] = \frac{\pi}{2}$. Deprojection factors can become arbitrarily large under the current model, because projected lengths can become arbitrarily small. As discussed in Sect. 6.5.2, this is not a very realistic set-up. In reality, an RG's projected length is bounded from below by its lobes, which have a non-vanishing volume and thus extend along all three spatial dimensions. Upon projection, the projected length therefore cannot shrink beyond some lower limit that depends on

the lobe geometry. In Appendix 6.A1.3, we show that by enriching the conventional stick-like geometry with spherical lobes, deprojection factors indeed become bounded.

6.2.4 Intrinsic proper length posterior and its moments

Because an RG's intrinsic length is more physically informative than its projected length, we ideally obtain the former. In this subsection, we quantify what a measurement $L_\mathrm{p} = l_\mathrm{p}$ already reveals about L.

We first note that the projected length bounds the intrinsic length from below. The intrinsic length can be much larger, however, but this is improbable for two reasons: large lengths are rarer than small lengths, although how drastic this effect is depends on ξ; in addition, viewing directions with large inclination angles are uncommon. The best we can do is to construct a posterior distribution for L given $L_\mathrm{p} = l_\mathrm{p}$. This posterior has a concise analytic form. If $l_\mathrm{p} > l_\mathrm{min}$, which is the relevant case for giants, the distribution of $L \mid L_\mathrm{p} = l_\mathrm{p}$ is

$$f_{L|L_\mathrm{p}=l_\mathrm{p}}(l) = \begin{cases} 0 & \text{if } l \leq l_\mathrm{p} \\ \frac{-\xi}{2^{1+\xi}\pi} \frac{\Gamma^2\left(-\frac{\xi}{2}\right)}{\Gamma(-\xi)} \frac{1}{l_\mathrm{p}} \frac{1}{\sqrt{\left(\frac{l}{l_\mathrm{p}}\right)^2 - 1}} \left(\frac{l}{l_\mathrm{p}}\right)^{\xi-1} & \text{if } l > l_\mathrm{p}. \end{cases} \tag{6.9}$$

For $l \gg l_\mathrm{p}$, $f_{L|L_\mathrm{p}=l_\mathrm{p}}(l) \propto \left(\frac{l}{l_\mathrm{p}}\right)^{\xi-2}$: the posterior PDF tends to a power law in l with index $\xi - 2$. In Fig. 6.2, we visualise the posterior PDF for several values of ξ. Clearly, to evaluate Eq. 6.9, one must choose l_p — however, the shape of the distribution is the same for all choices. We illustrate this by comparing the PDF for a comparatively small GRG ($l_\mathrm{p} = 1.0$ Mpc; top panel) to the PDF for Alcyoneus[5] ($l_\mathrm{p} = 5.0$ Mpc; bottom panel).

The posterior mean is

$$\mathbb{E}\left[L \mid L_\mathrm{p} = l_\mathrm{p}\right] = l_\mathrm{p} \cdot \frac{-\xi}{2^{2\xi+3}\pi} \frac{\Gamma^4\left(-\frac{\xi}{2}\right)}{\Gamma^2(-\xi)}; \tag{6.10}$$

[5] Alcyoneus is the projectively longest giant known to date (Oei et al., 2022a).

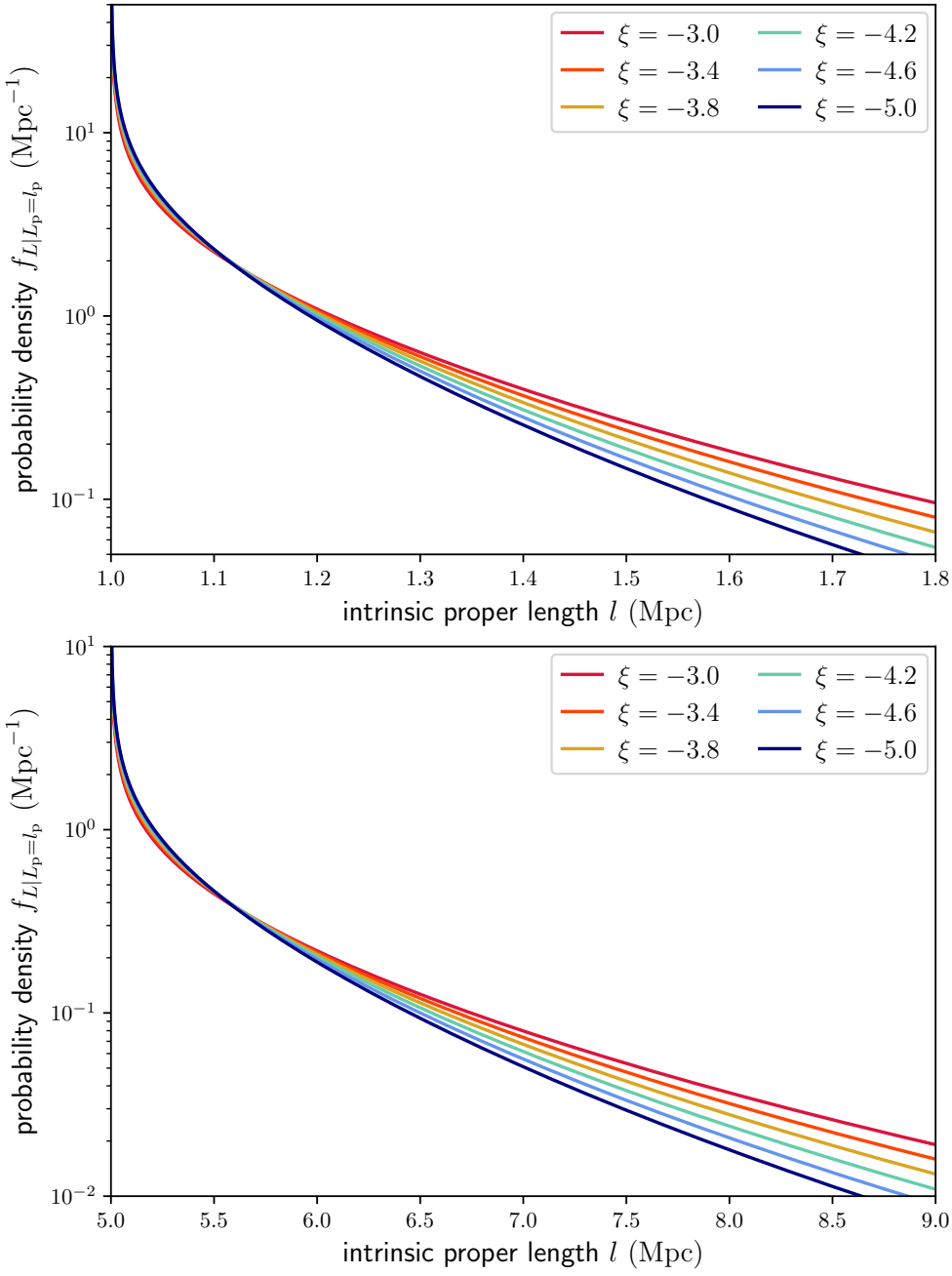


Figure 6.2: Posterior PDFs of intrinsic lengths for a given projected length $L \mid L_p = l_p$. If tail index ξ is known, then an RG's l_p fixes the probability distribution over its possible l. This distribution is strongly skewed, and the same for all l_p — save for horizontal translation and vertical scaling. We illustrate this point by showing posterior PDFs for giants with two different l_p. *Top:* $l_p = 1$ Mpc. *Bottom:* $l_p = 5$ Mpc. For $\xi = -4$, the posterior mean $\mathbb{E}[L \mid L_p = l_p] = 1.13\, l_p$ and the posterior standard deviation $\sqrt{\mathbb{V}[L \mid L_p = l_p]} = 0.23\, l_p$ (see Table 6.1).

the posterior variance is

$$\mathbb{V}\left[L \mid L_{\mathrm{p}} = l_{\mathrm{p}}\right] = l_{\mathrm{p}}^2 \left(\frac{\xi}{\xi+1} - \frac{\xi^2}{2^{4\xi+6}\pi^2} \frac{\Gamma^8\left(-\frac{\xi}{2}\right)}{\Gamma^4\left(-\xi\right)} \right). \tag{6.11}$$

Both mean and standard deviation scale linearly in l_{p}: the projection effect is a multiplicative noise source. In Table 6.1, we provide explicit values for various ξ.

Table 6.1: Intrinsic proper length posterior mean and standard deviation in multiples of projected proper length l_{p}, given for various tail indices ξ.

	$\mathbb{E}\left[L \mid L_{\mathrm{p}} = l_{\mathrm{p}}\right](\xi)\ \ (l_{\mathrm{p}})$	$\sqrt{\mathbb{V}\left[L \mid L_{\mathrm{p}} = l_{\mathrm{p}}\right](\xi)}\ \ (l_{\mathrm{p}})$
$\xi = -2$	$\frac{4}{\pi} \approx 1.27$	$\sqrt{2 - \frac{16}{\pi^2}} \approx 0.62$
$\xi = -3$	$\frac{3\pi}{8} \approx 1.18$	$\sqrt{\frac{3}{2} - \frac{9\pi^2}{64}} \approx 0.33$
$\xi = -4$	$\frac{32}{9\pi} \approx 1.13$	$\sqrt{\frac{4}{3} - \frac{1024}{81\pi^2}} \approx 0.23$
$\xi = -5$	$\frac{45\pi}{128} \approx 1.10$	$\sqrt{\frac{5}{4} - \frac{2025\pi^2}{16384}} \approx 0.17$

Higher moments exist up to order $\lceil -\xi \rceil$; because the PDF $f_{L \mid L_{\mathrm{p}} = l_{\mathrm{p}}}(l)$ is strongly skewed, such moments do further specify the distribution.

It is important to note that it is formally incorrect to statistically deproject RGs by drawing samples from deprojection factor D and multiplying them with some measurement $L_{\mathrm{p}} = l_{\mathrm{p}}$, or even more crudely, by multiplying the latter with $\mathbb{E}[D]$. The reason that renders such approaches invalid is that $L_{\mathrm{p}} = L \sin\Theta$ and $D = \sin^{-1}\Theta$ are not independent RVs. We refer to Appendix 6.A1.4 for an explicit proof of this fact, and for derivations of this subsection's expressions.

6.2.5 GRG inclination angle

Radio galaxies with jets that make a small angle with the plane of the sky are more likely to have a projected length exceeding $l_{\mathrm{p,GRG}}$ than those with jets that make a large angle with the plane of the sky. For this reason, the inclination angle distribution of giants is different from that of RGs: it is more peaked around $\theta = 90°$. More precisely, the PDF $f_{\Theta \mid L_{\mathrm{p}} > l_{\mathrm{p,GRG}}} : [0, \pi] \to \mathbb{R}_{\geq 0}$ of the GRG inclination angle

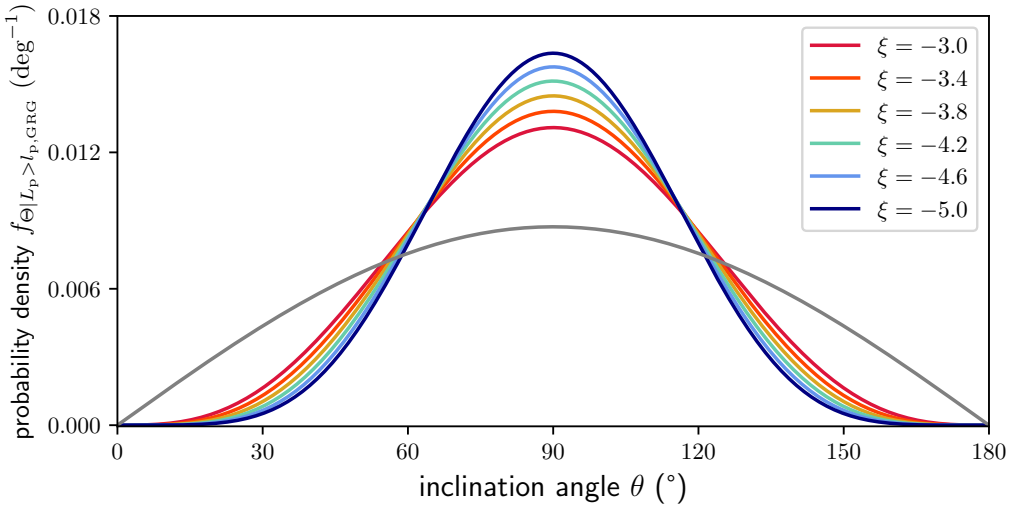


Figure 6.3: PDFs of GRG inclination angles $\Theta \mid L_p > l_{p,GRG}$ (*colours*) and RG inclination angles Θ (*grey*). Giants more often have orientations close to the sky plane. The strength of this tendency is governed by ξ, with larger ξ meaning more dispersion.

$\Theta \mid L_p > l_{p,GRG}$ has the general form

$$f_{\Theta \mid L_p > l_{p,GRG}}(\theta) = \frac{\left(1 - F_L\left(\frac{l_{p,GRG}}{\sin\theta}\right)\right) f_\Theta(\theta)}{1 - F_{L_p}(l_{p,GRG})}. \tag{6.12}$$

Under our Pareto distribution assumption for L, this concretises to

$$f_{\Theta \mid L_p > l_{p,GRG}}(\theta) = \frac{2}{-(\xi+1)\sqrt{\pi}} \frac{\Gamma\left(-\frac{\xi}{2}+1\right)}{\Gamma\left(-\frac{\xi}{2}-\frac{1}{2}\right)} \sin^{-\xi}\theta. \tag{6.13}$$

We note that $f_{\Theta \mid L_p > l_{p,GRG}}(\theta) \propto \sin^{-\xi}\theta$; the factor in front serves only as a normalisation constant. The distribution is independent of the choice of $l_{p,GRG}$ and depends on a single parameter: ξ. We visualise $f_{\Theta \mid L_p > l_{p,GRG}}(\theta)$ in Fig. 6.3. Appendix 6.A1.5 contains a brief derivation.

6.2.6 GRG angular length

The model predicts the distribution of GRG angular lengths in the Local Universe up to comoving distance r_{max}. The GRG angular length RV $\Phi \mid L_p > l_{p,GRG}$ relates to the GRG projected proper length RV $L_p \mid L_p > l_{p,GRG}$ and the comoving distance

RV R as

$$\Phi \mid L_\mathrm{p} > l_\mathrm{p,GRG} = L_\mathrm{p} \mid L_\mathrm{p} > l_\mathrm{p,GRG} \cdot \frac{1 + z(R)}{R}. \tag{6.14}$$

(We note that this relation is valid only in a *flat* Friedmann–Lemaître–Robertson–Walker (FLRW) universe.) We also assume that the GRG number density is constant in the Local Universe. The PDF of $\Phi \mid L_\mathrm{p} > l_\mathrm{p,GRG}$ has useful analytic forms under two different idealisations.

In a Euclidean universe, $z(R) = 0$, and the minimal GRG angular length $\varphi_\mathrm{GRG} = \frac{l_\mathrm{p,GRG}}{r_\mathrm{max}}$. Then

$$f_{\Phi \mid L_\mathrm{p} > l_\mathrm{p,GRG}}(\varphi) = \begin{cases} 0 & \text{if } \varphi \leq \varphi_\mathrm{GRG} \\ -3\frac{\xi+1}{\xi+4} \cdot \frac{1}{\varphi_\mathrm{GRG}} \left(\left(\frac{\varphi}{\varphi_\mathrm{GRG}}\right)^{\xi} - \left(\frac{\varphi}{\varphi_\mathrm{GRG}}\right)^{-4} \right) & \text{if } \varphi > \varphi_\mathrm{GRG}, \end{cases} \tag{6.15}$$

which is valid as long as $\xi \neq -4$.

In an expanding universe at low redshifts, the Hubble–Lemaître law $z(R) \approx \frac{R}{d_H}$ holds; the Hubble distance $d_H := \frac{c}{H_0}$. In this case,

$$f_{\Phi \mid L_\mathrm{p} > l_\mathrm{p,GRG}}(\varphi) = \begin{cases} 0 & \text{if } \varphi \leq \varphi_\mathrm{GRG} + \frac{l_\mathrm{p,GRG}}{d_H} \\ \frac{-3(\xi+1)\varphi^{\xi}}{r_\mathrm{max}^3 l_\mathrm{p,GRG}^{\xi+1}} \int_{\frac{1}{r_\mathrm{max}}+\frac{1}{d_H}}^{\frac{\varphi}{l_\mathrm{p,GRG}}} \frac{\mathrm{d}k}{k^{\xi+1}\left(k - \frac{1}{d_H}\right)^4} & \text{if } \varphi > \varphi_\mathrm{GRG} + \frac{l_\mathrm{p,GRG}}{d_H}. \end{cases} \tag{6.16}$$

Figure 6.4 shows GRG angular length PDFs under both idealisations. The PDFs undergo a minor shift upon changing universe type but are otherwise similar. For most current-day applications, it will therefore be unnecessary to calculate an even more refined version of $f_{\Phi \mid L_\mathrm{p} > l_\mathrm{p,GRG}}(\varphi)$. Appendix 6.A1.6 contains derivations and details.

6.2.7 Maximum likelihood estimation of the tail index

The GRG projected proper length distribution features just one parameter of physical interest: the tail index ξ. If observational selection effects are negligible, one can directly use maximum likelihood estimation (MLE) on GRG data to infer ξ. In particular, we consider a set of projected lengths $\{L_{\mathrm{p},1}, \ldots, L_{\mathrm{p},N}\} \sim L_\mathrm{p} \mid L_\mathrm{p} > l_\mathrm{p,GRG}$ from N giants. Appendix 6.A1.7 shows that the maximum likelihood estimate of ξ is

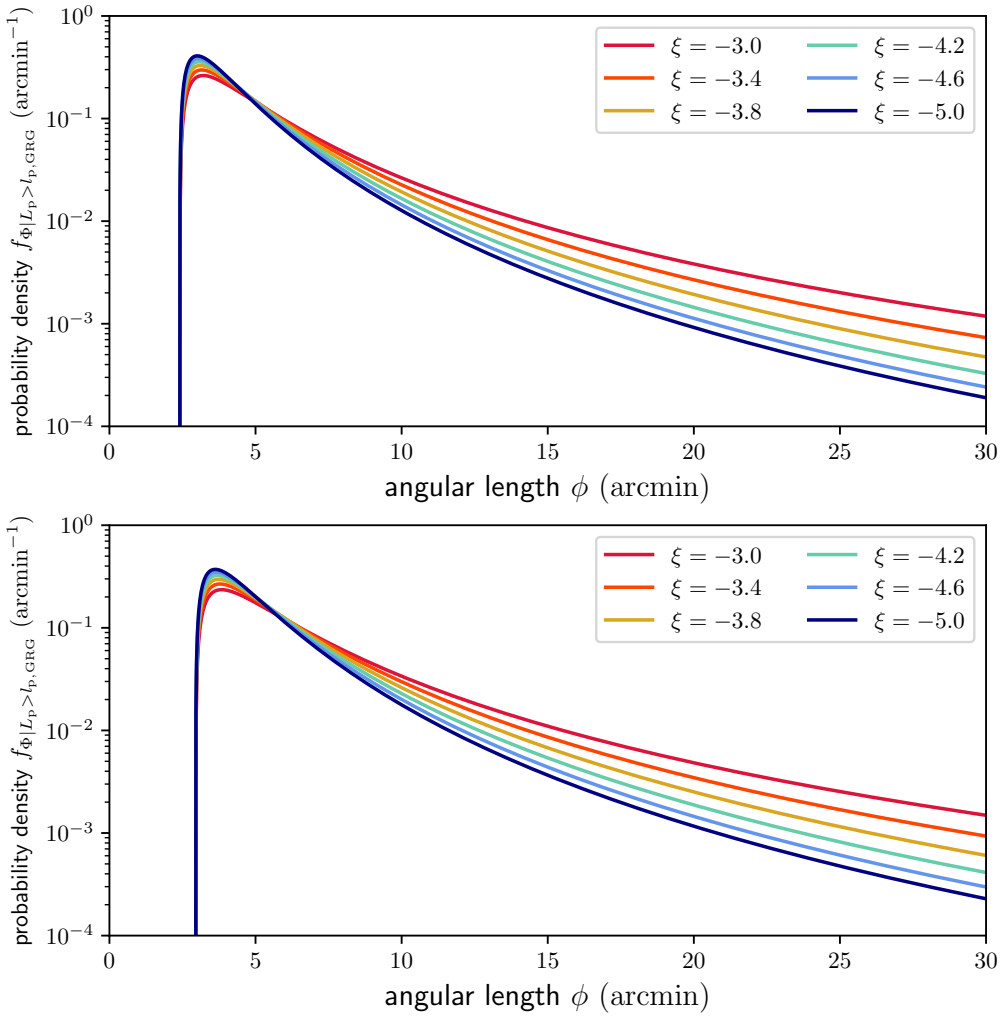


Figure 6.4: PDFs of GRG angular lengths $\Phi \mid L_p > l_{p,GRG}$. We fix $r_{max} = 1$ Gpc (and $l_{p,GRG} = 0.7$ Mpc), and vary ξ. *Top:* Euclidean universe. *Bottom:* expanding universe at low redshifts.

the RV ξ_{MLE}, given by

$$\xi_{MLE} = -\frac{N}{\sum_{i=1}^{N} \ln \frac{L_{p,i}}{l_{p,GRG}}} - 1. \tag{6.17}$$

6.2.8 Observed projected proper length

General considerations

In the preceding theory, we have ignored observational selection effects that favour some projected proper lengths over others. In practice, several such effects occur; the importance of each varies per survey and (G)RG search campaign within it. One of them is the bias against physically long RGs that the interferometer's largest detectable angular scale can induce.[6] As a result, the projected proper length of an observed RG might not be adequately modelled through RV L_{p}. Instead, we must introduce a new RV $L_{\mathrm{p,obs}}$.

We define the completeness $C : \mathbb{R}_{>0} \times \mathbb{R}_{>0} \to [0, 1]$ at $(l_{\mathrm{p}}, z_{\mathrm{max}})$ to be the fraction of all RGs with projected proper length l_{p} in the cosmological volume up to z_{max} that is detected in a particular RG search campaign. Then, assuming that the distribution of L_{p} does not evolve with redshift between $z = z_{\mathrm{max}}$ and $z = 0$ (i.e. ξ remains constant),

$$C\left(l_{\mathrm{p}}, z_{\mathrm{max}}\right) = \frac{\int_0^{z_{\mathrm{max}}} p_{\mathrm{obs}}\left(l_{\mathrm{p}}, z\right) r^2\left(z\right) E^{-1}\left(z\right)\, \mathrm{d}z}{\int_0^{z_{\mathrm{max}}} r^2\left(z\right) E^{-1}\left(z\right)\, \mathrm{d}z}, \tag{6.18}$$

where $p_{\mathrm{obs}}(l_{\mathrm{p}}, z)$ is the probability that an RG of projected proper length l_{p} at cosmological redshift z is detected through the campaign, and $r(z)$ is the comoving radial distance at cosmological redshift z. In a *flat* FLRW universe, the dimensionless Hubble parameter E is

$$E\left(z\right) := \frac{H\left(z\right)}{H_0} = \sqrt{\Omega_{\mathrm{R},0}\left(1+z\right)^4 + \Omega_{\mathrm{M},0}\left(1+z\right)^3 + \Omega_{\Lambda,0}}. \tag{6.19}$$

The PDF of the observed projected proper length RV $L_{\mathrm{p,obs}}$ becomes

$$f_{L_{\mathrm{p,obs}}}\left(l_{\mathrm{p}}\right) = \frac{C\left(l_{\mathrm{p}}\right) f_{L_{\mathrm{p}}}\left(l_{\mathrm{p}}\right)}{\int_0^{\infty} C\left(l'_{\mathrm{p}}\right) f_{L_{\mathrm{p}}}\left(l'_{\mathrm{p}}\right)\, \mathrm{d}l'_{\mathrm{p}}}, \tag{6.20}$$

where we suppress the z_{max}-dependence for succinctness. We note that multiplying

[6] For example, the Faint Images of the Radio Sky at Twenty Centimeters (FIRST) survey used the Very Large Array (VLA) in B-configuration, leading it to detect angular scales of at most two arcminutes. By contrast, the largest angular scale of the LoTSS — the survey relevant to this work — is about a degree. (For the 6″ and 20″ resolutions, the shortest baseline is 100 metres; for the 60″ and 90″ resolutions, the shortest baseline is 68 metres.) As virtually all giants are of subdegree angular length, we need not consider this bias in our case.

$p_{\mathrm{obs}}(l_{\mathrm{p}}, z)$ with an l_{p}- and z-independent factor affects the completeness $C(l_{\mathrm{p}}, z_{\mathrm{max}})$, but cancels in Eq. 6.20; $f_{L_{\mathrm{p,obs}}}$ will be independent of it. Finally, the PDF of the GRG observed projected proper length RV $L_{\mathrm{p,obs}} \mid L_{\mathrm{p,obs}} > l_{\mathrm{p,GRG}}$ is

$$f_{L_{\mathrm{p,obs}} \mid L_{\mathrm{p,obs}} > l_{\mathrm{p,GRG}}}\left(l_{\mathrm{p}}\right) = \begin{cases} 0 & \text{if } l_{\mathrm{p}} \leq l_{\mathrm{p,GRG}} \\ \frac{C\left(l_{\mathrm{p}}\right) f_{L_{\mathrm{p}}}\left(l_{\mathrm{p}}\right)}{\int_{l_{\mathrm{p,GRG}}}^{\infty} C\left(l'_{\mathrm{p}}\right) f_{L_{\mathrm{p}}}\left(l'_{\mathrm{p}}\right) \mathrm{d}l'_{\mathrm{p}}} & \text{if } l_{\mathrm{p}} > l_{\mathrm{p,GRG}}. \end{cases} \tag{6.21}$$

We derive these expressions in Appendix 6.A1.8.

Fuzzy angular length threshold

We provide a concrete example of an important observational selection effect in visual searches for GRG candidates in survey images. To cope with the sheer number of detectable RGs in modern surveys like the LoTSS, a natural criterion is to only add sources to a candidate list if they appear — by eye — to have an angular length larger than some threshold. However, it is hard to precisely assess the angular length of a candidate before actually measuring it; sometimes, a candidate with a smaller angular length than the threshold will feature in the list, while some candidates with a larger angular length than the threshold will not. This leads to the notion of a 'fuzzy angular length threshold', where the probability that an RG with angular length φ is observed through the visual search increases (e.g. linearly) from 0 to 1 between φ_{min} and φ_{max}:

$$p_{\mathrm{obs,AL}}\left(l_{\mathrm{p}}, z\right) = \min\left\{\max\left\{\frac{\varphi\left(l_{\mathrm{p}}, z\right) - \varphi_{\mathrm{min}}}{\varphi_{\mathrm{max}} - \varphi_{\mathrm{min}}}, 0\right\}, 1\right\}, \tag{6.22}$$

$$\varphi\left(l_{\mathrm{p}}, z\right) = \frac{l_{\mathrm{p}}\left(1+z\right)}{r\left(z\right)}. \tag{6.23}$$

See the top row of Fig. 6.5 for several examples of associated completeness curves $C(l_{\mathrm{p}})$ and observed projected proper length PDFs. See Appendix 6.A1.8 for additional information.

Surface brightness limitations

Another important observational selection effect is due to a survey's finite noise level. The noise determines the surface brightness threshold $b_{\nu,\mathrm{th}}$ (typically comparable to the noise level itself) below which radio galaxy features remain visually undetected.

Fanaroff–Riley class II We model the lobe surface brightness RV B_ν at the central observing frequency ν_{obs} as

$$B_\nu = \frac{b_{\nu,\mathrm{ref}} \cdot S}{(1+Z)^{3-\alpha}} \left(\frac{L}{l_{\mathrm{ref}}}\right)^{\zeta}, \tag{6.24}$$

where $b_{\nu,\mathrm{ref}}$ is the median surface brightness of RGs of intrinsic proper length l_{ref} at cosmological redshift $z = 0$ and frequency ν_{obs}, and S is a lognormally distributed RV with median 1 and dispersion parameter σ_{ref} that captures the variability in surface brightness among this population of RGs. The denominator models the fact that surface brightness is not conserved with distance in an FLRW universe; Z is the cosmological redshift RV up to $z = z_{\mathrm{max}}$ and α is the typical lobe spectral index. The exponent $\zeta < 0$ characterises the scaling between intrinsic proper length and surface brightness. (If RGs are self-similar, so that morphology does not predict length, one finds $\zeta = -2$.) For this selection effect, the observing probability is

$$p_{\mathrm{obs,SB}}\left(l_{\mathrm{p}}, z\right) = \int_{s_{\mathrm{min}}}^{\infty} \sqrt{1 - \left(\frac{s_{\mathrm{min}}}{s}\right)^{-\frac{2}{\zeta}}} f_S\left(s\right) \mathrm{d}s, \tag{6.25}$$

$$s_{\mathrm{min}} = \frac{b_{\nu,\mathrm{th}}}{b_{\nu,\mathrm{ref}}} \left(\frac{l_{\mathrm{p}}}{l_{\mathrm{ref}}}\right)^{-\zeta} (1+z)^{3-\alpha}, \tag{6.26}$$

$$f_S\left(s\right) = \frac{1}{\sqrt{2\pi}\sigma_{\mathrm{ref}} s} \exp\left(-\frac{\ln^2 s}{2\sigma_{\mathrm{ref}}^2}\right). \tag{6.27}$$

We note that $p_{\mathrm{obs,SB}}$ does not depend on $b_{\nu,\mathrm{th}}$ or $b_{\nu,\mathrm{ref}}$ separately, but on their ratio only. See the bottom row of Fig. 6.5 for several examples of associated completeness curves $C(l_{\mathrm{p}})$ and observed projected proper length PDFs.

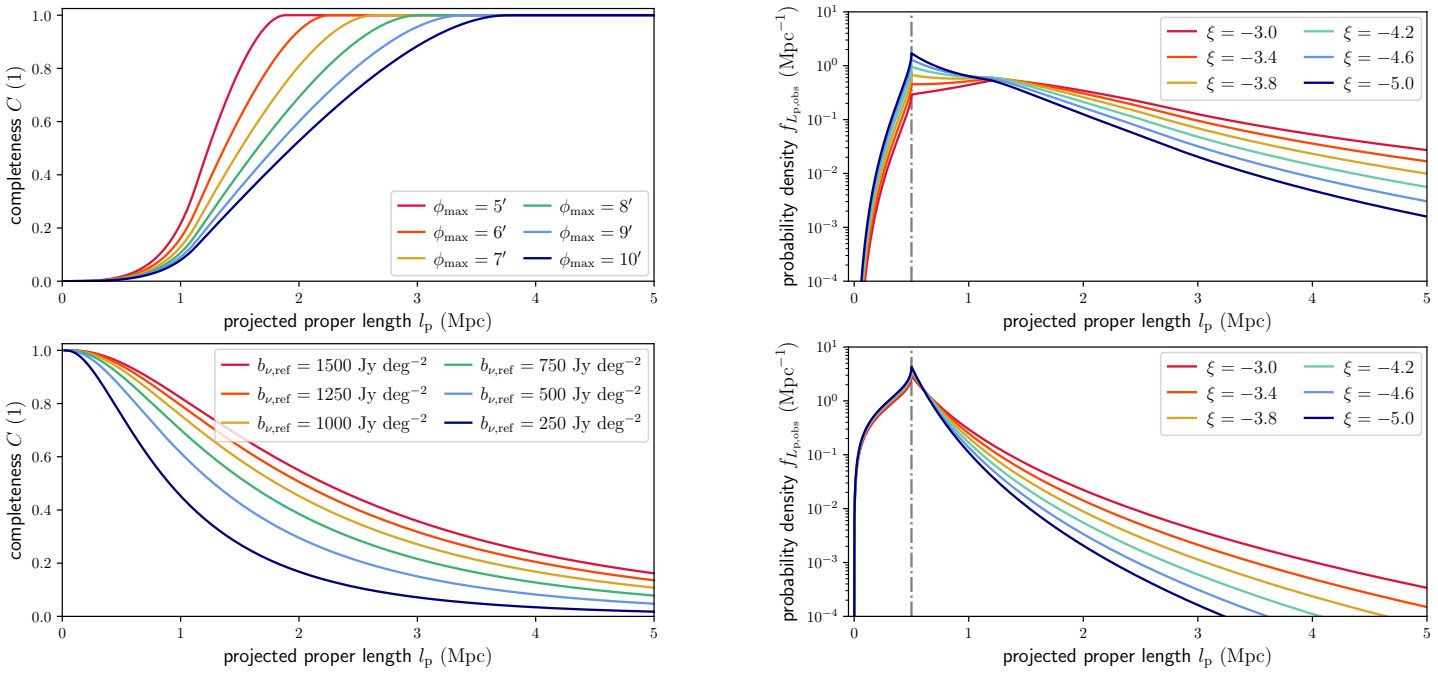


Figure 6.5: Completeness functions (*left column*) and PDFs of observed projected proper lengths $L_\mathrm{p,obs}$ (*right column*). Selection effects leave imprints on the distribution of radio galaxies' $L_\mathrm{p,obs}$. In the top row, we show how imposing an angular length threshold in a GRG search campaign leads to incompleteness *(left)*, which causes the PDF $f_{L_\mathrm{p,obs}}$ *(right)* to differ from f_{L_p}. We assume RGs with angular length $\varphi < \varphi_\mathrm{min}$ have probability 0 to be included in a sample, whilst RGs with angular length $\varphi > \varphi_\mathrm{max}$ have probability 1. The inclusion probability is assumed to increase linearly between φ_min and φ_max. In the left panel, we fix $\varphi_\mathrm{min} = 3'$ and vary φ_max; in the right panel, we also fix $\varphi_\mathrm{max} = 7'$. In the bottom row, we show how a survey's surface brightness limitations lead to incompleteness *(left)*, which causes the PDF $f_{L_\mathrm{p,obs}}$ *(right)* to differ from f_{L_p}. We assume that lobe surface brightnesses are lognormally distributed; we parametrise the distribution for RGs of intrinsic length $l_\mathrm{ref} = 0.7$ Mpc at $z = 0$ observed at $\nu_\mathrm{obs} = 144$ MHz with a median $b_{\nu,\mathrm{ref}}$ and dispersion parameter σ_ref. In the left panel, we fix $\sigma_\mathrm{ref} = 1.5$ and vary $b_{\nu,\mathrm{ref}}$; in the right panel, we also fix $b_{\nu,\mathrm{ref}} = 1000$ Jy deg^{-2}. We assume a lobe spectral index $\alpha = -1$, a surface brightness detection threshold $b_{\nu,\mathrm{th}} = 25$ Jy deg^{-2}, and self-similar growth: $\zeta = -2$. For both selection effects, we consider RGs up to cosmological redshift $z_\mathrm{max} = 0.5$ only.

Fanaroff–Riley class I For FRI RGs, the assumption of a constant surface brightness beyond the core is inaccurate. The simplest correction in which FRI RGs retain a well-defined notion of length assumes a linearly decreasing surface brightness profile, which peaks at the core and goes to zero at the RG's two endpoints. (A power-law profile does not work: in such case, the surface brightness only goes asymptotically to zero — but never actually reaches it.) In this case, we define RV B_ν to be the mean surface brightness along an RG's jets, which can be regarded as a typical value for that RG. As B_ν again obeys Eq. 6.24, we find that the formulaic structure of $p_{\mathrm{obs,SB}}(l_\mathrm{p}, z)$ is identical for FRI and FRII giants, except that FRI giants require a change

$$b_{\nu,\mathrm{th}} \to \frac{b_{\nu,\mathrm{th}}}{2\left(1 - \frac{l_{\mathrm{p,GRG}}}{l_\mathrm{p}}\right)}, \tag{6.28}$$

which affects $p_{\mathrm{obs,SB}}(l_\mathrm{p}, z)$ through s_{min}. There is no change for $l_\mathrm{p} = 2l_{\mathrm{p,GRG}}$. Although the formulaic structure might be the same, the best-fit parameters can differ. For example, it is possible that $\sigma_{\mathrm{ref,FRI}} \neq \sigma_{\mathrm{ref,FRII}}$ or $\zeta_{\mathrm{FRI}} \neq \zeta_{\mathrm{FRII}}$. See Appendix 6.A1.8 for derivations and numerical implementation considerations.

To include both aforementioned selection effects at the same time, a natural approximation is to assume that the observing probability simply factorises:

$$p_{\mathrm{obs}}\left(l_\mathrm{p}, z\right) \approx p_{\mathrm{obs,AL}}\left(l_\mathrm{p}, z\right) \cdot p_{\mathrm{obs,SB}}\left(l_\mathrm{p}, z\right). \tag{6.29}$$

6.2.9 GRG number density

A central question in the field of radio galaxies is: how intrinsically rare are giants? By counting giants in a search campaign with well-understood selection effects, we can give an answer. More precisely, one can estimate the comoving number density of giants in the Local Universe, n_{GRG}, through the number of observed giants up to cosmological redshift z_{max} in a uniformly searched region of sky of solid angle Ω, $N_{\mathrm{GRG,obs}}\left(z_{\mathrm{max}}, \Omega\right)$. Then, under the standard assumption $l_{\mathrm{p,GRG}} > l_{\mathrm{min}}$, Appendix 6.A1.9 shows that

$$n_{\mathrm{GRG}} = \frac{l_{\mathrm{p,GRG}}^{\xi+1}}{-(\xi+1)} \cdot \frac{H_0}{c} \cdot \frac{\frac{4\pi}{\Omega} N_{\mathrm{GRG,obs}}\left(z_{\mathrm{max}}, \Omega\right)}{\int_{l_{\mathrm{p,GRG}}}^{\infty} l_\mathrm{p}^{\xi} \int_0^{z_{\mathrm{max}}} p_{\mathrm{obs}}\left(l_\mathrm{p}, z\right) 4\pi r^2\left(z\right) E^{-1}\left(z\right)\, \mathrm{d}z\, \mathrm{d}l_\mathrm{p}}. \tag{6.30}$$

Although the appropriate $p_{\rm obs}(l_{\rm p}, z)$ varies per search campaign, it is always possible to bound $p_{\rm obs}(l_{\rm p}, z)$ from above — for example by 1. In such case, Eq. 6.30 bounds $n_{\rm GRG}$ from below.

6.2.10 GRG lobe volume-filling fraction

Because giants attain cosmological lengths, they might contribute to the energisation and magnetisation of Cosmic Web filaments in regions that smaller radio galaxies cannot reach. A key statistic that measures the enrichment of the Cosmic Web by giants is the volume-filling fraction (VFF) of their lobes. Assuming that lobes do not grow along with the expansion of the Universe, the proper VFF changes over cosmic time: $\mathrm{VFF}_{\rm GRG}(z) = \mathrm{VFF}_{\rm GRG}(z=0)\cdot(1+z)^3$, where

$$\begin{aligned}\mathrm{VFF}_{\rm GRG}(z=0) &:= n_{\rm GRG}\cdot\mathbb{E}[V\,|\,L_{\rm p} > l_{\rm p,GRG}]\\ &= n_{\rm GRG}\cdot\mathbb{E}[\Upsilon\cdot L^3\,|\,L_{\rm p} > l_{\rm p,GRG}],\end{aligned} \tag{6.31}$$

where V is the combined volume of the lobes and $\Upsilon := \frac{V}{L^3}$ is a dimensionless RV that captures the diversity in radio galaxy lobe shapes. We find under self-similar growth

$$\mathrm{VFF}_{\rm GRG}(z=0) = n_{\rm GRG}\cdot\mathbb{E}[\Upsilon]\cdot\mathbb{E}[L^3\,|\,L_{\rm p} > l_{\rm p,GRG}]. \tag{6.32}$$

$\mathbb{E}[\Upsilon]$ can be estimated from observations, but one must be wary of selection effects. Unfortunately, $\mathbb{E}[L^3\,|\,L_{\rm p} > l_{\rm p,GRG}]$ does not exist for $\xi \geq -4$, which is the regime supported by observations. A useful lower bound then is

$$\mathrm{VFF}_{\rm GRG}(z=0) > n_{\rm GRG}\cdot\mathbb{E}[\Upsilon]\cdot\mathbb{E}^3[L\,|\,L_{\rm p} > l_{\rm p,GRG}], \tag{6.33}$$

where

$$\mathbb{E}[L\,|\,L_{\rm p} > l_{\rm p,GRG}] = l_{\rm p,GRG}\frac{\Gamma\left(-\frac{\xi}{2}-1\right)\Gamma\left(-\frac{\xi}{2}+1\right)}{\Gamma\left(-\frac{\xi}{2}-\frac{1}{2}\right)\Gamma\left(-\frac{\xi}{2}+\frac{1}{2}\right)}. \tag{6.34}$$

This is the mean intrinsic proper length of giants, which is only defined when $\xi < -2$.[7]

An alternative is to deviate slightly from our Pareto ansatz and truncate the GRG

[7] For example, when $\xi = -3$, $\mathbb{E}[L\,|\,L_{\rm p} > l_{\rm p,GRG}] = \frac{3\pi}{4}l_{\rm p,GRG}$, which becomes 1.65 Mpc for $l_{\rm p,GRG} := 0.7$ Mpc and 2.36 Mpc for $l_{\rm p,GRG} := 1$ Mpc. When $\xi = -4$, $\mathbb{E}[L\,|\,L_{\rm p} > l_{\rm p,GRG}] = \frac{16}{3\pi}l_{\rm p,GRG}$, which becomes 1.19 Mpc for $l_{\rm p,GRG} := 0.7$ Mpc and 1.70 Mpc for $l_{\rm p,GRG} := 1$ Mpc.

projected proper length distribution at some $l_{\mathrm{p,max}}$; then

$$\mathrm{VFF}_{\mathrm{GRG}}(z=0) = n_{\mathrm{GRG}} \cdot \mathbb{E}[\Upsilon_{\mathrm{p}}] \cdot \frac{\xi+1}{\xi+4} \cdot \frac{l_{\mathrm{p,max}}^{\xi+4} - l_{\mathrm{p,GRG}}^{\xi+4}}{l_{\mathrm{p,max}}^{\xi+1} - l_{\mathrm{p,GRG}}^{\xi+1}}, \tag{6.35}$$

where $\Upsilon_{\mathrm{p}} := \frac{V}{L_{\mathrm{p}}^3}$. See Appendix 6.A1.10 for a derivation and further details.

6.2.11 Unification model constraints from quasar and non-quasar giants

The unification model and its extensions (e.g. Hardcastle & Croston, 2020) form an elegant family of hypotheses that aim to explain the observational diversity of active galaxies. It posits that active galaxies with quasars differ from those without quasars primarily because of differences in orientation of the dusty tori surrounding SMBHs. In particular, the central idea is that a quasar appears brighter to the observer than a non-quasar AGN because the axis of its dusty torus happens to be virtually parallel to the line of sight. As such, only quasars would offer an unobscured view of the luminous accretion disk surrounding the SMBH, whilst also beaming relativistic jet emission towards the observer. Using our statistical framework, we predict the general ramifications of the basic unification model on a GRG sample.

The basic unification model suggests to divide the radio galaxy population in two, distinguishing between RGs generated by AGN with quasar appearance (quasar RGs) and RGs generated by AGN without quasar appearance (non-quasar RGs). We assume that quasar RGs have inclination angles $\theta \leq \theta_{\mathrm{max}}$ or $\theta \geq 180^\circ - \theta_{\mathrm{max}}$ whilst non-quasar RGs have $\theta_{\mathrm{max}} < \theta < 180^\circ - \theta_{\mathrm{max}}$.[8] If quasar RGs are more closely aligned with the line of sight than non-quasar RGs but are otherwise similar, fewer of them will satisfy $l_{\mathrm{p}} \geq l_{\mathrm{p,GRG}}$ and thus be classified as giants. Therefore, the quasar GRG fraction f_{Q} — the fraction of quasar giants in an actual GRG sample — constrains θ_{max}. We model f_{Q} as an RV:

$$f_{\mathrm{Q}} := \frac{N_{\mathrm{Q}}}{N}; \quad N_{\mathrm{Q}} \sim \mathrm{Binom}\left(N, p_{\mathrm{Q}}\right), \tag{6.36}$$

where the RV N_{Q} is the number of quasar giants in the sample, the constant N is the total number of giants in the sample and the parameter p_{Q} is the quasar GRG

[8] Geometrically, θ_{max} represents the opening angle of the two coaxial conical gaps in the dusty torus of the AGN.

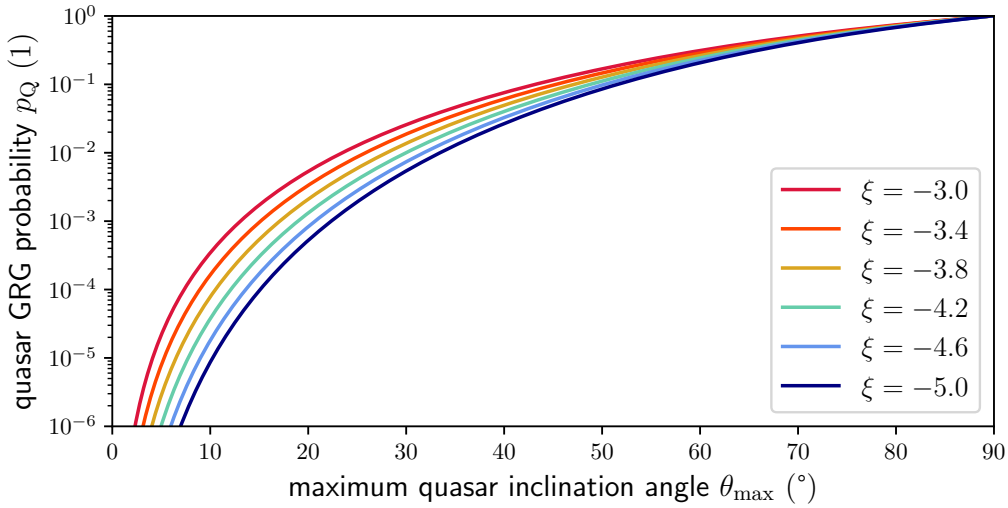


Figure 6.6: Probability p_Q that an observed giant is a quasar giant under the unification model. Under this model, giants generated by AGN with quasar appearance (quasar giants) have inclination angles $\theta \leq \theta_{max}$ or $\theta \geq 180° - \theta_{max}$ and giants generated by AGN without quasar appearance (non-quasar giants) have intermediate θ. As long as quasar giants and non-quasar giants are subject to the same selection effects, these selection effects do not affect p_Q. Instead, in such case, p_Q only depends on the maximum quasar inclination angle θ_{max} and the tail index ξ.

probability. Our framework predicts

$$\begin{aligned} p_Q &:= \frac{\mathbb{P}\left(L_{p,obs} \geq l_{p,GRG},\ \sin\Theta \leq \sin\theta_{max}\right)}{\mathbb{P}\left(L_{p,obs} \geq l_{p,GRG}\right)} \\ &= \frac{4\Gamma\left(-\frac{\xi}{2}+1\right)}{-(\xi+1)\sqrt{\pi}\,\Gamma\left(-\frac{\xi}{2}-\frac{1}{2}\right)} \int_0^{\sin\theta_{max}} \frac{x^{-\xi}\,\mathrm{d}x}{\sqrt{1-x^2}}. \end{aligned} \tag{6.37}$$

See Appendix 6.A1.11 for a derivation. Interestingly, as long as quasar giants and non-quasar giants are subject to the same selection effects, these selection effects do not affect p_Q. Figure 6.6 shows that, for all relevant ξ, p_Q is a steeply and monotonically increasing function of θ_{max}. Thus, knowing ξ, one can use an empirical f_Q to determine p_Q and in turn θ_{max}.

Does one expect quasar giants to have a different distribution for $L_{p,obs} \mid L_{p,obs} \geq l_{p,GRG}$ than non-quasar giants? Interestingly, our framework allows us to prove that the inclination angle differences between the two classes affect their relative rarity, but not their observed projected proper length distribution. Under the basic unification model, quasar giants and non-quasar giants obey the same $L_{p,obs} \mid L_{p,obs} \geq l_{p,GRG}$ if they are subject to the same selection effects.

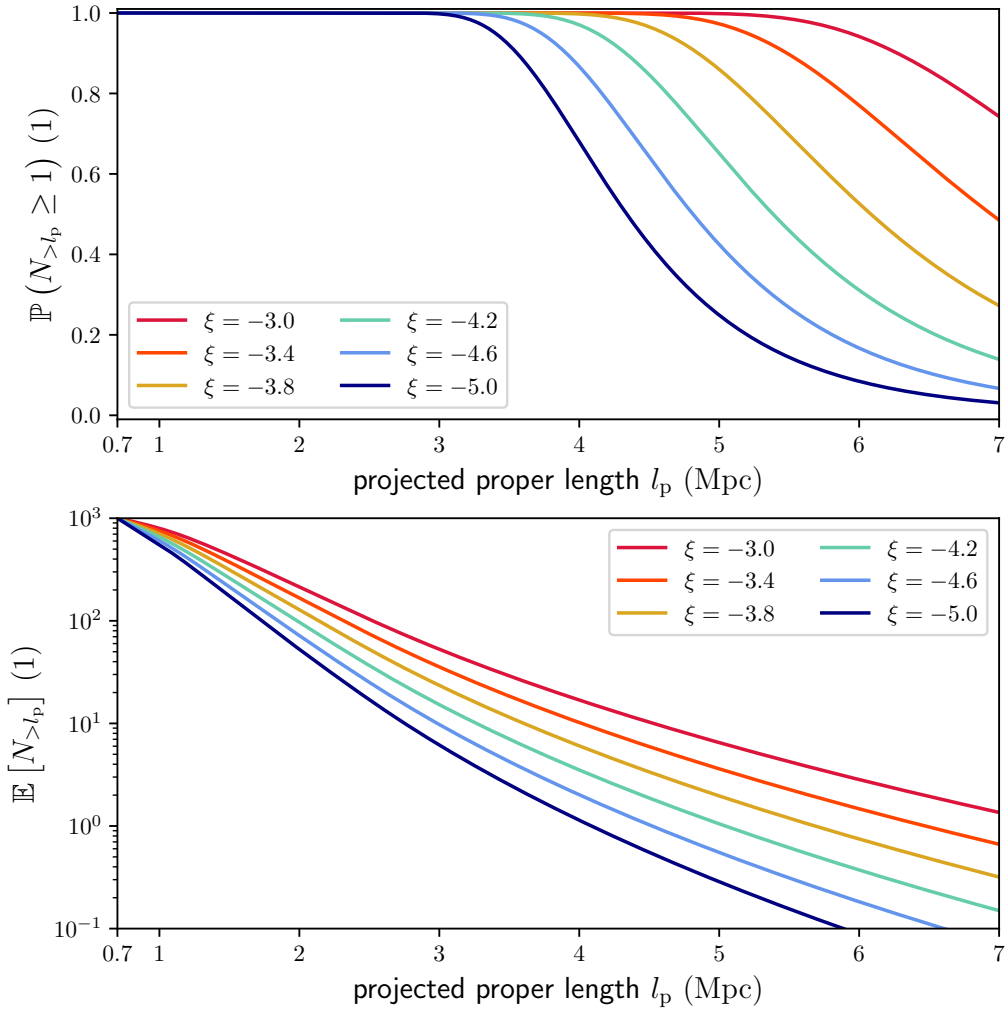


Figure 6.7: Predictions of the existence and expected number of giants that exceed projected length l_p in a sample of cardinality N, as functions of l_p. Both tail index ξ and selection effect parameters affect these predictions. We consider a sample of $N = 1000$ giants with redshifts below $z_\mathrm{max} = 0.5$, use $\varphi_\mathrm{min} = 3'$, $\varphi_\mathrm{max} = 7'$, $b_{\nu,\mathrm{ref}} = 1000\,\mathrm{Jy\,deg^{-2}}$, and $\sigma_\mathrm{ref} = 1.5$, and keep other parameters as in Fig. 6.5. *Top:* the probability that at least one observed giant has a projected length of at least l_p. *Bottom:* the expected number of observed giants with a projected length of at least l_p.

6.2.12 Extreme giants in a sample

The model can predict the occurrence of giants with extreme projected proper lengths in a sample of cardinality N. The probability that an observed GRG has a projected

proper length exceeding l_p is

$$p_{>l_\mathrm{p}} := \mathbb{P}\left(L_\mathrm{p,obs} > l_\mathrm{p} \mid L_\mathrm{p,obs} > l_\mathrm{p,GRG}\right) = \frac{1 - F_{L_\mathrm{p,obs}}\left(l_\mathrm{p}\right)}{1 - F_{L_\mathrm{p,obs}}\left(l_\mathrm{p,GRG}\right)}, \tag{6.38}$$

so that the number of observed giants with a projected proper length exceeding l_p is $N_{>l_\mathrm{p}} \sim \mathrm{Binom}(N, p_{>l_\mathrm{p}})$. Its expectation is $\mathbb{E}[N_{>l_\mathrm{p}}] = N \cdot p_{>l_\mathrm{p}}$. Furthermore, the probability that the sample contains at least one giant with projected proper length l_p or higher is

$$\mathbb{P}\left(N_{>l_\mathrm{p}} \geq 1\right) = 1 - \left(1 - p_{>l_\mathrm{p}}\right)^N. \tag{6.39}$$

See Appendix 6.A1.12 for details. Figure 6.7 shows $\mathbb{E}[N_{>l_\mathrm{p}}]$ and $\mathbb{P}(N_{>l_\mathrm{p}} \geq 1)$ for various ξ. As an example, the case $\xi = -3.0$ predicts that a sample of $N = 1000$ giants with redshifts below $z_\mathrm{max} = 0.5$ should contain almost ten giants of $l_\mathrm{p} > 5$ Mpc, and still several of $l_\mathrm{p} > 6$ Mpc. Such predictions are useful as they can be directly compared to elementary sample statistics.

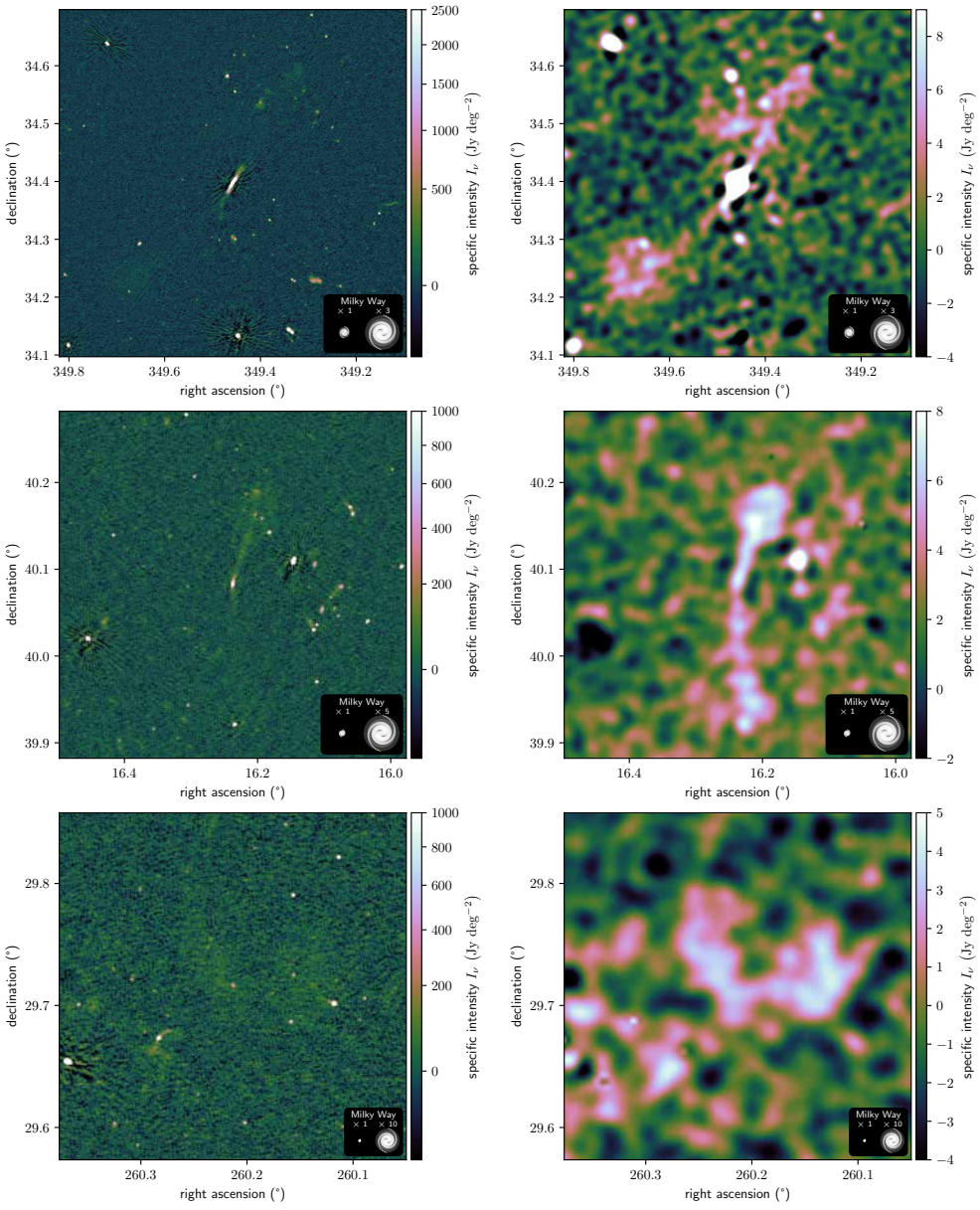


Figure 6.8: LoTSS DR2 cutouts of three newly discovered giants at 6″ *(left column)* and 60″ *(right column)*. By subtracting compact sources from calibrated 144 MHz visibility data and imaging at low resolution (60″ and 90″), we reveal otherwise speculative giant radio galaxies at the unexplored ∼1 Jy deg^{-2} surface brightness level. The claimed host galaxy is in the image centre. *Top:* a GRG of projected proper length $l_\mathrm{p} = 1.4 \pm 0.3$ Mpc, whose angular length $\varphi = 32.3 \pm 0.2'$ is larger than that of the full Moon. *Middle:* a GRG of $l_\mathrm{p} = 1.6 \pm 0.6$ Mpc and $\varphi = 16.4 \pm 0.2'$. *Bottom:* a GRG of $l_\mathrm{p} = 3.6 \pm 0.1$ Mpc and $\varphi = 8.5 \pm 0.2'$. For scale, we show the stellar Milky Way disk (with a diameter of 50 kpc) generated using the Ringermacher & Mead (2009) formula, alongside a 3, 5, or 10 times inflated version.

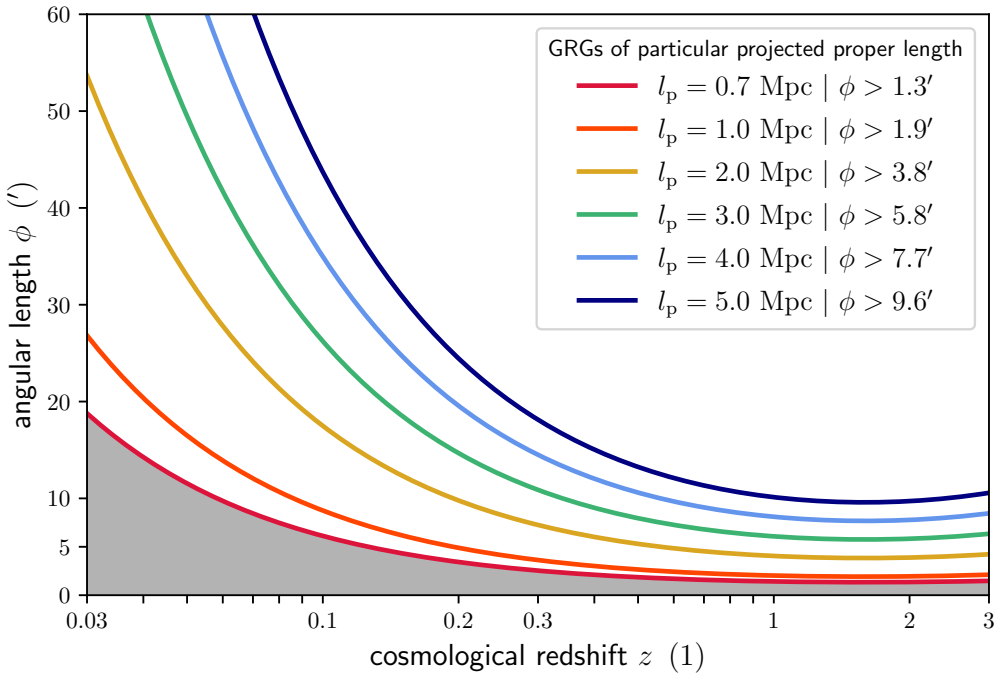


Figure 6.9: Relations between cosmological redshift z and angular length φ for six giants of different projected lengths l_p. Due to the expansion of the Universe, there is a minimum angular length for each l_p. If one defines giants as RGs with $l_p \geq l_{p,GRG} = 0.7$ Mpc, all giants have an angular length of $1.3'$ or above. If one instead defines giants as RGs with $l_p \geq l_{p,GRG} = 1$ Mpc, all giants have an angular length of $1.9'$ or above.

6.3 Sample compilation and properties

To measure the intrinsic GRG length distribution, we also require a large sample of giants collected from a single survey through a systematic approach. This ensures approximately homogeneous selection effects, which we can correct for in subsequent analysis using the statistical framework of Sect. 6.2.

6.3.1 LoTSS DR2

The Low-Frequency Array (LOFAR; van Haarlem et al., 2013) is a powerful, Pan-European radio interferometer that features both (sub)arcsecond-scale resolution and sensitivity to degree-scale structures. Dabhade et al. (2020b) have already demonstrated that this combination is ideal for detecting giants: these authors found a record 225 new specimina in the LoTSS DR1, the first data release of the LOFAR's North-

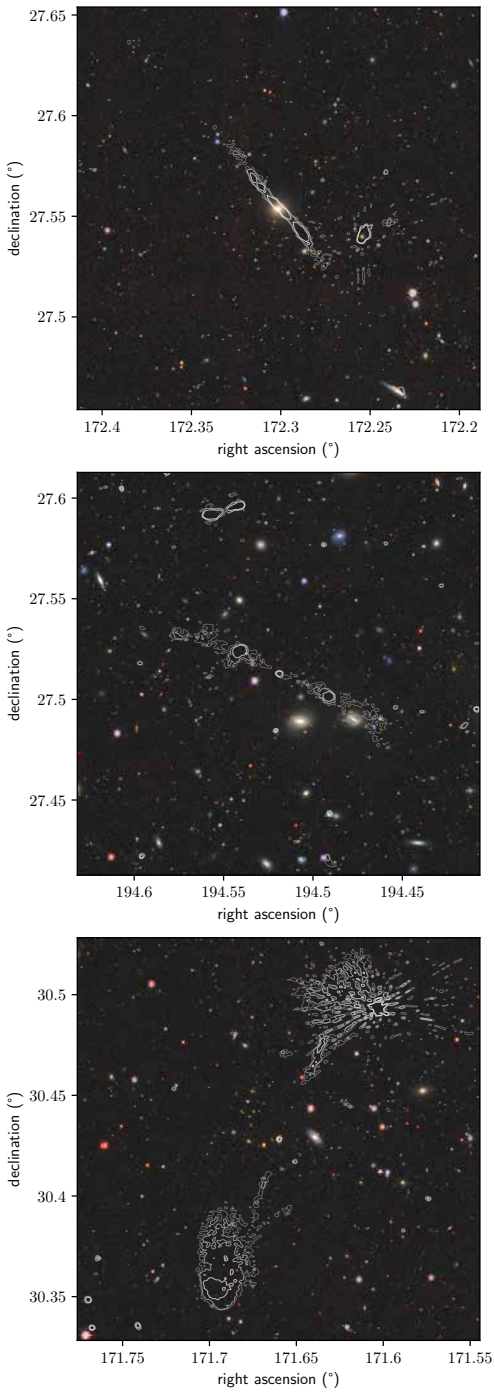


Figure 6.10: $12' \times 12'$ DESI Legacy Imaging Surveys DR9 (g, r, z)-details with LoTSS DR2 $6''$ contours $(3\sigma, 5\sigma, 10\sigma)$ overlaid. At $6''$ resolution, LoTSS images allow for more accurate host galaxy identification in SDSS, Pan-STARRS, and Legacy Survey images than was possible before. *Top:* the jet of the giant at rank 33 of Table 6.2 and shown in the middle-left panel of Fig. 6.24. *Middle:* the giant at rank 37 of Table 6.2. *Bottom:* the giant at rank 43 of Table 6.2.

ern Sky survey at central frequency $\nu_{\rm obs} = 144$ MHz.[9] Excitingly, the recent LoTSS DR2 (Shimwell et al., 2022) improves the data calibration and increases the survey footprint from 424 deg^2 to 5635 deg^2 — that is by more than a factor 13. By default, the LoTSS features imagery at 6″ and 20″ resolutions. To further facilitate the discovery of giants (among other goals), we reprocessed the LoTSS by subtracting compact sources and imaging at 60″ and 90″ resolution; more details are given in Oei et al. (2022a) and Oei et al. (prepa). This 60″ and 90″ imagery has turned out to be effective in highlighting jets and lobes of RGs of large angular and physical extent, whose surface brightnesses are low and which therefore have remained undetected in shallower surveys, and even in the LoTSS DR2 at higher resolutions. We demonstrate this fact in Fig. 6.8 by comparing the LoTSS DR2 at 6″ and 60″ for three giants whose discovery has relied on the lower-resolution images. After the serendipitous discovery of several such hitherto unknown giants in the 60″ and 90″ images, we decided to initiate a systematic, multi-resolution, visual GRG search through the area covered by LoTSS DR2 pipeline products as of September 2022.[10] This search comprised of a hundreds-of-hours-long inspection of the LoTSS maps at 6″ and 60″, alongside Pan-STARRS DR1 (Chambers et al., 2016) and SDSS DR9 (Ahn et al., 2012) maps, in *Aladin Desktop 11.0* (Bonnarel et al., 2000).

Reliable automated search strategies do not yet exist for several reasons. Giants showcase a rich morphological variety (see Sect. 6.3.6) and are easily confused with other types of astrophysical sources (see Sects. 6.3.3 and 6.3.4). Moreover, the known population is too small to effectively apply supervised learning techniques. However, it appears possible to find giants in morphological outlier lists of unsupervised learning techniques such as self-organising maps (SOMs; Mostert et al., 2021). The efficacy of such techniques in GRG searches has not yet been quantified.

6.3.2 Angular length threshold

To limit the amount of manual work, we decided to search only for GRG candidates whose angular length exceeds some threshold. In Fig. 6.9, we show how the angular length φ of giants with six different projected proper lengths $l_{\rm p}$ varies as a function of cosmological redshift z. Because of the expansion of the Universe, giants have an arcminute-scale minimum angular length: if $l_{\rm p,GRG} = 0.7$ Mpc, all giants obey $\varphi > 1.3'$.[11] For the purpose of finding a GRG, it is therefore never useful to inspect a

[9] As in the current work, the authors defined giants using the projected proper length threshold $l_{\rm p,GRG} = 0.7$ Mpc.

[10] In this process, we have skipped the enclosed LoTSS DR1 area, which has already been analysed by Dabhade et al. (2020b).

[11] If $l_{\rm p,GRG} = 1$ Mpc, all giants obey $\varphi > 1.9'$.

source with an angular length less than $1.3'$. Our highest priority has been to find giants with $z < 0.2$, which lie in a volume for which the total matter density field is known or will be known in the coming years through the combined power of deep spectroscopic surveys and Bayesian inference frameworks, such as the Bayesian Origin Reconstruction from Galaxies (BORG; Jasche & Wandelt, 2013; Jasche et al., 2015; Jasche & Lavaux, 2019). In an upcoming publication, we combine a sample of low-redshift giants with the BORG to analyse the large-scale environments of giants (Oei et al., prepb). Figure 6.9 shows that all giants with $l_{\mathrm{p}} > 1$ Mpc at $z < 0.2$ have an angular length $\varphi > 5'$. For this reason, we have chosen $5'$ as the angular length threshold of our search campaign. This choice has kept the visual inspection duration to order $\sim 10^2$ h, while still enabling us to target all Mpc-exceeding giants in the Local Universe ($z < 0.2$). In practice, this threshold is 'fuzzy': it is hard to accurately estimate angular lengths by eye before performing an actual measurement, so that our list of GRG candidates does contain some with a smaller angular length than the specified threshold. Inversely, it will presumably lack some GRG candidates with an angular length exceeding the threshold.

rank	host name SDSS DR12	host coordinates J2000 (°)	redshift (1)	redshift type	angular length (′)	projected proper length (Mpc)	host stellar mass ($10^{11} M_\odot$)	host SMBH mass ($10^9 M_\odot$)	host quasar
1	J081956.41+323537.6	124.9851, 32.5938	0.749 ± 0.073	*p*	11.2	5.07 ± 0.20	...	...	...
2	J081421.68+522410.0	123.5904, 52.4028	$0.2467 \pm 6 \cdot 10^{-5}$	*s*	20.8	4.99 ± 0.04	2.4 ± 0.4	0.4 ± 0.2	*n*
3	J142910.70+311245.0	217.2946, 31.2125	0.5921 ± 0.0001	*s*	11.7	4.80 ± 0.06	...	2.3 ± 2.0	*n*
4	J131823.42+262622.8	199.5976, 26.4397	$0.6230 \pm 5 \cdot 10^{-5}$	*s*	11.0	4.62 ± 0.06	...	...	*y*
5	J152634.77+262003.2	231.6449, 26.3342	$0.1507 \pm 2 \cdot 10^{-5}$	*s*	28.0	4.56 ± 0.03	3.7 ± 0.6	1.4 ± 0.3	*n*
6	J121815.66+382407.5	184.5653, 38.4021	0.634 ± 0.064	*p*	10.6	4.49 ± 0.21	...	...	...
7	J175735.88+405154.2	269.3995, 40.8651	0.585 ± 0.036	*p*	10.5	4.29 ± 0.14	...	...	...
8	J161622.52+111135.7	244.0939, 11.1933	$0.3574 \pm 7 \cdot 10^{-5}$	*s*	13.4	4.15 ± 0.05	9.5 ± 1.8	5.7 ± 3.1	*n*
9	J154709.22+353846.1	236.7884, 35.6462	$0.0794 \pm 1 \cdot 10^{-5}$	*s*	43.8	4.08 ± 0.01	4.6 ± 0.1	3.9 ± 0.9	*n*
10	J013406.32+301537.2	23.5264, 30.2604	0.884 ± 0.138	*p*	8.5	4.06 ± 0.22	...	...	...
11	J082747.88+662813.6	126.9495, 66.4705	0.968 ± 0.160	*p*	8.2	4.02 ± 0.20	...	...	...
12	J012440.54+194003.9	21.1689, 19.6678	0.578 ± 0.162	*p*	9.6	3.90 ± 0.58	...	...	...
13	...	238.4466, 28.4763	1.094 ± 0.122	*p*	7.7	3.88 ± 0.12	...	...	...
14	...	275.3624, 26.6599	0.0850 ± 0.0001	*s*	39.0	3.86 ± 0.02	...	...	...
15	J162656.58+543421.3	246.7358, 54.5726	0.4887 ± 0.0001	*s*	10.3	3.84 ± 0.06	...	1.0 ± 1.6	*n*
16	J220605.67+275100.3	331.5237, 27.8501	0.317 ± 0.116	*p*	13.2	3.78 ± 0.99	...	...	...
17	J023544.96+310447.5	38.9373, 31.0799	0.541 ± 0.063	*p*	9.6	3.77 ± 0.23	...	...	...
18	...	136.9661, 67.1071	0.754 ± 0.081	*p*	8.3	3.77 ± 0.17	...	...	...
19	J180117.72+510722.4	270.3239, 51.1229	0.448 ± 0.086	*p*	10.3	3.66 ± 0.42	...	...	...
20	J123900.69+360924.5	189.7529, 36.1568	$0.5935 \pm 6 \cdot 10^{-5}$	*s*	8.8	3.62 ± 0.06	...	...	*y*
21	J102430.93+381842.8	156.1289, 38.3119	0.411 ± 0.028	*p*	10.7	3.62 ± 0.16	...	...	...
22	J172051.08+294256.8	260.2129, 29.7158	$\geq 0.620 \pm 0.044$	*p*	8.5	$\geq 3.56 \pm 0.13$	...	...	...
23	J090534.54+563052.0	136.3939, 56.5145	0.898 ± 0.056	*p*	7.4	3.55 ± 0.10	...	...	...
24	J133105.80+293435.7	202.7742, 29.5766	0.734 ± 0.034	*p*	7.9	3.55 ± 0.09	...	...	...
25	J223649.76+251242.5	339.2074, 25.2118	0.749 ± 0.075	*p*	7.8	3.53 ± 0.15	...	...	...
26	J230125.38+240148.2	345.3558, 24.0301	0.450 ± 0.089	*p*	9.9	3.53 ± 0.41	...	...	...
27	J004848.01+021003.1	12.2001, 2.1675	$\geq 0.3604 \pm 9 \cdot 10^{-5}$	*s*	11.2	$\geq 3.49 \pm 0.05$	...	3.0 ± 1.5	*n*
28	J220239.13+070656.7	330.6631, 7.1158	$0.4649 \pm 7 \cdot 10^{-5}$	*s*	9.6	3.48 ± 0.05	...	0.9 ± 0.6	*n*
29	J165113.78+320943.4	252.8074, 32.1621	0.744 ± 0.045	*p*	7.7	3.48 ± 0.10	...	...	...
30	...	102.0173, 70.8276	0.714 ± 0.049	*p*	7.8	3.47 ± 0.12	...	...	...
31	...	11.2869, 28.7951	0.668 ± 0.042	*p*	8.0	3.46 ± 0.11	...	...	...
32	...	127.9215, 67.1934	0.451 ± 0.027	*p*	9.5	3.39 ± 0.12	...	...	...
33	J112912.14+273313.9	172.3006, 27.5539	$0.0732 \pm 1 \cdot 10^{-5}$	*s*	38.6	3.34 ± 0.01	3.7 ± 0.1	3.6 ± 0.8	*n*
34	J163659.07+541725.4	249.2461, 54.2904	$0.5027 \pm 5 \cdot 10^{-5}$	*s*	8.8	3.33 ± 0.06	...	53.4 ± 59.1	*n*
35	J092826.93+230448.0	142.1122, 23.0800	0.491 ± 0.045	*p*	8.9	3.33 ± 0.17	...	...	...

rank	host name SDSS DR12	host coordinates J2000 (°)	redshift (1)	redshift type	angular length (′)	projected proper length (Mpc)	host stellar mass ($10^{11} M_\odot$)	host SMBH mass ($10^9 M_\odot$)	host quasar
36	J001152.65+310024.3	2.9694, 31.0068	0.757 ± 0.083	*p*	7.3	3.32 ± 0.15	...	...	...
37	J125804.46+273046.0	194.5186, 27.5128	0.741 ± 0.083	*p*	7.3	3.29 ± 0.16	...	...	...
38	...	182.5080, 44.0903	1.031 ± 0.163	*p*	6.6	3.28 ± 0.15	...	...	...
39	J084127.02+554627.1	130.3626, 55.7742	$0.7912 \pm 3 \cdot 10^{-5}$	*s*	7.1	3.28 ± 0.07	...	25.2 ± 27.0	*n*
40	J143011.92+410404.2	217.5497, 41.0678	0.5868 ± 0.0001	*s*	8.0	3.27 ± 0.06	...	2.1 ± 1.4	*n*
41	J135119.31+340844.1	207.8305, 34.1456	0.923 ± 0.121	*p*	6.7	3.24 ± 0.14	...	...	...
42	J134436.33+291239.6	206.1514, 29.2110	0.766 ± 0.134	*p*	7.1	3.24 ± 0.23	...	...	...
43	J112638.34+302541.9	171.6598, 30.4283	$0.3049 \pm 4 \cdot 10^{-5}$	*s*	11.6	3.23 ± 0.04	...	1.5 ± 0.6	*n*
44	J012342.20+293633.1	20.9259, 29.6092	2.525 ± 0.300	*p*	6.5	3.22 ± 0.11	...	...	...
45	J223224.15+285753.3	338.1007, 28.9648	0.566 ± 0.038	*p*	8.0	3.21 ± 0.12	...	...	...
46	J103731.47+312948.9	159.3811, 31.4969	0.5228 ± 0.0001	*s*	8.3	3.21 ± 0.06	...	2.2 ± 1.4	*n*
47	J154742.69+384119.4	236.9279, 38.6887	0.280 ± 0.024	*p*	12.2	3.21 ± 0.20	...	...	...
48	J114333.93+425800.5	175.8914, 42.9668	0.802 ± 0.075	*p*	6.9	3.20 ± 0.12	...	...	...
49	J122329.86+313116.0	185.8744, 31.5211	0.666 ± 0.042	*p*	7.4	3.20 ± 0.11	...	...	...
50	...	212.1410, 67.8036	1.103 ± 0.181	*p*	6.3	3.17 ± 0.14	...	...	...

Table 6.2: Properties of the 50 projectively longest giants out of a total of 2060 discovered during our LoTSS DR2 search campaign. The giants are ranked by projected proper length. The column 'rank' thus denotes each giant's projected proper length rank within this new sample, not within the total known population. Lying outside of the coverage, some GRG host galaxies have no SDSS DR12 name. The column 'host coordinates' contains the central right ascension and declination of the host galaxy. The columns 'redshift' and 'redshift type' provide cosmological redshift estimates, derived from spectroscopy *s* or photometry *p*. The column 'angular length' denotes the largest great-circle distance between two LoTSS DR2–detectable GRG (end)points. These angular lengths may increase in future deeper surveys. We have not measured angular length errors on a case by case basis, but estimate them to be 0.15′. The column 'projected proper length' propagates both redshift and angular length uncertainty and assumes the Planck Collaboration et al. (2020) cosmology. The columns 'host stellar mass' and 'host SMBH mass', further discussed in Section 6.3.6 and Appendix 6.A3, provide SDSS-derived estimates of host stellar and supermassive black hole masses. Finally, the column 'host quasar' indicates whether the host's AGN has a quasar appearance: *y* (yes) or *n* (no). In cases where only a set of candidates containing the host galaxy could be established beyond reasonable doubt, we list the properties of the lowest-redshift candidate. In this way, the provided projected proper length bounds the actual projected proper length from below. In such cases, to signify uncertainty, we mark the host name, host coordinates, and physical host properties in grey. We note that this lowest-redshift candidate is often, but not always, also the most probable host. For access to these data for all 2060 newly discovered giants, see Appendix 6.A6.

6.3.3 GRG candidates in the radio

We first identified GRG candidates in the LoTSS at 6″ and 60″. These maps serve complementary roles. The 6″ images reveal the precise morphology of radio galaxy cores and jets, which are necessary to pinpoint the host galaxy. Figure 6.10 provides a representative sense of the host galaxy identification accuracy these data allow for when combined with modern optical surveys such as the DESI Legacy Imaging Surveys (Dey et al., 2019) DR9. In contrast, the 60″ images have such compact sources removed or highly suppressed, but better highlight diffuse structures, such as RG lobes. Being similar in morphology, we made sure not to interpret diffuse emission from low-redshift spiral galaxies and their circumgalactic media, or radio halos and relics in galaxy clusters, as RG lobe emission. We required that all new RGs feature a detection of at least two[12] lobes, or of at least one lobe and one jet oriented towards the lobe(s), at at least one of the resolutions used in this work (6″, 20″, 60″, and 90″).

6.3.4 GRG candidates in the optical

To confirm that a radio structure really is a radio galaxy, we compared the radio images with optical images of the same sky region. If a patch of radio emission is indeed due to RG jets or lobes, the patch itself must have *no* codirectional galactic counterpart in the optical. If the radio emission is due to a low-redshift spiral galaxy or a galaxy cluster instead, a corresponding easily recognisable counterpart *will* exist. We also took care not to erroneously associate the lobes of two distinct RGs. For this reason, we were more cautious to associate a pair of lobes to a suspected host galaxy when, in the optical, one could discern other galaxies in the angular vicinity of the lobes that could have generated them instead.

The Pan-STARRS and SDSS images used for these purposes complement each other, as they differ in quality throughout the sky — and in particular around sources of high optical flux density. Neither consistently outperforms the other. Only Pan-STARRS covers the full Northern Sky and could thus always be relied upon.

6.3.5 Host galaxy identification

We also used the Pan-STARRS and SDSS maps for the identification of host galaxies. We collected host redshifts from the SDSS DR12 (Alam et al., 2015) and *Gaia* (Gaia Collaboration et al., 2016) DR3 (Gaia Collaboration et al., 2021) through automated *VizieR* queries, from the Galaxy List for the Advanced Detector Era (GLADE) 2.4 (Dálya et al., 2018), and from the DESI DR9 photometric redshift catalogue (Zou

[12] A small fraction of observed RGs are double-double radio galaxies, which show *four* lobes.

et al., 2022). If redshifts from multiple sources were available, we favoured SDSS over GLADE data, GLADE over *Gaia* data, and *Gaia* over DESI data. Similarly, we only adopted photometric redshifts if spectroscopic ones were not available.

For a small subset of RGs, a definite host galaxy could not be established beyond reasonable doubt, but a set of candidates containing the host galaxy could. In these cases, the lowest-redshift candidate provides a lower bound to the RG's projected proper length.[13] If this lower bound exceeds $l_{\mathrm{p,GRG}} = 0.7$ Mpc, then the actual projected proper length certainly does so too, and the RG can be classified as a GRG — despite persisting uncertainty concerning the identity of the host galaxy.

6.3.6 LoTSS DR2 GRG sample

Our search campaign has led to the identification of 2060 hitherto unknown giants. To establish novelty, we assembled a literature catalogue with all known giants as of September 2022, combining the catalogue of Dabhade et al. (2020a)[14] with the giants discovered in Galvin et al. (2020), Ishwara-Chandra et al. (2020), Tang et al. (2020), Bassani et al. (2021), Brüggen et al. (2021), Delhaize et al. (2021), Masini et al. (2021), Kuźmicz & Jamrozy (2021), Andernach et al. (2021), Mahato et al. (2022), Gürkan et al. (2022), and Simonte et al. (2022). Fusing our sample with this literature catalogue, we obtain a final catalogue with $N = 3341$ giants.

Figure 6.11 shows the locations of all known giants in the sky. We list basic properties of the 50 projectively largest new discoveries in Table 6.2, and refer to Appendix 6.A6 for access to these data for our entire sample. In Fig. 6.12 and Figs. 6.24–6.26, we present images for discoveries with projected proper lengths l_{p} in the ranges 5.1–4 Mpc, 4–3 Mpc, 3–2 Mpc, and 2–0.7 Mpc, respectively. For each giant, we use the LoTSS resolution $\theta_{\mathrm{FWHM}} \in \{6'', 20'', 60'', 90''\}$ that most clearly conveys the morphology through a single image. The selection reflects our sample's diversity in shapes and sizes and provides a sense of the data quality.

[13] This is true as long as all candidates have a redshift below that of the angular diameter distance maximum: $z = 1.59$ for the cosmology adopted.

[14] This catalogue, complete up to and including April 2020, contains the giants found in 40 prior publications (for a list, see Dabhade et al. (2020a)'s Sect. 1: Introduction), alongside their own discoveries.

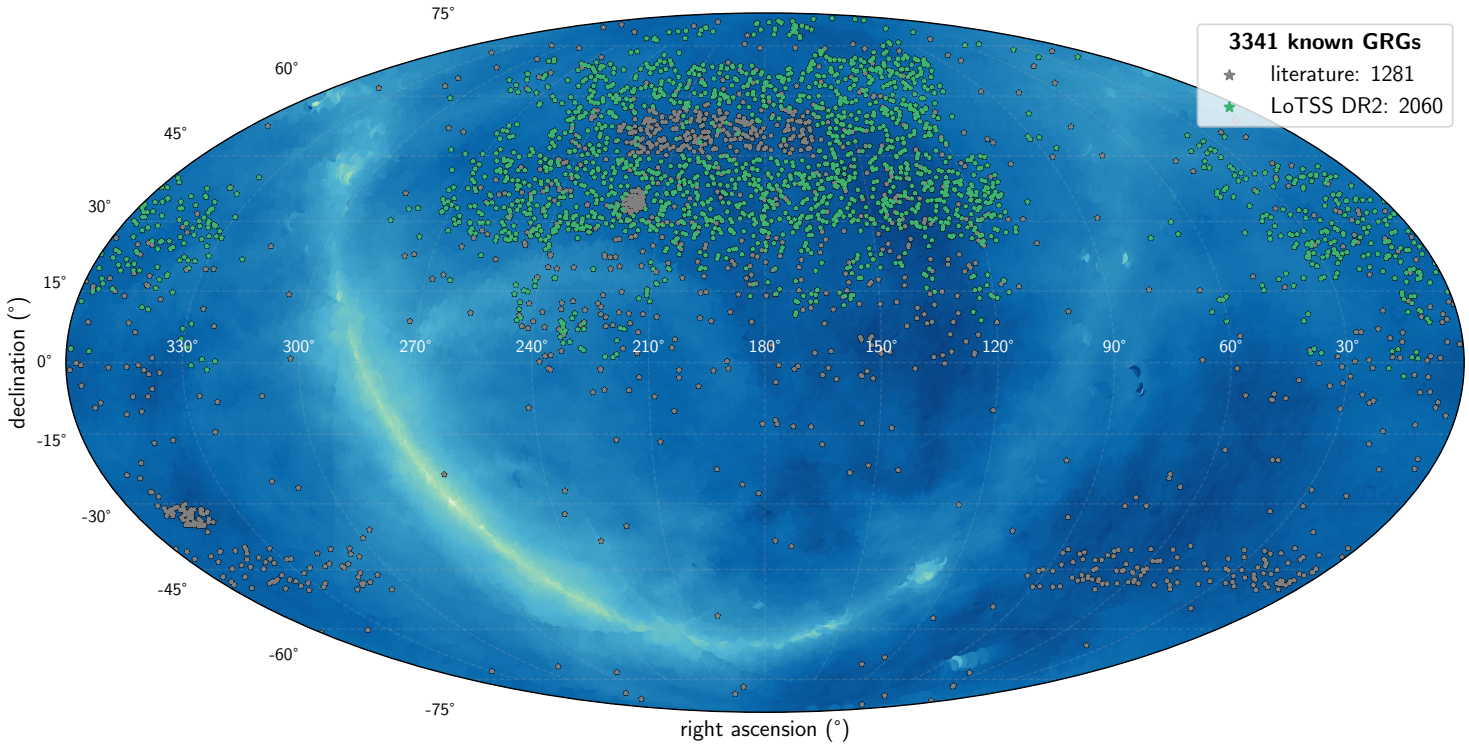


Figure 6.11: Mollweide view of the sky showing locations of all known giants, of which 62% are discoveries presented in this work. In the background, we show the specific intensity function of the Milky Way at $\nu_{\rm obs} = 150$ MHz (Zheng et al., 2017). The LoTSS DR2 has avoided the Galactic Plane, where extended emission complicates calibration and deconvolution. Our search footprint encloses a grey spherical rectangle, which represents the LoTSS DR1 search by Dabhade et al. (2020b), and a grey spherical cap, which represents the Boötes LOFAR Deep Field search by Simonte et al. (2022).

Angular lengths

The angular length distribution of the newly found giants is as follows: the smallest $\varphi = 1.5'$, the median $\varphi = 5.2'$, the largest $\varphi = 2.2°$, and 80% of angular lengths fall within $[3.4', 9.8']$. Thirteen of our discoveries — listed in Table 6.3 — are larger than the Moon in the sky (whose angular diameter varies over time, but here taken to be $\varphi = 30'$). Our search more than doubles the known number of such spectacular giants — from 10 to 23.

The GRG associated with NGC 2300 (see the middle-left panel of Fig. 6.13) is the giant with the largest angular length ever found, and the radio galaxy with the largest angular length in the Northern Sky.[15] It remains possible that the GRG has been generated by spiral galaxy NGC 2276 instead, with which elliptical galaxy NGC 2300 is interacting. However, this scenario seems unlikely, as only a fraction $\sim 10^{-3}$ of known giants are hosted by spirals. Its discovery emphasises that low-frequency interferometers like the LOFAR and the MWA, which are sensitive to degree-scale angular scales, are important to complete a low-redshift census of giant radio galaxies. Sky-wide, a simple extrapolation of our findings suggests that several ($\sim 10^1$) degree-scale angular length giants similar to the GRG of NGC 2300 still await discovery at the LoTSS DR2 depth.

Table 6.3: (Extended) Table 6.2 ranks, host names, angular lengths, and LoTSS DR2 image references for all 13 newly discovered giants that appear larger in the sky than the Moon ($\varphi > 30'$).

rank Table 6.2	host name NGC or LEDA	angular length ($'$)	LoTSS DR2 image reference
1441	NGC 2300	2.2°	Fig. 6.13, middle left
141	NGC 6185	1.0°	Fig. 6.13, top left
9	LEDA 56028	43.8′	Fig. 6.12, bottom left
144	LEDA 54794	40.0′	-
14	LEDA 5060619	39.0′	Fig. 6.24, top left
33	LEDA 1811497	38.6′	Fig. 6.24, middle left
1692	NGC 2789	34.5′	Fig. 6.13, bottom right
1643	LEDA 38523	33.8′	Fig. 6.13, middle right
1770	NGC 1044	32.7′	Fig. 6.13, bottom left
178	LEDA 2048533	32.3′	Fig. 6.8, top row
1486	NGC 7385	32.0′	-
326	LEDA 3090801	31.0′	Fig. 6.25, bottom left
465	LEDA 37801	30.1′	Fig. 6.13, top right

[15] At $\varphi = 8°$, the Southern Sky's Centaurus A (Cooper et al., 1965) is the radio galaxy with the largest angular length overall (e.g. McKinley et al., 2022); despite this, at $l_p = 0.48$ Mpc, it is not a giant.

Redshifts

The redshift distribution of the newly found giants is as follows: the lowest $z = 0.00635 \pm 6 \cdot 10^{-5}$, the median $z = 0.29$, the highest $z = 2.6394 \pm 6 \cdot 10^{-4}$, and 80% of redshifts fall within [0.12, 0.68]. Because of our focus on giants of large angular length (see Sect. 6.3.2), we have found only 36 giants beyond $z > 1$. One of these, the GRG at rank 13 of Table 6.2, is the largest secure giant found beyond redshift 1. It lies at $z = 1.1 \pm 0.1$ and spans $l_\mathrm{p} = 3.9 \pm 0.1$ Mpc. Its host galaxy does not appear to contain a quasar.

Projected proper lengths

With $l_\mathrm{p} = 5.1 \pm 0.2$ Mpc and $l_\mathrm{p} = 4.99 \pm 0.04$ Mpc, our LoTSS DR2 sample contains the first two 5 Mpc–scale giants. We have presented a dedicated analysis of the latter GRG in Oei et al. (2022a). 11 discoveries have $l_\mathrm{p} \geq 4$ Mpc, 53 have $3 \leq l_\mathrm{p} < 4$ Mpc, 291 have $2 \leq l_\mathrm{p} < 3$ Mpc, 1215 have $1 \leq l_\mathrm{p} < 2$ Mpc, and 490 have $0.7 \leq l_\mathrm{p} < 1$ Mpc. The median $l_\mathrm{p} = 1.35$ Mpc, and 80% of projected proper lengths fall within [0.82 Mpc, 2.29 Mpc].

Stellar and supermassive black hole masses

Following Oei et al. (2022a), we collected host stellar masses $M_\star$ from Chang et al. (2015) and Salim et al. (2018), and estimated host SMBH masses $M_\bullet$ via SDSS DR12 stellar velocity dispersions (Alam et al., 2015) and the M-sigma relation of Kormendy & Ho (2013)'s Eq. (7). From all 3341 giants in our final catalogue, only 732 (22%) could be assigned a stellar mass in this way, and only 1115 (33%) an SMBH mass; for both quantities, our LoTSS DR2 sample accounts for four-fifths of the resulting subpopulation. Fig. 6.27 shows both $M_\star$ and $M_\bullet$ in relation to projected proper length l_p.

The median $M_\star = 3.4 \cdot 10^{11}\, M_\odot$, and 80% of stellar masses fall within $[1.8 \cdot 10^{11}\, M_\odot,\ 5.3 \cdot 10^{11}\, M_\odot]$. We discover two giants whose hosts, J150329.07+374850.3 and J073505.24+415827.5, are the least massive known: both have a stellar mass $M_\star = 4.8 \cdot 10^{10}\, M_\odot$. These are small giants, with $l_\mathrm{p} = 0.8$ Mpc and $l_\mathrm{p} = 0.7$ Mpc, respectively. The top panel of Fig. 6.27 hints at a weak positive correlation between $M_\star$ and l_p, which future work should confirm or reject.

The median $M_\bullet = 1.0 \cdot 10^{9}\, M_\odot$, and 80% of SMBH masses fall within $[0.4 \cdot 10^{9}\, M_\odot,\ 2.2 \cdot 10^{9}\, M_\odot]$. The SMBH masses of J123703.24+275819.5, J163659.07+-541725.4, and J103129.54+502959.1 are the highest estimated yet, with $M_\bullet = 2 \cdot$

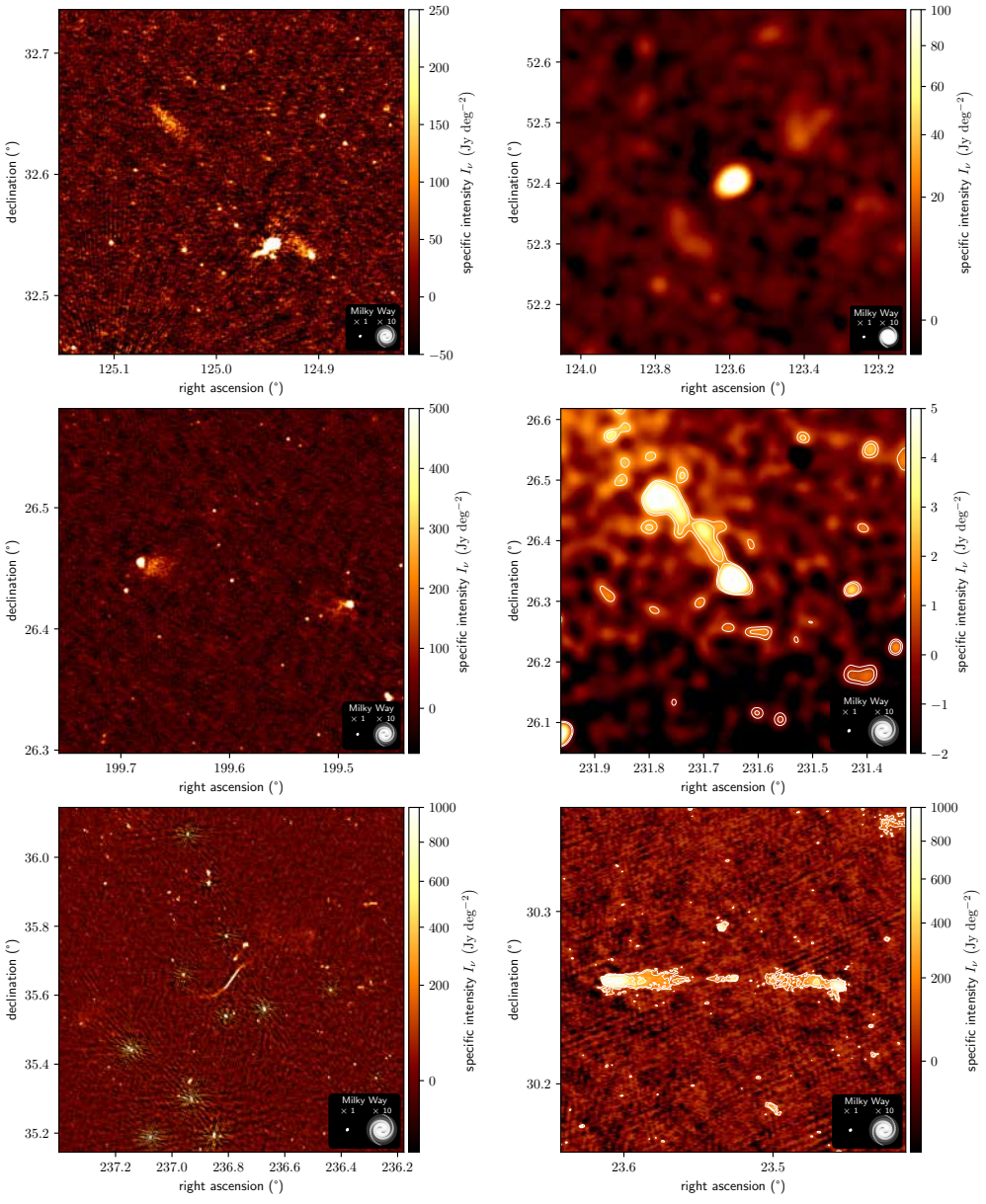


Figure 6.12: Details of the LoTSS DR2–estimated specific intensity function $I_\nu(\hat{r})$ at central observing frequency $\nu_{\rm obs} = 144$ MHz and resolutions $\theta_{\rm FWHM} \in \{6'', 90''\}$, centred around the hosts of newly discovered giants. Row-wise from left to right, from top to bottom, the projected proper length $l_{\rm p}$ is 5.1 Mpc, 5.0 Mpc (Oei et al., 2022a), 4.6 Mpc, 4.6 Mpc, 4.1 Mpc, and 4.1 Mpc; in the same order, $\theta_{\rm FWHM}$ is $6''$, $90''$, $6''$, $90''$, $6''$, and $6''$. The GRG in the bottom-left panel appears larger in the sky than the Moon. In the middle-right panel, contours signify 2.5 and 3.5 sigma-clipped standard deviations (SDs) above the sigma-clipped median; in the bottom-right panel, they signify 3, 5, and 10 such SDs. For scale, we show the stellar Milky Way disk (with a diameter of 50 kpc) generated using the Ringermacher & Mead (2009) formula, alongside a 10 times inflated version.

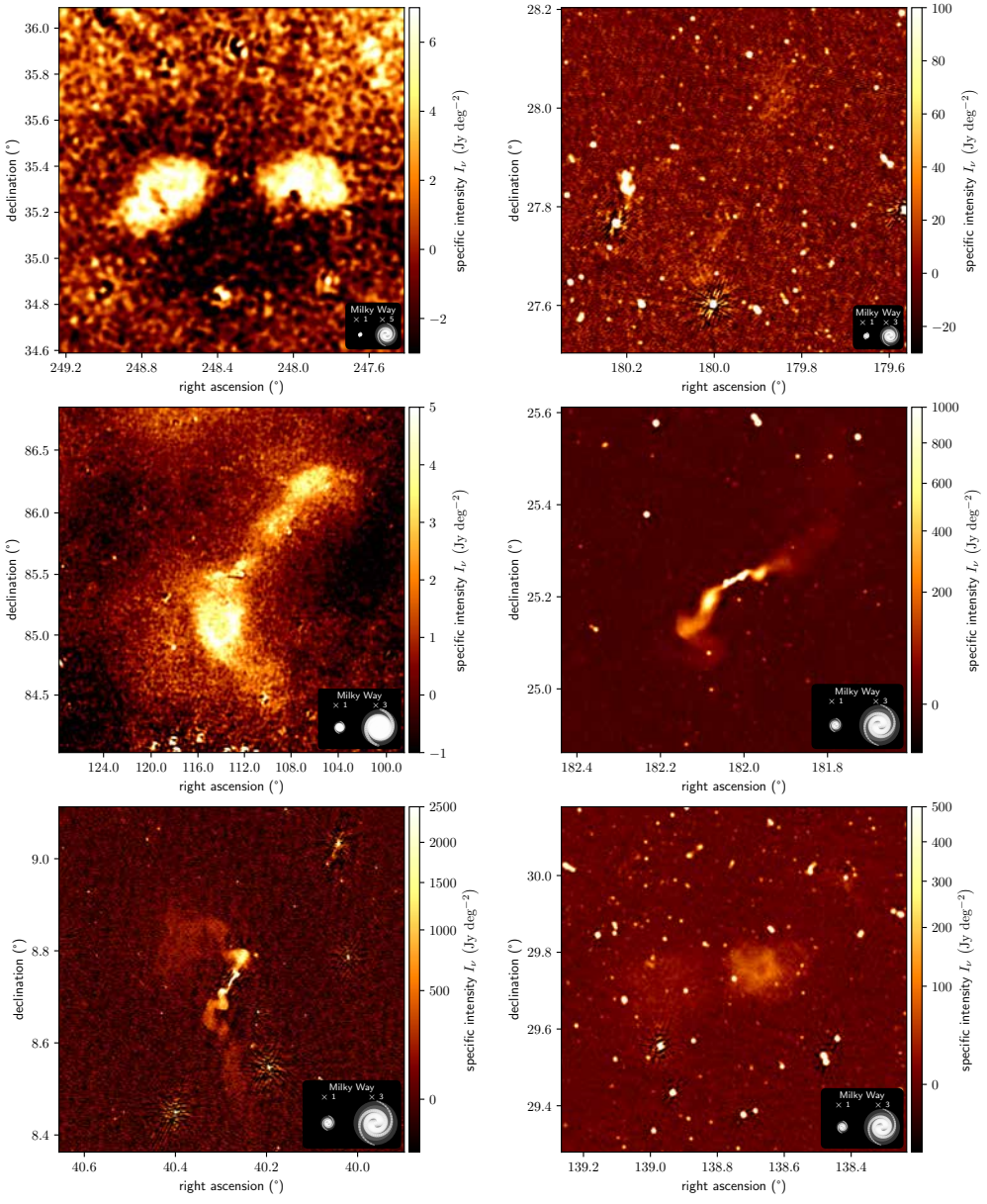


Figure 6.13: Details of the LoTSS DR2–estimated specific intensity function $I_\nu(\hat{r})$ at central observing frequency $\nu_{\rm obs} = 144$ MHz and resolutions $\theta_{\rm FWHM} \in \{6'', 20'', 90''\}$, centred around the hosts of newly discovered giants. Row-wise from left to right, from top to bottom, the projected proper length $l_{\rm p}$ is 2.5 Mpc, 1.9 Mpc, 1.1 Mpc, 1.0 Mpc, 0.9 Mpc, and 0.9 Mpc; in the same order, $\theta_{\rm FWHM}$ is 90″, 20″, 90″, 20″, 6″, and 20″. All appear larger in the sky than the Moon. The top-left panel shows the giant of NGC 6185, a *spiral* galaxy. This is the first spiral galaxy–hosted giant known with $l_{\rm p} > 2$ Mpc. The middle-left panel shows a structure we interpret as a radio galaxy belonging to the elliptical galaxy NGC 2300. At $\varphi = 2.2°$, this giant has the largest angular length of all uncovered thus far.

$10^{11}\ M_\odot$, $M_\bullet = 5 \cdot 10^{10}\ M_\odot$, and $M_\bullet = 5 \cdot 10^{10}\ M_\odot$, respectively. The latter masses equal the theoretical maximum mass of accreting black holes of typical spin (King, 2016). Curiously, although J103129.54+502959.1's $M_\bullet$ is among the highest estimated SMBH mass of any GRG host, the GRG itself is relatively small: $l_\mathrm{p} = 0.81 \pm 0.06$ Mpc. Conversely, the bottom panel of Fig. 6.27 shows that multi-Mpc giants can have hosts with SMBH masses that are two orders of magnitude lower than J103129.54+502959.1's.

Spiral or lenticular host galaxies

Remarkably, although NGC 6185 is a spiral galaxy of Hubble–de Vaucouleurs class SAa (Jansen et al., 2000), it appears to have generated the giant shown in the top-left panel of Fig. 6.13. Such systems are exceedingly rare: not only are few RGs giants, but also virtually all giants have an elliptical galaxy as their host. With $l_\mathrm{p} = 2.54 \pm 0.01$ Mpc, this is the largest known spiral galaxy–hosted radio galaxy. Hitherto, the largest spiral galaxy–hosted giant in the literature has been J2345-0449 (Bagchi et al., 2014), with $l_\mathrm{p} = 1.6$ Mpc. Given the favourably low redshift of $z = 0.03430 \pm 7 \cdot 10^{-5}$, NGC 6185 and its enigmatic giant solicit a dedicated analysis (Oei et al., 2023b). Besides NGC 6185, spiral or lenticular galaxies J080403.40+404809.3 and J091459.66+294348.8 (known alternatively as NGC 2789; see the bottom-right panel of Fig. 6.13) have also generated giants; these have projected proper lengths $l_\mathrm{p} = 1.1$ Mpc and $l_\mathrm{p} = 0.9$ Mpc, respectively. Morphological host classification through SDSS, Pan-STARRS, and DESI imagery is reliable only up to $z \sim 0.1$–0.2, depending on viewing angle and various other factors. Our LoTSS DR2 sample contains 342 giants with definite hosts at $z = 0.15$ or below, among which are all 3 giants with spiral or lenticular hosts discussed here. It thus appears that the fraction of GRG hosts that is of such non-elliptical nature is $\sim$1%. A more detailed morphological characterisation of the sample appears possible, for example using data from Hart et al. (2016), but this is beyond the scope of the current work.

6.4 Results

After first building a statistical framework and then collecting a large sample of giants from a single survey through a systematic approach, we were ready to infer the intrinsic GRG length distribution. In particular, we aimed to establish whether the RG intrinsic proper length RV L is well described by a Pareto distribution, and if so, what its tail index ξ is. Subsequently, we inferred derived quantities, such as the comoving GRG number density in the Local Universe.

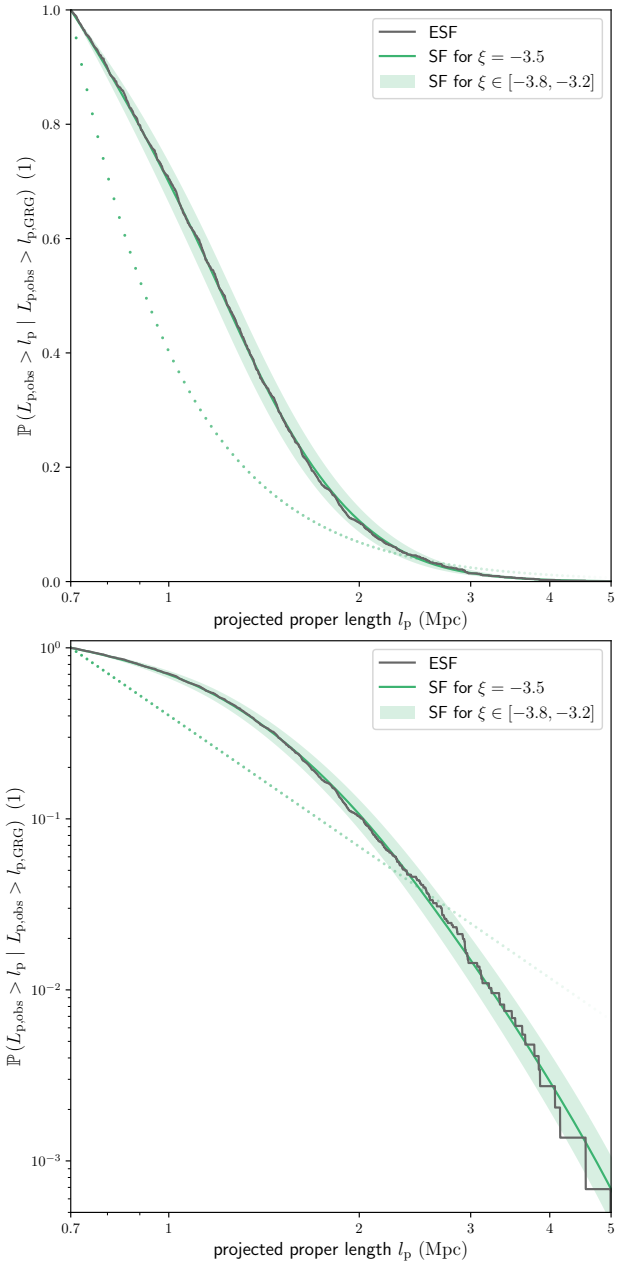


Figure 6.14: Empirical survival function of the observed giant radio galaxy projected proper length RV (ESF; dark grey) and the corresponding survival function $1 - F_{L_{\mathrm{p,obs}} \mid L_{\mathrm{p,obs}} > l_{\mathrm{p,GRG}}}$ (SF; green curve) using the maximum a posteriori probability parameters (MAP; see Table 6.4). Observed GRG projected lengths are well described by a Pareto distribution modified to include selection effects. Keeping the selection effect parameters fixed, we show how models vary with tail index ξ (green range). We also show the selection effect–free SF $1 - F_{L_{\mathrm{p}} \mid L_{\mathrm{p}} > l_{\mathrm{p,GRG}}}$ using the MAP ξ (green dots). We included all LoTSS DR2 search campaign giants up to $z_{\mathrm{max}} = 0.5$. *Left:* logarithmic horizontal axis and linear vertical axis. *Right:* logarithmic horizontal axis and logarithmic vertical axis.

6.4.1 Giant radio galaxy length distribution

Empirical survival function

From our LoTSS DR2 GRG sample, we computed the empirical cumulative distribution function (ECDF) of the GRG observed projected proper length RV $L_{\mathrm{p,obs}} \mid L_{\mathrm{p,obs}} > l_{\mathrm{p,GRG}}$ (see Eq. 8.11). We only included giants for which l_{p} itself — rather than a lower bound — is known, and set either $z_{\mathrm{max}} = 0.5$ or $z_{\mathrm{max}} = 0.25$; this retained 1473 or 811 giants for analysis, respectively.[16] The empirical survival function (ESF), which equals one minus the ECDF, is shown in dark grey in both panels of Fig. 6.14. Just as the ECDF approximates the CDF, the ESF approximates the survival function (SF). In this case, for any l_{p}, the ESF provides the probability that a randomly drawn LoTSS DR2 GRG will have a projected proper length exceeding l_{p}. If $L_{\mathrm{p,obs}} \mid L_{\mathrm{p,obs}} > l_{\mathrm{p,GRG}}$ were Paretian, its ESF (hereafter: 'the' ESF) would resemble a straight line in Fig. 6.14's bottom panel — both axes have logarithmic scaling. However, the ESF clearly displays curvature.

Selection effects

Under the ansatz that L is Paretian, as proposed in Sect. 6.2.1, the aforementioned ESF's curvature implies a significant role for observational selection effects. The reason is the following. The ansatz implies that the GRG projected proper length RV $L_{\mathrm{p}} \mid L_{\mathrm{p}} > l_{\mathrm{p,GRG}}$ is also Paretian (see Eq. 6.5), as is $L_{\mathrm{p,obs}} \mid L_{\mathrm{p,obs}} > l_{\mathrm{p,GRG}}$ when selection effects are negligible (see Sect. 6.2.8 and set $C(l_{\mathrm{p}}) = 1$). But in our case $L_{\mathrm{p,obs}} \mid L_{\mathrm{p,obs}} > l_{\mathrm{p,GRG}}$ is not Paretian: its ESF is curved. To avoid contradiction, we must relax at least one assumption. If the Pareto ansatz is held, then selection effects must be at play.

When selection effects are non-negligible, we must devise a procedure to disentangle them from the data if our ξ estimate is to be uncontaminated. To this end, we performed joint Bayesian inference with a model that includes both ξ *and* parameters that describe the selection effects.

In particular, we considered the roles of the fuzzy angular length threshold and surface brightness selection effects, introduced in Sects. 6.2.8 and 6.2.8, respectively. In Sect. 6.3.2, we explain that we have attempted to maintain a 5′ angular length threshold during our LoTSS DR2 GRG search. If we want to use Sect. 6.2.8 to correct for the bias against faraway and physically small giants that this threshold has imprinted onto our sample, we must estimate parameters φ_{min} and φ_{max}. A natural choice is to

[16] We explored two choices for z_{max} as our model assumes that ξ remains constant between $z \in [0, z_{\mathrm{max}}]$. We further discuss this assumption in Section 6.5.5.

assume that they lie symmetrically around the intended angular length threshold of $5'$ — but at what distance from it? We propose to consider this distance, $\frac{1}{2}(\varphi_{\rm max} - \varphi_{\rm min})$, as a yet unknown model parameter that we fitted to the data. Similarly, if we want to use Sect. 6.2.8 to correct for the bias against physically large giants that the limited depth of the LoTSS DR2 imprints onto our sample, we must estimate parameters $b_{\nu,\rm ref}$ and $\sigma_{\rm ref}$. Our approach was to, again, fit these parameters — possibly making use of any available prior knowledge.

The meaning of $b_{\nu,\rm ref}$ depends on the choice of $l_{\rm ref}$ and $\nu_{\rm obs}$; we defined $l_{\rm ref} := 0.7$ Mpc and used $\nu_{\rm obs} = 144$ MHz. We furthermore assumed $b_{\nu,\rm th} = 1 \cdot \sigma_{I_\nu}$, with $\sigma_{I_\nu} = 25$ Jy deg^{-2} being the typical LoTSS DR2 $6''$ RMS noise (Shimwell et al., 2022).

Prior

We exploited two sources of prior knowledge. Firstly, we attempted to directly estimate $b_{\nu,\rm ref}$ and $\sigma_{\rm ref}$ by selecting from all LoTSS DR2 giants with $l_{\rm p} \leq 1$ Mpc a random subset of size 50 (10%). For these giants, we estimated the mean surface brightnesses of both lobes from the LoTSS DR2 $6''$ imagery, differentiating between the brighter and the fainter lobe. Because our goal was to estimate $b_{\nu,\rm ref}$, we attempted to undo cosmological and growth-induced surface brightness dimming assuming a universal lobe spectral index $\alpha = -1$ and self-similar growth: $\zeta = -2$. The resulting surface brightnesses correspond to $z = 0$, and to the epoch in each giant's life when $l = l_{\rm ref}$. The median of the corrected bright lobe mean surface brightnesses is $b_{\nu,\rm ref} = 1.3 \cdot 10^3$ Jy deg^{-2}, whilst the median of the corrected faint lobe mean surface brightnesses is $b_{\nu,\rm ref} = 0.7 \cdot 10^3$ Jy deg^{-2}. All lobes taken together, the median becomes $b_{\nu,\rm ref} = 1.0 \cdot 10^3$ Jy deg^{-2}. We performed maximum likelihood estimation assuming the surface brightness distribution is lognormal and found $\sigma_{\rm ref} = 1.3$, again using all lobes. Note, however, that we have only used observed giants here, whilst $b_{\nu,\rm ref}$ and $\sigma_{\rm ref}$ should correspond to the entire population of giants. As fainter giants will have preferentially fallen out, we might have overestimated $b_{\nu,\rm ref}$ and underestimated $\sigma_{\rm ref}$.

To probe whether we had overestimated $b_{\nu,\rm ref}$ and underestimated $\sigma_{\rm ref}$, we used data from Dabhade et al. (2020b) and a Monte Carlo approach. First, for their sample of 239 LoTSS DR1 giants, we computed total luminosity densities L_ν at rest-frame frequency $\nu = 144$ MHz. (Given that LoTSS DR1 and DR2 noise levels are similar, this population is also representative of the LoTSS DR2.) In the top panel of Fig. 6.15, we show L_ν versus $l_{\rm p}$ for all 139 for which $z < 0.6$. To increase the range of lengths covered, we additionally show data on Alcyoneus (Oei et al., 2022a). Next, under the

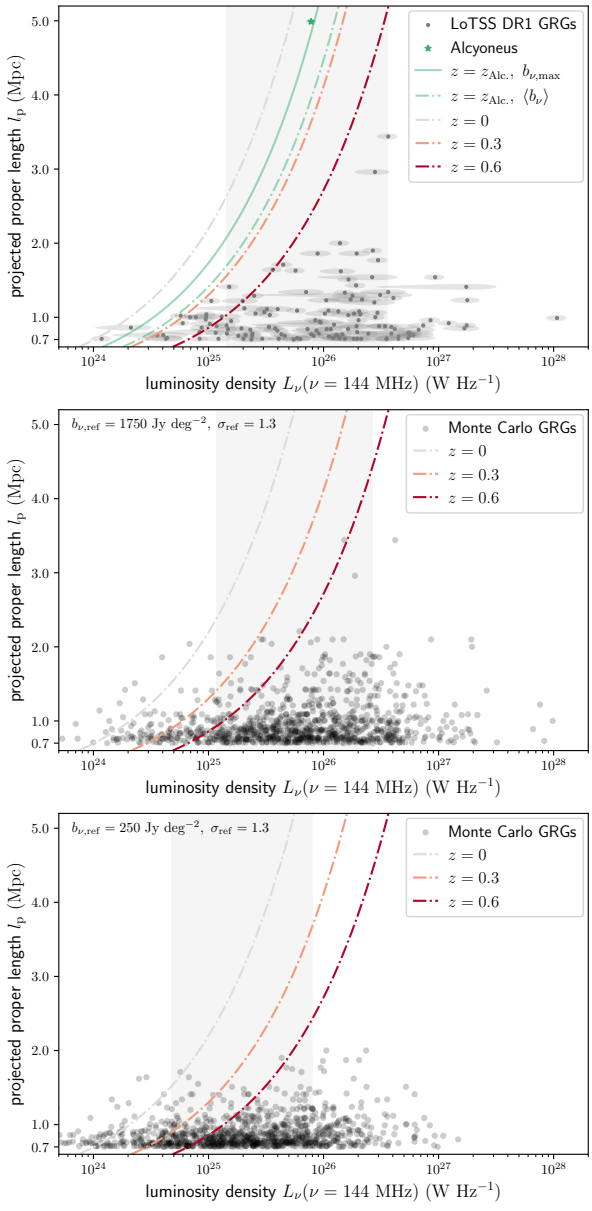


Figure 6.15: Comparison between luminosity density–projected proper length relations for observed and simulated giants. Each dash-dotted curve denotes a family of giants at a given redshift, assuming $f_{L_\nu} = 0.3$ and $f_l = 0.3$, whose mean lobe surface brightnesses equal the LoTSS DR2 noise level and who are thus borderline-detectable. The grey band denotes the median-centred luminosity density range that contains 68% of giants. *Top:* 139 LoTSS DR1 giants (Dabhade et al., 2020b), alongside Alcyoneus, with redshift $z_{\rm Alc.} = 0.25$ (Oei et al., 2022a). The solid green curve is similar to the dash-dotted green curve, but represents maximum instead of mean lobe surface brightness. *Middle:* 1000 simulated giants, assuming $b_{\nu,\rm ref} = 1750$ Jy deg^{-2} and $\sigma_{\rm ref} = 1.3$. *Bottom:* 1000 simulated giants, assuming $b_{\nu,\rm ref} = 250$ Jy deg^{-2} and $\sigma_{\rm ref} = 1.3$.

same assumptions of a constant spectral index and self-similar growth, we derived a simple luminosity density–surface brightness relationship that allows for back-and-forth conversion between the two — at least, given projected lengths and redshifts. The mean lobe surface brightness $\langle b_\nu \rangle$ is proportional to the total luminosity density L_ν, and assuming a pair of spherical lobes

$$\langle b_\nu \rangle = \frac{2 f_{L_\nu} \cdot L_\nu}{\pi^2 \cdot \mathbb{E}[D](\eta(f_l)) \cdot f_l^2 \cdot l_\mathrm{p}^2 \cdot (1+z)^{3-\alpha}}. \tag{6.40}$$

Here f_{L_ν} is the fraction of the total luminosity density that belongs to the lobes, f_l is the fraction of the RG's axis length that lies inside the lobes, and $\mathbb{E}[D](\eta(f_l))$ is the mean deprojection factor as given by Eq. 6.71. The peak surface brightness $b_{\nu,\mathrm{max}}$ relates to $\langle b_\nu \rangle$ as $b_{\nu,\mathrm{max}} = \frac{3}{2}\langle b_\nu \rangle$. Section 6.A4 contains derivations for both these results. For Alcyoneus, at $z_{\mathrm{Alc.}} = 0.25$, $f_{L_\nu} = 0.3$ and $f_l = 0.3$ (Oei et al., 2022a). Assuming these parameter values, again in the top panel of Fig. 6.15, we show luminosity density–projected length pairs of RGs at $z = z_{\mathrm{Alc.}}$ whose peak (solid green curve) or mean (dash-dotted green curve) surface brightness equals the LoTSS DR2 noise level. Thus, each curve represents a family of borderline-detectable giants at Alcyoneus's redshift. The other dash-dotted curves indicate similar barely detectable families, but for other redshifts. Without optimising any free parameters, the curves correctly predict that Alcyoneus's lobes have surface brightnesses comparable to the $6''$ LoTSS DR2 noise level and explain the absence of observations in the top-left corner of the figure. We conclude that Eq. 6.40 appears reasonable, but note that RGs may significantly differ in their values of f_{L_ν} and f_l.[17]

Bolstered, we made use of Eq. 6.40 to Monte Carlo simulate — for particular values of $b_{\nu,\mathrm{ref}}$ and σ_ref — luminosity density–projected length relationships as they appear to observers. The simulated giants have projected lengths adopted from the observed giants, randomly sampled redshifts up to $z = 0.6$ assuming a spatially constant GRG number density, and randomly sampled reference surface brightnesses (i.e. those for RGs at $z = 0$ that have intrinsic length $l = l_\mathrm{ref}$) whose distribution is determined by $b_{\nu,\mathrm{ref}}$ and σ_ref. We then used Eq. 6.24 to compute surface brightnesses as they would be observed, and retained only those giants whose surface brightness exceeds the LoTSS DR2 noise level. For these simulated detectable giants, we finally generated luminosity densities using Eq. 6.40, assuming wide uniform distributions $f_{L_\nu} \sim f_l \sim \mathrm{Uniform}(0.1, 0.9)$. The middle and bottom panels of Fig. 6.15 show results for $b_{\nu,\mathrm{ref}} = 1750\,\mathrm{Jy\,deg^{-2}}$ and $b_{\nu,\mathrm{ref}} = 250\,\mathrm{Jy\,deg^{-2}}$, respectively; we adopted

[17] This is also the reason that some giants in Fig. 6.15 cross their redshift's dash-dotted curve, which represents $f_{L_\nu} = 0.3$ and $f_l = 0.3$ only.

$\sigma_{\text{ref}} = 1.3$ from our LoTSS DR2 GRG surface brightness measurements. The median luminosity density of the $z < 0.6$ LoTSS DR1 giants is $L_\nu = 1.1 \cdot 10^{26}$ W Hz^{-1}, that of the $b_{\nu,\text{ref}} = 1750$ Jy deg^{-2} simulated giants is $L_\nu = 0.6 \cdot 10^{26}$ W Hz^{-1}, and that of the $b_{\nu,\text{ref}} = 250$ Jy deg^{-2} simulated giants is $L_\nu = 0.2 \cdot 10^{26}$ W Hz^{-1}.[18] Interestingly, the higher reference surface brightness median provides a better fit to the data. Even in case our Monte Carlo approach would predict luminosity densities that are biased low by a factor two, the higher median remains favoured.

In conclusion, it seems reasonable to suppose that our measurement $b_{\nu,\text{ref}} = 1.0 \cdot 10^3$ Jy deg^{-2} is not, or only mildly, biased high by selection effects. Still, we take a conservative approach and in setting priors we assume a 75% error on our measurement of $b_{\nu,\text{ref}}$ and a 50% error on our measurement of σ_{ref}. Thus, the priors for $b_{\nu,\text{ref}}$ and σ_{ref} — which we choose to be Gaussian — have 68% credible intervals $[250 \text{ Jy deg}^{-2}, 1750 \text{ Jy deg}^{-2}]$ and $[0.65, 1.95]$. We retain flat priors for ξ and $\frac{1}{2}(\varphi_{\text{max}} - \varphi_{\text{min}})$.

Inference

To compute the posterior distribution for ξ, $\frac{1}{2}(\varphi_{\text{max}} - \varphi_{\text{min}})$, $b_{\nu,\text{ref}}$ and σ_{ref}, we first brute-force evaluated the likelihood function over a regular grid that covers a total of 3.3 million parameter combinations.[19] For each proposed parameter quartet, we computed the PDF of $L_{\text{p,obs}} \mid L_{\text{p,obs}} > l_{\text{p,GRG}}$, and obtained the likelihood assuming that the LoTSS DR2 GRG projected proper lengths are IID draws from it. To obtain the PDF, we successively evaluated Eqs. 6.22, 8.15, 6.29, 8.12, 6.3, and 8.11, alongside their direct dependencies. This required the numerical evaluation of five integrals. Compared with using Riemann sums, we achieved substantial accuracy improvements at virtually no added numerical cost by approximating these integrals with the trapezoid rule and the composite Simpson's rule.

[18] For another interesting case, $b_{\nu,\text{ref}} = 500$ Jy deg^{-2} (not shown), the median $L_\nu = 0.3 \cdot 10^{26}$ W Hz^{-1}.

[19] The computation took a few thousand CPU hours to complete, but can be trivially distributed among nodes, and within a node among CPUs. Model extensions that introduce additional parameters shall necessitate more efficient inference techniques, such as Markov chain Monte Carlo.

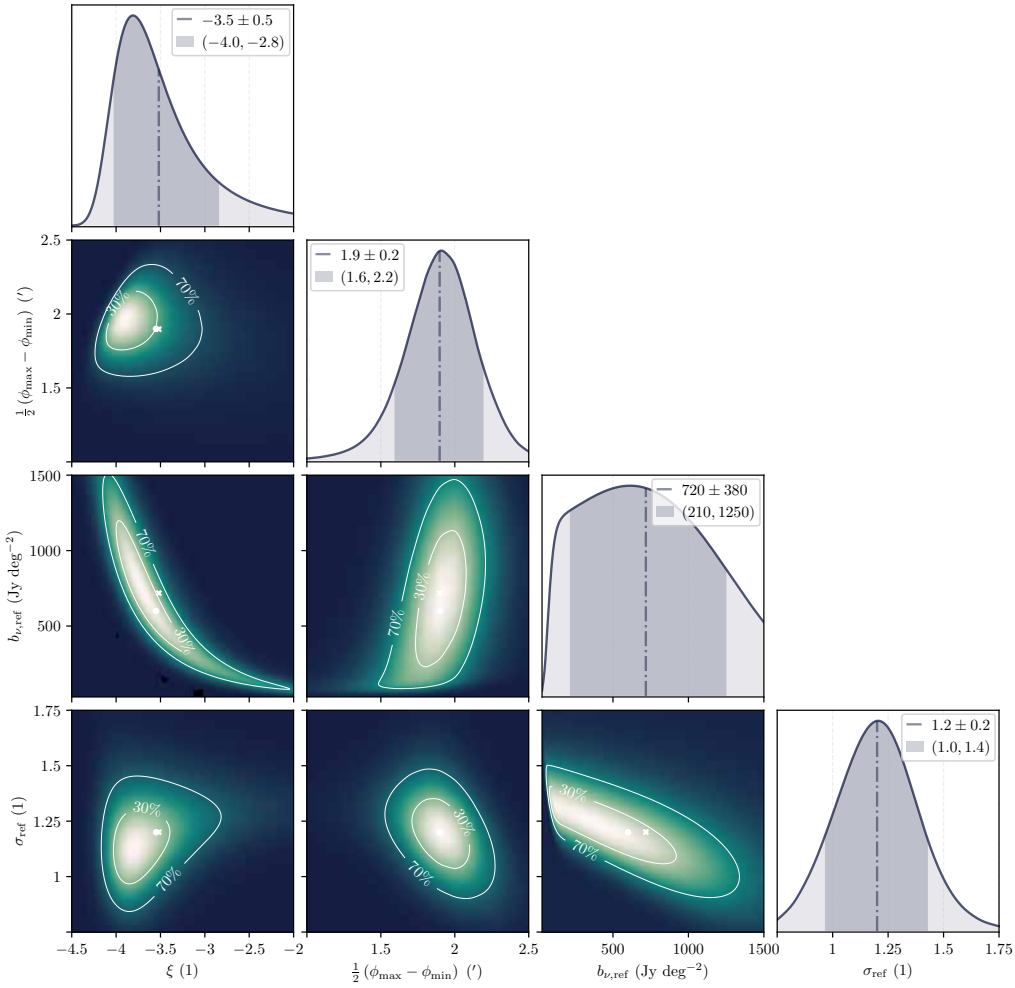


Figure 6.16: Joint posterior distribution over ξ — the parameter of interest — and $\frac{1}{2}(\varphi_{\max} - \varphi_{\min})$, $b_{\nu,\mathrm{ref}}$ and σ_{ref} — the selection effect parameters, based on 1473 projected lengths of LoTSS DR2 giants up to $z_{\max} = 0.5$. We show all two-parameter marginals of the posterior, with contours enclosing 30% and 70% of total probability. We mark the maximum a posteriori probability parameters (white circle) and the posterior mean parameters (white cross). The single-parameter marginals again show the estimated posterior mean, now marked by a vertical line, alongside shaded median-centred 80% credible intervals. To compare the posterior to the likelihood function, which is also the posterior for a uniform prior, see Fig. 6.28.

We summarise the likelihood function in Table 6.5 and Fig. 6.28. To obtain the posterior, we simply multiplied the likelihood function by the prior and normalised the result.

Table 6.4: Maximum a posteriori probability (MAP) and posterior mean and standard deviation (SD) estimates of the free parameters in intrinsic GRG length distribution inference.[20]

$z_{\mathrm{max}} = 0.5$:

parameter	MAP	posterior mean and SD
ξ	-3.55	-3.5 ± 0.5
$\frac{1}{2}(\varphi_{\mathrm{max}} - \varphi_{\mathrm{min}})$	$1.9'$	$1.9 \pm 0.2'$
$b_{\nu,\mathrm{ref}}$	$600\ \mathrm{Jy\ deg}^{-2}$	$720 \pm 380\ \mathrm{Jy\ deg}^{-2}$
σ_{ref}	1.2	1.2 ± 0.2

$z_{\mathrm{max}} = 0.25$:

parameter	MAP	posterior mean and SD
ξ	-3.5	-3.5 ± 0.4
$\frac{1}{2}(\varphi_{\mathrm{max}} - \varphi_{\mathrm{min}})$	$1.85'$	$1.7 \pm 0.3'$
$b_{\nu,\mathrm{ref}}$	$900\ \mathrm{Jy\ deg}^{-2}$	$1020 \pm 490\ \mathrm{Jy\ deg}^{-2}$
σ_{ref}	1.15	1.3 ± 0.4

In Table 6.4, for each parameter, we list the maximum a posteriori probability (MAP) estimate, alongside estimates for the posterior mean and standard deviation. In Fig. 6.16, we visualise all one- and two-dimensional posterior marginals, in which we mark the MAP (white circle) and the posterior mean (white cross). The joint marginal for ξ and $b_{\nu,\mathrm{ref}}$ shows that these parameters have a strong negative correlation, indicating that with current data, the steep slope of the ESF at high l_{p} can equally be described with a steep intrinsic slope and mild surface brightness selection (i.e. ξ low and $b_{\nu,\mathrm{ref}}$ high), or by a shallow intrinsic slope and strong surface brightness selection (i.e. ξ high and $b_{\nu,\mathrm{ref}}$ low). We leave it up to future studies to break this degeneracy, either by using larger samples, by measuring $b_{\nu,\mathrm{ref}}$ directly, or by improving survey sensitivities so that surface brightness selection effect modelling becomes superfluous altogether.

Goodness of fit

In both panels of Fig. 6.14, we compare the ESF and SF of $L_{\mathrm{p,obs}} \mid L_{\mathrm{p,obs}} > l_{\mathrm{p,GRG}}$ for $z_{\mathrm{max}} = 0.5$, using the MAP parameters for the latter. The model appears able to produce a tight fit to the data. The mean and standard deviation of the ESF–SF residuals are 0.01% and 0.3%, whilst the mean and standard deviation of the absolute ESF–SF residuals are both 0.2%. Using a Kolmogorov–Smirnov test, we formally verified that our *best* parameters are indeed *good* parameters — in the sense that they represent a plausible model underlying the data. The Kolmogorov–Smirnov statistic

[20] The model assumes ξ is constant for $z \in [0, z_{\mathrm{max}}]$. We determined the posterior twice: for $z_{\mathrm{max}} = 0.5$, using 1473 giants, and for $z_{\mathrm{max}} = 0.25$, using 811 giants.

is the maximum deviation between the ESF and SF, and equals 1% in our case. The p-value — the probability that an ESF–SF discrepancy of at least this magnitude would occur if the SF represents the true underlying distribution — is $p = 99\%$. For any reasonable significance level, we do not reject the null hypothesis. The model, given our best parameters, indeed represents a possible description of the data. We conclude that the distribution of GRG intrinsic proper lengths, after correcting for selection effects, is consistent with a single Pareto distribution with tail index $\xi = -3.5$. We show the SF of this distribution in both panels of Fig. 6.14 (fading green dots). For low l_p, the observed slope is shallower (due to angular length selection), whilst for high l_p, the observed slope is steeper (due to surface brightness selection).

A quasi-Pareto distribution can arise naturally as the tail of a lognormal distribution (e.g. Malevergne et al., 2011), and there are reasons to believe that the entire radio galaxy length distribution is indeed approximately lognormal (Oei et al., prepb). This provides an explanation of the approximately Paretian nature of the giant radio galaxy length distribution found in this section. The specific value of the tail index ξ is set by both the physics of radio galaxy growth and the distribution of radio galaxies over large-scale environments, the latter of which we measure in Oei et al. (prepb). Our result $\xi = -3.5 \pm 0.5$ is a new constraint for dynamical models such as those of Turner & Shabala (2015) and Hardcastle (2018).

6.4.2 Giant radio galaxy number density

If in addition to our discoveries, we know how many giants our search campaign has missed, then we can infer the true comoving GRG number density in the Local Universe. The posterior distribution over selection effect parameters $\frac{1}{2}(\varphi_\mathrm{max} - \varphi_\mathrm{min})$, $b_{\nu,\mathrm{ref}}$ and σ_ref induces a probability distribution over the search completeness function $C(l_\mathrm{p})$. $C(l_\mathrm{p})$ denotes the probability that a giant of projected proper length l_p in comoving space up to $z = z_\mathrm{max}$ is detected through the search. We first generated parameter samples from our posterior using rejection sampling, and then used each to calculate a $C(l_\mathrm{p})$ sample. We show the distribution over $C(l_\mathrm{p})$ for $z_\mathrm{max} = 0.5$ in Fig. 6.17. For small l_p, C is low as many giants drop out due to angular length selection; for large l_p, C is low as many giants drop out due to surface brightness selection. The completeness peaks around $l_\mathrm{p} \sim 2$ Mpc; however, even there the majority of giants remains undetected.

We inferred a probability distribution over the true comoving GRG number density n_GRG by combining Eqs. 8.12 and 6.30 with the LoTSS DR2 GRG catalogue and samples from Sect. 6.4.1's posterior. The resulting skewed distribution, with mean and SD $n_\mathrm{GRG} = 4.6 \pm 2.4\,(100\,\mathrm{Mpc})^{-3}$ and 80% credible interval 3.1–6.7 $(100\,\mathrm{Mpc})^{-3}$,

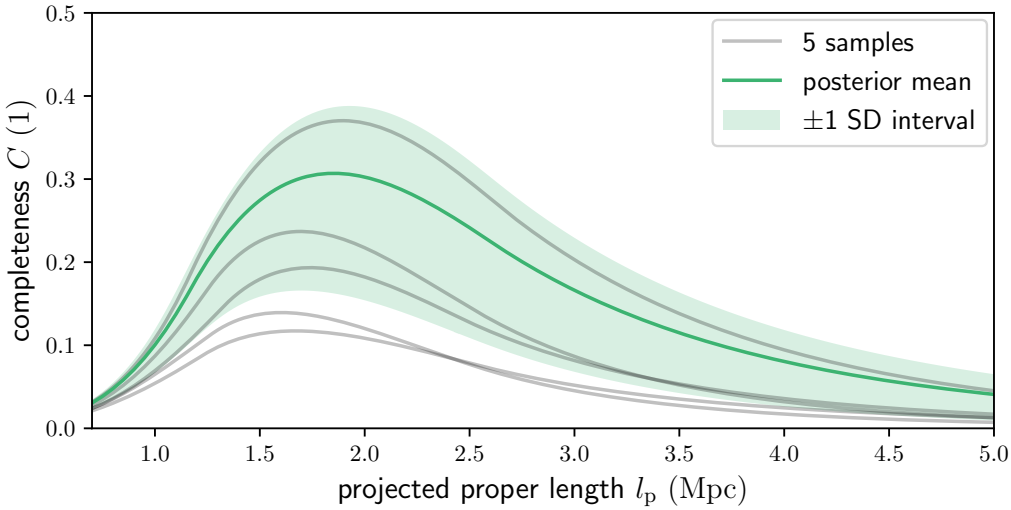


Figure 6.17: Completeness C of a sample of giant radio galaxies up to cosmological redshift $z_{\rm max}$ as a function of projected proper length $l_{\rm p}$. From samples of the posterior distribution, we infer the LoTSS DR2 GRG search campaign completeness up to $z_{\rm max} = 0.5$. We show completeness curves for five randomly selected samples (grey) and for the posterior mean (dark green). We also show an interval around the completeness mean with the completeness standard deviation (SD) as the half-width (light green). The completeness peaks around $l_{\rm p} = 2$ Mpc.

is shown in Fig. 6.18. We note that, although the uncertainty in $b_{\nu,\rm ref}$ induces a large uncertainty in C from $l_{\rm p} \sim 1.5$ Mpc onwards, the completeness uncertainty at large projected lengths does not substantially contribute to the uncertainty in $n_{\rm GRG}$. This is because the GRG population is dominated by smaller giants, for which the completeness appears better constrained.

What picture arises regarding the abundance of giant radio galaxies in the Local Universe's large-scale structure? If we model the Cosmic Web through comoving cubic unit cells (Oei et al., 2022b) with 50 Mpc sides, and each cubic unit cell contributes one cluster and three filaments, then a cube with 100 Mpc sides features a total of eight clusters and 24 filaments. For comparison, in a $(100\ \mathrm{Mpc})^3$ volume up to $z_{\rm max} = 0.5$, the SDSS-III cluster catalogue of Wen et al. (2012) contains on average 11.2 clusters of any mass, and 4.5 clusters of mass $M_{200} > 10^{14}\ M_\odot$. Since clusters contain $\sim$20% of giants (Oei et al., prepb), we find the average number of giants per cluster to be $\sim 10^{-1}$. If one assumes that filaments contain the remaining $\sim$80% of giants, and uses the fact that the average number of filaments per cluster is of order unity, it follows that the average number of giants per filament is also $\sim 10^{-1}$. In all likelihood, most clusters and filaments do not currently contain a giant.

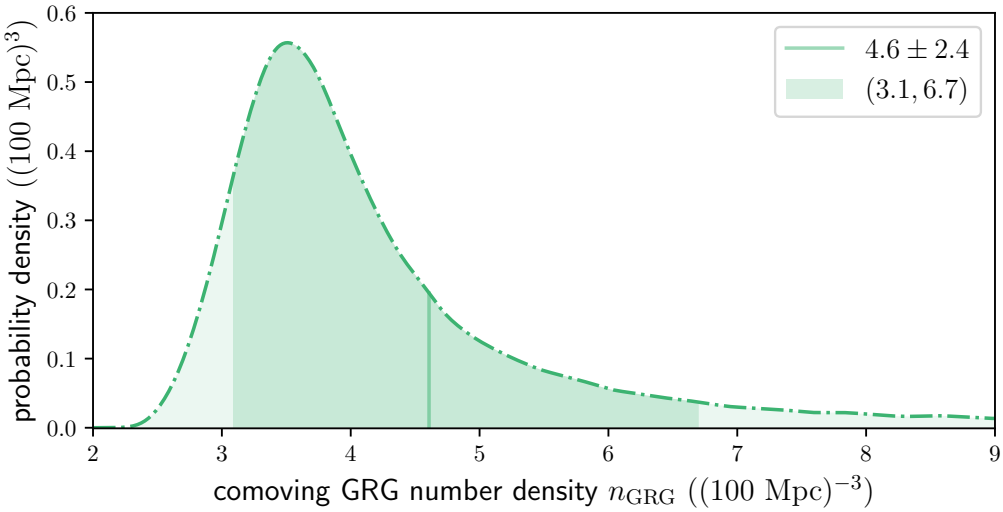


Figure 6.18: PDF of the comoving GRG number density $n_{\rm GRG}$. We mark the mean (vertical line) and the median-centred 80% credible interval (darker range). In the Local Universe, the average number of giants per comoving cube with 100 Mpc sides is 4.6 ± 2.4. We define giants through $l_{\rm p,GRG} = 0.7$ Mpc, and define the Local Universe to be up to cosmological redshift $z_{\rm max} = 0.5$.

6.4.3 Giant radio galaxy lobe volume-filling fraction

Because giant radio galaxies enrich the IGM with hot plasma and magnetic fields far beyond the circumgalactic media of their hosts, they may provide a meaningful contribution to the heating and magnetisation of — in particular — the most rarefied parts of the filament IGM. By combining the GRG number density and the GRG jet power distribution (e.g. Dabhade et al., 2020a), one could estimate the instantaneous heating and magnetisation contributions directly. We recommend such analysis for future research.

We evaluated Eq. 6.35 to obtain an estimate of the fraction of the Local Universe's proper volume that GRG lobes occupy. We used Alcyoneus as a reference giant, for which $V = 2.5 \pm 0.3$ Mpc^3 and $l_{\rm p} = 4.99 \pm 0.04$ Mpc (Oei et al., 2022a); this suggests $\mathbb{E}[\Upsilon_{\rm p}] \approx 2\%$. Future work should determine whether Alcyoneus's case is typical, as observations, such as those shown in Figs. 6.12–6.13, suggest that giants exhibit a large variety of shapes — and thus total lobe volume–cubed length ratios. Interestingly, simulations by Krause et al. (2012) have found that these shapes also depend on environmental parameters such as ambient pressure and density. Truncating the GRG projected length distribution at $l_{\rm p,max} = 7$ Mpc, so that its support is exactly an order of magnitude, Eq. 6.35 predicts $\mathrm{VFF}_{\rm GRG}(z=0) = 3^{+4}_{-1} \cdot 10^{-7}$.[21]

[21] We found the weaker constraints $\mathrm{VFF}_{\rm GRG}(z=0) > 13^{+21}_{-4} \cdot 10^{-8}$, $\mathrm{VFF}_{\rm GRG}(z=0) > 11^{+16}_{-3} \cdot$

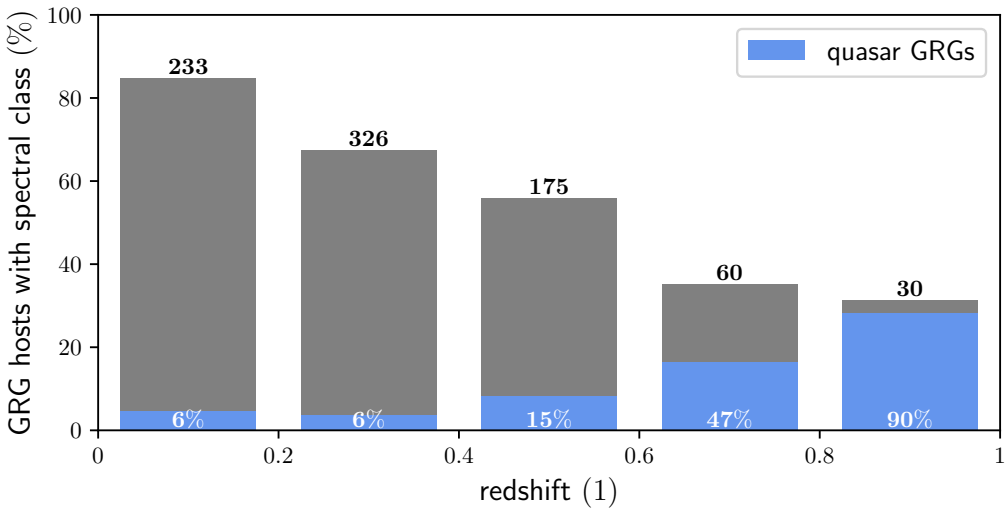


Figure 6.19: Binary classification into quasar and non-quasar giants based on SDSS DR12 host spectra, for 5 redshift intervals. We selected all giants with definite hosts within the SDSS-covered sky patch bounded by right ascensions 120° and 250°, and declinations 27° and 62°. As redshift increases, the fraction of hosts with a known spectral class decreases (grey bars, with the absolute number of such hosts in black), whilst the fraction of quasar identifications increases (blue bars and white percentages).

Whether this result is sensitive to changes in $l_{\mathrm{p,max}}$ depends on ξ, with $\xi = -3$ being a special value under self-similar growth. In that case, small and large giants contribute equally to $\mathrm{VFF_{GRG}}$: although large giants are rarer ($f_{L_\mathrm{p}} \propto l_\mathrm{p}^{-3}$), their larger lobe volumes ($V \propto l_\mathrm{p}^3$) exactly compensate. For $\xi < -3$, small giants provide the dominant contribution to $\mathrm{VFF_{GRG}}$ and the choice of $l_{\mathrm{p,max}}$ can be irrelevant; for $\xi > -3$, large giants dominate and the choice of $l_{\mathrm{p,max}}$ always matters.

If we assume that giants occur in clusters and filaments only and use the fact that clusters and filaments comprise about 5% of the Local Universe's volume (Forero-Romero et al., 2009), then the GRG lobe VFF within clusters and filaments specifically is $\mathrm{VFF_{GRG}}(z = 0) = 5^{+8}_{-2} \cdot 10^{-6}$. We conclude that, at each given moment, GRG lobes occupy just a small fraction of the WHIM and ICM. If the enrichment of the IGM by giants is to affect the WHIM and ICM on a large scale, mixing processes in the IGM are necessary and many galaxies must be able to form giants at some point in their evolution.

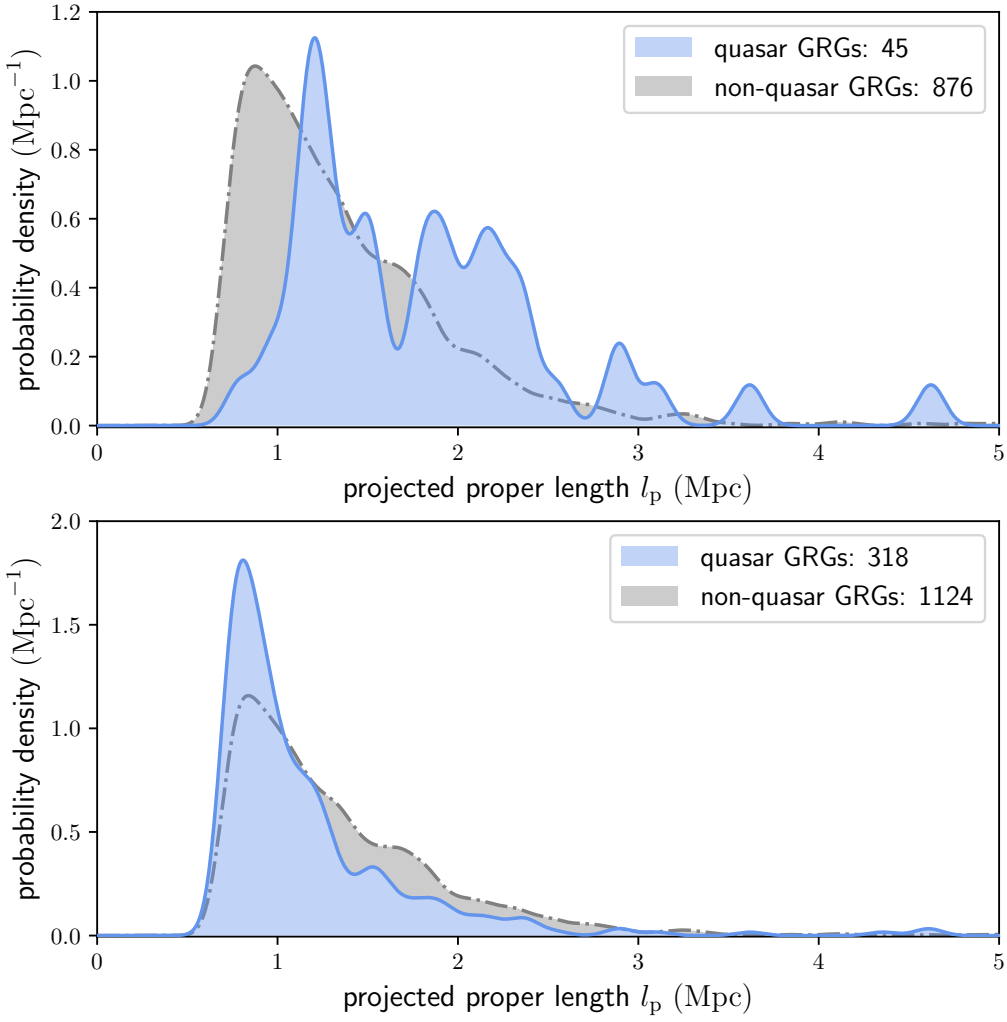


Figure 6.20: Observed projected proper length PDFs for SDSS-classified quasar and non-quasar giants, obtained through kernel density estimation. We used a Gaussian kernel with $\sigma_{\mathrm{KDE}} = 75$ kpc. For both panels, two-sample Kolmogorov–Smirnov tests yield $p < 1‰$. However, given the severe impact of selection effects, we could not reject the null hypothesis that quasar giants and non-quasar giants have the same underlying projected proper length distribution. *Top:* newly discovered (LoTSS DR2) giants with SDSS spectral class labels. *Bottom:* all known giants with SDSS spectral class labels.

6.4.4 Unification model constraints from quasar and non-quasar giants

In Section 6.2.11, we have predicted general ramifications of the basic unification model on a GRG sample. We constrained and tested this model with our LoTSS

10^{-8}, and $\mathrm{VFF}_{\mathrm{GRG}}(z=0) > 6^{+4}_{-1} \cdot 10^{-8}$ using Eqs. 6.33, 6.142, and 6.143, respectively.

DR2 GRG sample.

First, for all 3198 giants with definitively identified hosts, we queried the SDSS DR12 spectral class `spCl`: a Boolean label indicating whether or not the host contains a quasar. As many hosts have no SDSS DR12 spectrum, or even fall outside of SDSS DR12 coverage, we retrieved host classifications for just 1442 of these giants (45%). Of these classified giants, 318 are quasar giants (22%) (of which 45 (14%) are discoveries presented in this work) and 1124 are non-quasar giants (78%) (of which 876 (78%) are discoveries presented in this work). Therefore, the apparent LoTSS DR2 quasar GRG fraction $f_{\mathrm{Q}} = \frac{45}{45+876} = 5\%$. However, spectral class labels are preferentially available for GRG hosts with higher optical flux densities, such as those at low redshifts or those containing quasars, because they are more probable spectroscopic targets. Through Fig. 6.19, we demonstrate that the fraction of observed GRG hosts with spectral class labels indeed decreases with redshift, whilst the fraction of quasar identifications increases. For each redshift interval, in white, we denote the fraction of quasar giants within the classified population. If spectral class labels would have been available for the non-classified observed populations as well, the quasar GRG fractions would probably have been lower. Assuming that all observed giants whose hosts have an unknown spectral class are non-quasar giants, we find quasar GRG fractions $f_{\mathrm{Q}} = 5\%$, 4%, 8%, 16%, and 28%, for redshift intervals 0–0.2, 0.2–0.4, 0.4–0.6, 0.6–0.8, and 0.8–1, respectively. The true quasar GRG fraction for a given redshift interval might still differ from the aforementioned *observed* quasar GRG fraction: namely, if selection effects make a given quasar GRG easier (or harder) to detect than a given non-quasar GRG. At higher redshifts, quasar giants certainly appear easier to detect than non-quasar giants, as hosts without quasars often become too faint to optically identify. For the lowest redshift intervals, this problem does not exist, and we therefore consider the observed quasar GRG fraction $f_{\mathrm{Q}} = 5\%$ to be closest to the true one. A thorough analysis of the impact of selection effects on f_{Q} is a topic for future research.

Using Eq. 6.37, and assuming a quasar GRG probability $p_{\mathrm{Q}} = 5\%$, we found a maximum inclination angle $\theta_{\mathrm{max}} = 39^{+2^\circ}_{-3^\circ}$. For $p_{\mathrm{Q}} \sim \mathrm{Uniform}(4\%, 6\%)$, the result remained the same. In conclusion, if the basic unification model considered in Sect. 6.2.11 is correct, then observations of giants predict that quasars are AGN seen along lines-of-sight that make an angle of at most $\theta_{\mathrm{max}} = 39^{+2^\circ}_{-3^\circ}$ with the black hole rotation axis.

Finally, we tested whether the RV $L_{\mathrm{p,obs}} \mid L_{\mathrm{p,obs}} \geq l_{\mathrm{p,GRG}}$ has the same distribution for quasar and non-quasar giants, as predicted by the unification model. In the top panel of Fig. 6.20, we show PDFs approximated through kernel density estimation (KDE) for newly discovered (LoTSS DR2) SDSS-classified quasar giants and non-

quasar giants up to $z_{\max} = \infty$. Despite the small number of quasar giants, $N_Q = 45$, the two-sample Kolmogorov–Smirnov (KS) test yielded a low p-value $p < 1‰$: we rejected the null hypothesis that these distributions stem from a single underlying one. However, this does not mean that the unification model hypothesis should immediately be rejected as well, because the selective availability of SDSS DR12 spectral class labels induces a severe selection effect.[22] This effect can be tempered by choosing a lower $z_{\max}$, but for choices such as $z_{\max} = 0.5$ and 0.25, just $N_Q = 16$ and 10 quasar giants remain — too few to extract meaningful information. The bottom panel of Fig. 6.20 again shows observed projected proper length KDE PDFs, but now for *all* SDSS-classified quasar giants and non-quasar giants up to $z_{\max} = \infty$; in this case, $N_Q = 318$. Again, the two-sample KS test yields $p < 1‰$, but this time quasar giants appear smaller than non-quasar giants. Because we aggregated samples here that have different selection effects imprinted, it becomes hard to draw clear conclusions. Limiting $z_{\max}$ reduces the severity of most selection effects, but of course comes at the cost of reducing the sample size. Interestingly, the corresponding quasar GRG and non-quasar GRG projected length distributions do become more alike; for instance for $z_{\max} = 0.5$ and 0.25, the two-sample KS test yields $p = 1\%$ and 4%, respectively. In conclusion, because quasar giants are intrinsically rare and the availability of spectral class labels is biased towards low redshifts and hosts containing quasars, it is challenging to robustly test the unification model with current GRG observations. We refrain from drawing final conclusions, and recommend a careful future analysis.

6.5 Discussion

6.5.1 Radio galaxy length definitions

How large are radio galaxies? Despite the simplicity of this question and more than half a century of research on radio galaxies, their intrinsic length distribution has not yet been rigorously characterised. In this work, we have carried out the first precision analysis of the tail of the radio galaxy intrinsic length distribution. Precision analyses tend to raise questions; firstly, whether the studied observable is well defined, and

[22] In fact, the top panel of Fig. 6.20 shows expected behaviour for a GRG search campaign with a (fuzzy) angular length threshold selection effect if the unification model is correct. At high redshifts, a GRG must have a larger projected length to pass the angular length threshold than at low redshifts. Thus, sampled high-redshift giants are physically larger than sampled low-redshift giants. Because the fraction of high-redshift quasar giants that is detectable and spectrally classifiable is higher than the fraction of high-redshift non-quasar giants that is detectable and spectrally classifiable, sampled quasar giants will be physically larger than sampled non-quasar giants. To draw this latter conclusion, we must also invoke the fact that, under the unification model, the projected length distributions of quasar and non-quasar giants are the same.

secondly whether one could conceive of more informative observables — that is to say those that make it easier to reveal underlying physical mechanisms. This work's main observable is the radio galaxy projected proper length; we argue that it is neither well defined nor maximally informative.

The current length definition: survey-dependence

Contemporary research uses a survey-dependent definition for radio galaxy angular lengths, which then makes projected proper lengths survey-dependent too.

The angular length is canonically defined as the largest possible angular separation between two directions for which the RG's specific intensity function $I_{\nu,\mathrm{RG}}$ exceeds $b_{\nu,\mathrm{th}}$: some specified factor of order unity times the image noise σ_{I_ν}. A complication is that not only $I_{\nu,\mathrm{RG}}$, but also σ_{I_ν} varies with observing frequency; the latter because of observational factors such as (u, v)-coverage, bandpass, radio-frequency interference (RFI), ionospheric weather, the sky density of bright calibrators and the performance of calibration algorithms. $I_{\nu,\mathrm{RG}}$ additionally depends on resolution, at least for point sources; σ_{I_ν} additionally depends on resolution and integration time. As a result, both $I_{\nu,\mathrm{RG}}$ and the concrete value of $b_{\nu,\mathrm{th}}$ used in the angular length definition change from image to image. Each study thus far has therefore implicitly used a different definition for angular length, instead of a shared, absolute one. As the projected proper length follows from combining the angular length with the host redshift, it suffers from the same problem.

Whether the survey-dependence of the current length definition is problematic, depends on the radio galaxy. For archetypal FRII RGs, the angular length is roughly equal to the angular distance between the hotspots. These are an FRII RG's brightest morphological components, and are thus the first to be picked up by a survey. In contrast, archetypal FRI RGs, which gradually fade with distance from the host, can have significantly larger angular lengths in surveys of higher sensitivity. The giants in the middle-right and bottom-left panel of Fig. 6.13 are good examples: these radio galaxies were known before the LoTSS DR2, but were not known to be giants; similarly, more sensitive surveys are poised to assign them even larger extents. If we are to move towards precision science, it therefore makes sense — at least for FRI RGs — to more explicitly recognise that the angular and projected proper lengths are functions of the observing frequency ν_{obs} and a surface brightness threshold $b_{\nu,\mathrm{th}}$. If catalogues would explicitly state for what combination $(\nu_{\mathrm{obs}}, b_{\nu,\mathrm{th}})$ they provide $\varphi = \varphi\,(\nu_{\mathrm{obs}}, b_{\nu,\mathrm{th}})$ and $l_{\mathrm{p}} = l_{\mathrm{p}}\,(\nu_{\mathrm{obs}}, b_{\nu,\mathrm{th}})$, it is possible to homogenise a collection of data sets by using universal angular and projected proper length definitions for all RGs.[23] If two angular

[23] Our LoTSS DR2 GRG φ and l_{p} correspond to $b_{\nu,\mathrm{th}} := 1 \cdot \sigma_{I_\nu}$; on average, $\sigma_{I_\nu}(\nu_{\mathrm{obs}} =$

lengths have been measured for the same RG, for instance φ_1 at $(\nu_{\rm obs,1}, b_{\nu,\rm th,1})$ and φ_2 at $(\nu_{\rm obs,2}, b_{\nu,\rm th,2})$, then we can estimate φ for any desired $(\nu_{\rm obs}, b_{\nu,\rm th})$ through interpolation. For example, the interpolation formula for a symmetric radio galaxy with jets or lobes of constant spectral index α, and with a specific intensity function contribution which fades to zero linearly with angular distance from the host, is

$$\varphi = \max\left\{\varphi_1 + \frac{b_{\nu,\rm th} - b_{\nu,\rm th,1}\left(\frac{\nu_{\rm obs}}{\nu_{\rm obs,1}}\right)^{\alpha}}{b_{\nu,\rm th,2}\left(\frac{\nu_{\rm obs}}{\nu_{\rm obs,2}}\right)^{\alpha} - b_{\nu,\rm th,1}\left(\frac{\nu_{\rm obs}}{\nu_{\rm obs,1}}\right)^{\alpha}}\left(\varphi_2 - \varphi_1\right), 0\right\}. \tag{6.41}$$

The ideal length definition: physical relevance

Alcyoneus, shown in Fig. 6.12's top-right panel, is a 5 Mpc giant whose ageing lobes are revealed for the first time by the LoTSS DR2 (Oei et al., 2022a). Future image sensitivity improvements shall reveal more hitherto unseen, fading lobes around known RGs that formed in the aftermath of earlier AGN activity episodes. Could a large fraction of sufficiently old RGs turn out to be giants, once such sensitivity improvements start providing evidence of earlier and earlier AGN activity episodes?

Up to now, it has been informative to include all visible plasma in the angular length measurement. In future images, some of the visible plasma might be of such low pressure that it has become physically insignificant, in the sense that it does not affect the thermodynamics of the surrounding IGM anymore; we propose to exclude such plasma from a radio galaxy length. Practically, one option is to introduce an absolute threshold: for example to include plasma of pressure $P \geq 10^{-17}$ Pa only — this is the pressure of the warm–hot intergalactic medium (WHIM) in the $\rho_{\rm BM} \sim 10\ \Omega_{\rm BM,0}\rho_{\rm c,0}$ and $T \sim 5 \cdot 10^5$ K regime. Another option is to introduce an environment-dependent pressure threshold; this would mean that a cluster RG sees its length measured against a higher pressure threshold than a filament RG, because its plasma becomes thermodynamically irrelevant sooner. A problem with (equipartition or minimum energy) pressure–based length definitions is that pressure is a derived quantity: the specific intensity only determines the product of pressure and line-of-sight length through the lobe.

6.5.2 Moving beyond line segment projection

In this work, we have adopted a classical approach to treating projection, by modelling radio galaxies as line segments. When a radio galaxy's inclination angle θ is close to

144 MHz, $\theta_{\rm FWHM} = 6''$) $= 25$ Jy deg^{-2}.

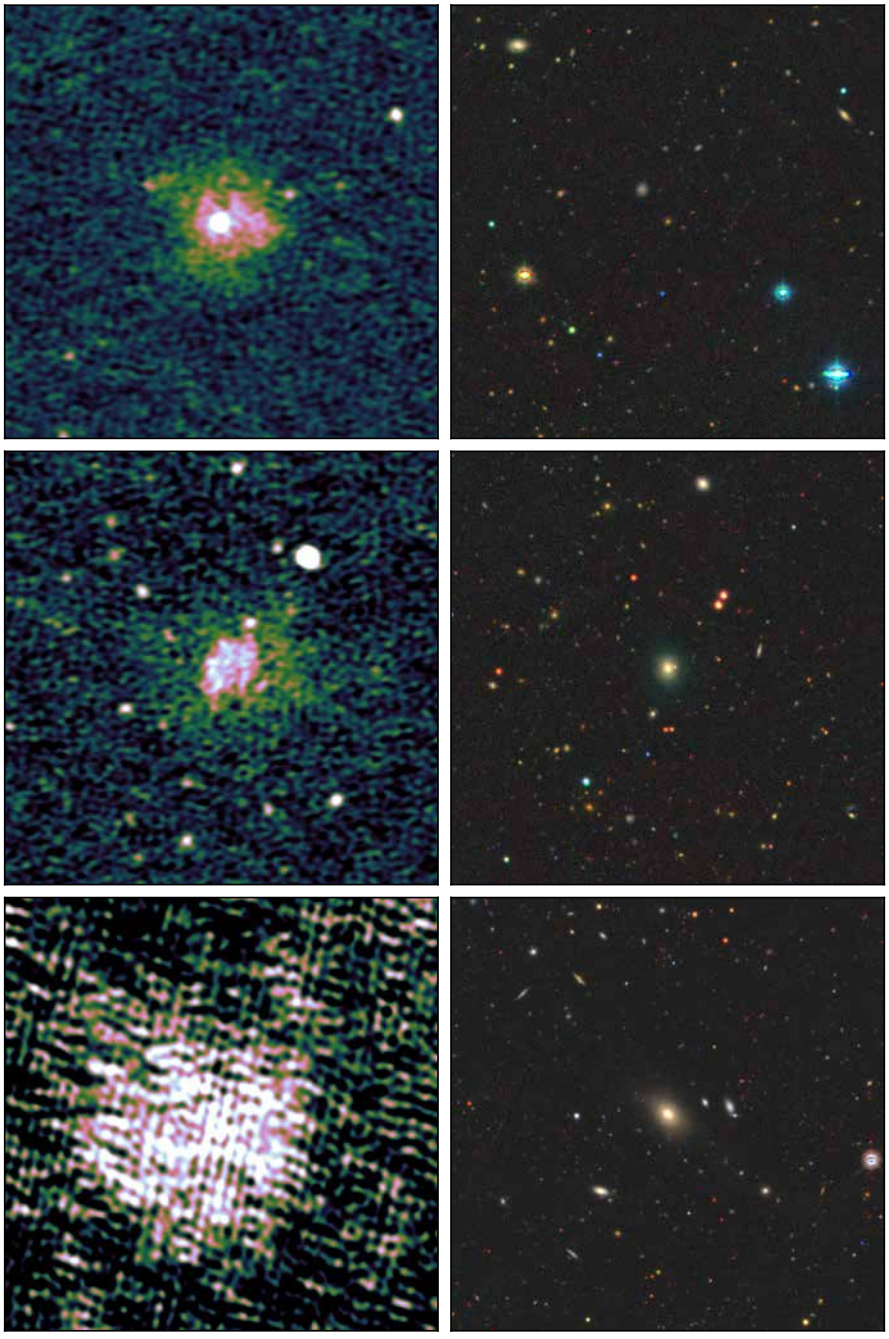


Figure 6.21: Suspected line-of-sight RGs in the radio and optical. We show $5' \times 5'$ LoTSS DR2 $6''$ cutouts *(left column)* and corresponding DESI Legacy Imaging Surveys (g, r, z)-cutouts *(right column)*. Due to its lobes, an RG whose axis aligns closely to the line of sight has a non-zero projected length. From top to bottom row, the SDSS host names are J125027.47+642034.3, J092220.72+560234.9, and J023100.61+032922.1.

0° or 180°, this approach predicts that the angular length φ and thus the projected proper length l_{p} vanishes. In reality however, because radio galaxy lobes have non-zero volumes, the RV L_{p} will never tend to zero. Figure 6.21 illustrates this point through three sources interpreted to be radio galaxies aligned closely with the line of sight. In each case, φ remains of arcminute scale.

A more realistic approach, outlined in Appendix 6.A1.3, moves beyond the simplistic line segment geometry by adding two lobes of radius R to the endpoints of the line segment, whose length is given by RV L. The ratio $\frac{L}{L_{\mathrm{p}}}$, which was unbounded under the classical approach, is now at most $1 + \frac{1}{\eta}$, where $\eta := \frac{2R}{L}$. (The classical approach simply is the limit $\eta = 0$.) We suggest a recalculation of this work's theoretical and applied results under this more realistic RG geometry as a direction for future research. To obtain the applied results, one must either fix η as a hyperparameter, or include it as an additional model parameter. We note that even using conservatively low values of η, such as $\eta = \frac{1}{10}$, will represent an improvement in realism over $\eta = 0$.

6.5.3 Unmodelled selection effects

In this work, we have modelled both an angular length and a surface brightness selection effect. Several other plausible selection effects have not been included in the forward model, as we have judged each to be of minor importance. However, in unison, they could have a non-negligible influence on the distribution of the observed projected proper GRG length RV $L_{\mathrm{p,obs}} \mid L_{\mathrm{p,obs}} > l_{\mathrm{p,GRG}}$. Some of their influence might have been absorbed by the parameters of the *included* selection effects — that is by $\frac{1}{2}(\varphi_{\mathrm{max}} - \varphi_{\mathrm{min}})$, $b_{\nu,\mathrm{ref}}$, and σ_{ref} — or, worse still, by ξ.

One of these unmodelled selection effects is that RGs whose axes are oriented almost parallel to the line of sight are more likely to be rejected from a sample than RGs whose axes are closer to the plane of the sky, as the former do not always have a characteristic double-lobe appearance. For instance, perhaps not all readers would regard our identification of the sources in the left column of Fig. 6.21 as RGs convincing. Nevertheless, as remarked in Sect. 6.2.11, conditioning $L_{\mathrm{p,obs}} \mid L_{\mathrm{p,obs}} > l_{\mathrm{p,GRG}}$ on inclination angle does not affect its distribution. As a result, this selection effect does not necessitate forward model modifications.

Furthermore, as shown in Sect. 6.4.4, there is a selection effect at play that favours the selection of quasar giants over non-quasar giants at high redshift, as host galaxies with quasars are more luminous in the optical and therefore have a better chance to be picked up in optical imagery. This selection effect will get less severe once deeper photometric surveys become available.

We encountered three more selection effects during our LoTSS DR2 GRG search.

The larger an RG — and especially an FRII RG — becomes, the harder it is for an observer to identify its host galaxy, as an increasing number of plausible host candidates can lie interspersed in the strip of sky between the lobes. The severity of this effect, which diminishes the prevalence of the largest giants in a sample, depends on the balance chosen between avoiding false discoveries and avoiding rejections of true discoveries. Another selection effect runs against RGs in galaxy clusters. Such environments can contain multiple adjacent RGs, making it at times unclear which lobe belongs to which RG. When no confident double-lobe associations can be made, the RGs involved fail to make it into the sample. If galaxy clusters contain primarily smaller giants, this selection effect induces a bias against smaller giants. A final unmodelled bias runs against RGs at the end of their life cycle. Once the AGN stops launching jets for a prolonged period, it becomes hard to identify the host galaxy, which will no longer present as a bright, compact radio source. This effect preferentially deselects larger giants, which are even more likely to approach the end of their life than smaller giants.

6.5.4 Does the choice of prior matter?

We have taken a conservative approach to constraining the posterior distribution through the prior: we have left tail index ξ and angular length selection half-width $\frac{1}{2}(\varphi_{\max} - \varphi_{\min})$ fully unconstrained, and have adopted wide Gaussian priors for reference surface brightness parameters $b_{\nu,\mathrm{ref}}$ and σ_{ref}, despite measuring them explicitly under assumptions. We provide summary statistics of the posterior in Table 6.4 and visualise its one- and two-parameter marginals in Fig. 6.16. Does our choice of prior significantly affect the inferences? To explore this question, we chose a different reasonable prior and compared results. One such prior is the fully uniform prior, which equivalises the posterior and the likelihood function. We provide analogous summary statistics of the likelihood function in Table 6.5 and visualise analogous marginals in Fig. 6.28. Reassuringly, no statistically significant parameter changes occur. In particular, for $z_{\max} = 0.5$, $\xi = -3.5 \pm 0.5$ becomes $\xi = -3.4 \pm 0.5$ upon changing to the uniform prior; for $z_{\max} = 0.25$, $\xi = -3.5 \pm 0.4$ even remains the same. However, given the strong likelihood degeneracy between ξ and $b_{\nu,\mathrm{ref}}$ apparent in Fig. 6.28, more stringent priors on $b_{\nu,\mathrm{ref}}$ *are* able to meaningfully shift ξ's posterior mean. Such priors shall be appropriate only after studying the surface brightness properties of large radio galaxies — and the associated selection effect — in more detail.

6.5.5 Cosmological evolution of the GRG length distribution

In this work we have assumed that the parameter ξ, which fully characterises the intrinsic GRG length distribution under the ansatz of Paretianity, remains constant throughout the chosen redshift range $[0, z_{\max}]$. To test this assumption, we have analysed our LoTSS DR2 GRG sample in Sect. 6.4 up to both $z_{\max} = 0.5$ and $z_{\max} = 0.25$. For $z_{\max} = 0.5$, we included giants that existed in the last 5 Gyr of the Universe's history, and found $\xi = -3.5 \pm 0.5$. Meanwhile, for $z_{\max} = 0.25$, we included giants that existed in the last 3 Gyr of the Universe's history only, and found the very similar $\xi = -3.5 \pm 0.4$. Thus, our analysis did not produce evidence that ξ evolves over cosmic time. However, given the large error bars, a modest time evolution cannot be excluded. Furthermore, the data sets that underlie these inferences are not disjoint: the 811 giants that inform the lower-maximum-redshift analysis make up 55% of the 1473 giants that inform the higher-maximum-redshift analysis.

Whether a time evolution of ξ is expected is presumably tied to whether giant radio galaxy growth varies with environmental density at a given epoch, because the combined effects of the Universe's expansion and ongoing large-scale structure formation can similarly change environmental density. Combining our LoTSS DR2 GRG sample with Cosmic Web reconstructions to explore giant growth as a function of environmental density is the topic of a forthcoming work (Oei et al., prepb). Interestingly, the recent exploration by Lan & Prochaska (2021) that compared the environments of giants and non-giants did not find significant differences.

The most straightforward model extension is to again assume that RG lengths are Pareto distributed with tail index ξ, but now $\xi = \xi(z)$. This function's first-degree Maclaurin polynomial, which provides the linearisation at the present day, is

$$\xi(z) \approx \xi(z = 0) + \frac{\mathrm{d}\xi}{\mathrm{d}z}(z = 0) \cdot z. \tag{6.42}$$

One would adopt $\xi(z = 0)$ and $\frac{\mathrm{d}\xi}{\mathrm{d}z}(z = 0)$ as the parameters of interest, replacing what used to be a constant ξ; the number of model parameters would thus increase by one. However, an attempt to infer the cosmic evolution of ξ appears promising only once the major selection effects are better constrained.

6.6 Conclusions

In this work, we have performed Bayesian inference on a LoTSS-derived sample of 2060 giant radio galaxy projected proper lengths, using a one-parameter model that assumes a spatially homogeneous, non-evolving population of radio galaxies with

stick-like geometry, Pareto-distributed lengths, and isotropic inclination angles. Before fitting to data, we extended the forward model with two selection effects typical of contemporary manual GRG search campaigns. The best-fit survival function tightly reproduces the empirical one, leaving permille-scale absolute residuals. Having quantified the most important selection effects, we estimated the true comoving giant radio galaxy number density in the Local Universe.

1. We developed an analytical model through which statistical questions about radio galaxy (RG) lengths can be rigorously answered. In the current work, we applied this model to giant radio galaxies. We adopted the ansatz that the RG intrinsic proper length L, as measured in three spatial dimensions, is a random variable (RV) with a Pareto Type I distribution (i.e. a simple power-law distribution) characterised by tail index ξ. Next, by assuming that RGs have no preferential orientation with respect to the observer, we derived the distribution of the RG projected proper length L_{p}. By conditioning, one obtains the version relevant for giants, $L_{\mathrm{p}} \mid L_{\mathrm{p}} > l_{\mathrm{p,GRG}}$ (where we chose $l_{\mathrm{p,GRG}} := 0.7$ Mpc). This RV is again Paretian, with the same tail index ξ. In summary, for giant radio galaxies, projection retains Paretianity. Finally, observers face selection effects; we modelled the observed projected proper length $L_{\mathrm{p,obs}}$ by considering an angular length threshold selection effect and a surface brightness selection effect. The angular length threshold selection effect assumes a linearly increasing angular-length-dependent probability of sample inclusion around a particular pre-defined threshold, meant to emulate manual visual searches that only target RGs of some angular length and above. The surface brightness selection effect assumes that giants are self-similar, and that their lobes have lognormally distributed surface brightnesses which must be above-noise to secure sample inclusion. The GRG observed projected proper length $L_{\mathrm{p,obs}} \mid L_{\mathrm{p,obs}} > l_{\mathrm{p,GRG}}$ again follows through conditioning. We assumed our data to be realisations of this RV.

2. The model also yielded explicit expressions for the (posterior) distribution of $L \mid L_{\mathrm{p}} = l_{\mathrm{p}}$. This allows one to deproject RGs in a statistical sense, providing the intrinsic proper length given the projected proper length in the limit of negligible selection effects. We also present practical expressions for the mean and variance of $L \mid L_{\mathrm{p}} = l_{\mathrm{p}}$. To unravel the driving factors that allow some RGs to become giants, most authors search for correlations between host or environmental physical parameters and the GRG *projected* length. However, if there is a causal chain that connects host or environmental parameters to GRG length, the connection will be to the *intrinsic* length; the observer's van-

tage point does not play a role in the physics. Therefore, the projection effect merely serves as a multiplicative noise source. We suggest that future analyses should recognise the projection effect as such, and correlate host or environmental parameters with the intrinsic, rather than projected, proper length — using statistical deprojection.

3. Through a manual visual search of the LoTSS DR2 pipeline products, the latest version of the LOFAR's Northern Sky survey at 144 MHz, we discovered a population of 2060 previously unknown giants. This is the largest single contribution to the literature yet, and increases the community-wide census by a factor 2.6. We present 11 discoveries with $l_\mathrm{p} \geq 4$ Mpc, 53 with $3 \leq l_\mathrm{p} < 4$ Mpc, 291 with $2 \leq l_\mathrm{p} < 3$ Mpc, 1215 with $1 \leq l_\mathrm{p} < 2$ Mpc, and 490 with $0.7 \leq l_\mathrm{p} < 1$ Mpc. Our study extends the known breadth of the giant radio galaxy phenomenon. Among the findings are both the giant hosted by J081956.41+323537.6 and Alcyoneus (Oei et al., 2022a), at $l_\mathrm{p} = 5.1$ Mpc and $l_\mathrm{p} = 5.0$ Mpc the projectively largest giants ever found. We discover that multi-Mpc radio galaxies can be generated before redshift 1, despite the Universe's mean density being an order of magnitude higher, and by spiral galaxies, whose stellar masses are typically an order of magnitude lower than those of ellipticals. We discover giants whose hosts have a record-low stellar mass $M_\star = 4.8 \cdot 10^{10}\ M_\odot$. We also discover giants whose hosts have a record-high supermassive black hole mass $M_\bullet \gtrsim 5 \cdot 10^{10}\ M_\odot$; interestingly, with $l_\mathrm{p} = 0.8$ Mpc, one of these giants is relatively small. We more than double the number of known giants with angular lengths exceeding that of the Moon; one discovery, at 2.2°, is the angularly largest radio galaxy in the Northern Sky and the angularly largest giant overall. Excitingly, our LoTSS DR2 search has been far from exhaustive: many thousands of readily identifiable giants still await discovery in this public data set.

4. Using our LoTSS DR2 GRG sample up to $z_\mathrm{max} = 0.5$, we generated a posterior distribution over ξ and three selection effect parameters. Our model provides an excellent fit to the data, with absolute residuals being on average 2‰. We inferred that the intrinsic proper length distribution of the largest radio galaxies resembles a Pareto distribution with tail index $\xi = -3.5 \pm 0.5$. Our analysis did not yield evidence for an evolving ξ in the last 5 Gyr of cosmic time.

5. The selection effect parameters estimated through the posterior are far from nuisance parameters, as they allowed us to statistically undo the selection effects imprinted on our LoTSS DR2 GRG data. As a result, we could for the

first time estimate the true comoving giant radio galaxy number density $n_{\rm GRG}$ in the Local Universe up to $z_{\rm max}$. We relied on the crucial assumption that the surface brightness distribution of RGs with intrinsic length $l_{\rm ref} = 0.7$ Mpc at $z = 0$ and frequency $\nu_{\rm obs} = 144$ MHz is unimodal — and lognormal in particular. We furthermore assumed a lobe spectral index $\alpha = -1$ and self-similar growth. We found $n_{\rm GRG}(l_{\rm p,GRG} = 0.7\ {\rm Mpc}, z_{\rm max} = 0.5) = 5 \pm 2\ (100\ {\rm Mpc})^{-3}$. The implication is that giant radio galaxies are truly rare — not only from a current observational perspective, but also from a cosmological one. Current GRG lobes occupy just a few millionths of the IGM volume. At any given moment in time, most clusters and filaments — the building blocks of modern large-scale structure — do not harbour giants.

Giants embody the most extreme known mechanism by which galaxies can affect the Cosmic Web around them. Whereas this work has explored the geometric properties of giants, a thorough exploration of their Cosmic Web energisation and magnetisation potential is a future frontier. Excitingly, the interactions between giants and the ethereal intergalactic medium may also allow for new constraints on the thermodynamics in filaments.

M.S.S.L. Oei warmly thanks Frits Sweijen for coding the very useful https://github.com/tikk3r/legacystamps and for tireless ICT advice, and Aleksandar Shulevski for comments that have improved the manuscript. In dear memory of Nicoline. By staying strong despite life's challenges, you showed what it means to be a giant.
M.S.S.L. Oei, R.J. van Weeren, and A. Botteon acknowledge support from the VIDI research programme with project number 639.042.729, which is financed by the Dutch Research Council (NWO). A. Drabent acknowledges support by the BMBF Verbundforschung under the grant 05A20STA.
The LOFAR is the Low-Frequency Array designed and constructed by ASTRON. It has observing, data processing, and data storage facilities in several countries, which are owned by various parties (each with their own funding sources), and which are collectively operated by the ILT Foundation under a joint scientific policy. The ILT resources have benefited from the following recent major funding sources: CNRS–INSU, Observatoire de Paris and Université d'Orléans, France; BMBF, MIWF–NRW, MPG, Germany; Science Foundation Ireland (SFI), Department of Business, Enterprise and Innovation (DBEI), Ireland; NWO, The Netherlands; the Science and Technology Facilities Council, UK; Ministry of Science and Higher Education, Poland; the Istituto Nazionale di Astrofisica (INAF), Italy.
This research made use of the Dutch national e-infrastructure with support of the SURF Cooperative (e-infra 180169) and the LOFAR e-infra group. The Jülich LOFAR Long Term Archive and the German LOFAR network are both coordinated and operated by the Jülich Supercomputing Centre (JSC), and computing resources on the supercomputer JUWELS at JSC were provided by the Gauss Centre for Supercomputing e.V. (grant CHTB00) through the John von Neumann Institute for Computing (NIC).
This research made use of the University of Hertfordshire high-performance computing facility and the LOFAR-UK computing facility located at the University of Hertfordshire and supported by STFC [ST/P000096/1], and of the Italian LOFAR IT computing infrastructure supported and operated by INAF, and by the Physics Department of the University of Turin (under an agreement with Consorzio Interuniversitario per la Fisica Spaziale) at the C3S Supercomputing Centre, Italy.
Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is http://www.sdss3.org/. SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University,

the Instituto de Astrofísica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.
The Pan-STARRS1 Surveys (PS1) and the PS1 public science archive have been made possible through contributions by the Institute for Astronomy, the University of Hawaii, the Pan-STARRS Project Office, the Max-Planck Society and its participating institutes, the Max Planck Institute for Astronomy, Heidelberg and the Max Planck Institute for Extraterrestrial Physics, Garching, The Johns Hopkins University, Durham University, the University of Edinburgh, the Queen's University Belfast, the Harvard–Smithsonian Center for Astrophysics, the Las Cumbres Observatory Global Telescope Network Incorporated, the National Central University of Taiwan, the Space Telescope Science Institute, the National Aeronautics and Space Administration under Grant No. NNX08AR22G issued through the Planetary Science Division of the NASA Science Mission Directorate, the National Science Foundation Grant No. AST-1238877, the University of Maryland, Eötvös Loránd University (ELTE), the Los Alamos National Laboratory, and the Gordon and Betty Moore Foundation.
This work has made use of data from the European Space Agency (ESA) mission *Gaia* (https://www.cosmos.esa.int/gaia), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, https://www.cosmos.esa.int/web/gaia/dpac/consortium). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement.
The Legacy Surveys consist of three individual and complementary projects: the Dark Energy Camera Legacy Survey (DECaLS; Proposal ID #2014B-0404; PIs: David Schlegel and Arjun Dey), the Beijing–Arizona Sky Survey (BASS; NOAO Prop. ID #2015A-0801; PIs: Zhou Xu and Xiaohui Fan), and the Mayall z-band Legacy Survey (MzLS; Prop. ID #2016A-0453; PI: Arjun Dey). DECaLS, BASS and MzLS together include data obtained, respectively, at the Blanco telescope, Cerro Tololo Inter-American Observatory, NSF's NOIRLab; the Bok telescope, Steward Observatory, University of Arizona; and the Mayall telescope, Kitt Peak National Observatory, NOIRLab. The Legacy Surveys project is honored to be permitted to conduct astronomical research on Iolkam Du'ag (Kitt Peak), a mountain with particular significance to the Tohono O'odham Nation. NOIRLab is operated by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation. This project used data obtained with the Dark Energy Camera (DECam), which was constructed by the Dark Energy Survey (DES) collaboration. Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundacao Carlos Chagas Filho de Amparo, Financiadora de Estudos e Projetos, Fundacao Carlos Chagas Filho de Amparo a Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Cientifico e Tecnologico and the Ministerio da Ciencia, Tecnologia e Inovacao, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey. The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energeticas, Medioambientales y Tecnologicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenössische Technische Hochschule (ETH) Zürich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciencies de l'Espai (IEEC/CSIC), the Institut de Fisica d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig Maximilians Universität München and the associated Excellence Cluster Universe, the University of Michigan, NSF's NOIRLab, the University of Nottingham, the Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, and Texas A&M University. BASS is a key project of the Telescope Access Program (TAP), which has been funded by the National Astronomical Observatories of China, the Chinese Academy of Sciences (the Strategic Priority Research Program "The Emergence of Cosmological Structures" Grant # XDB09000000), and the Special Fund for Astronomy from the Ministry of Finance. The BASS is also supported by the External Cooperation Program of Chinese Academy of Sciences (Grant # 114A11KYSB20160057), and Chinese National Natural Science Foundation (Grant # 11433005). The Legacy Survey team makes use of data products from the Near-Earth Object Wide-field Infrared Survey Explorer (NEOWISE), which is a project of the Jet Propulsion Laboratory/California Institute of Technology. NEOWISE is funded by the National Aeronautics and Space Administration. The Legacy Surveys imaging of the DESI footprint is supported by the Director, Office of Science, Office of High Energy Physics of the U.S. Department of Energy under Contract No. DE-AC02-05CH1123, by the National Energy Research Scientific Com-

puting Center, a DOE Office of Science User Facility under the same contract; and by the U.S. National Science Foundation, Division of Astronomical Sciences under Contract No. AST-0950945 to NOAO.

6.A1 Framework derivations and details

6.A1.1 Intrinsic proper length

Let $f_L : \mathbb{R} \to \mathbb{R}_{\geq 0}$ be the PDF of the distribution of intrinsic (i.e. 3D) proper (i.e. not comoving) radio galaxy lengths L. We assume that f_L follows a power law between $l_{\min}$ and $l_{\max}$:

$$f_L(l) = \begin{cases} f_L(l_{\text{ref}}) \left(\frac{l}{l_{\text{ref}}}\right)^{\xi} & l \in [l_{\min}, l_{\max}), \\ 0 & \text{otherwise.} \end{cases} \tag{6.43}$$

Then, when $\xi \neq -1$,

$$\begin{aligned} 1 &= \int_{\mathbb{R}} f_L(l)\, \mathrm{d}l = \int_{l_{\min}}^{l_{\max}} f_L(l_{\text{ref}}) \left(\frac{l}{l_{\text{ref}}}\right)^{\xi} \mathrm{d}l \\ &= \frac{f_L(l_{\text{ref}})\, l_{\text{ref}}}{\xi + 1} \left(\left(\frac{l_{\max}}{l_{\text{ref}}}\right)^{\xi+1} - \left(\frac{l_{\min}}{l_{\text{ref}}}\right)^{\xi+1} \right). \end{aligned} \tag{6.44}$$

This equation provides the normalisation factor $f_L(l_{\text{ref}})$ given $l_{\min}$, $l_{\max}$, ξ and an arbitrary choice for $l_{\text{ref}} \neq 0$.

When $l_{\max} = \infty$, L has a Pareto Type I distribution and we must have $\xi < -1$ for the integrals of Eq. 6.44 to converge. Throughout the remaining analysis, for the sake of simplicity, we assume $l_{\max} = \infty$ and $\xi < -1$ and choose $l_{\text{ref}} = l_{\min}$. Then $f_L(l_{\text{ref}}) = f_L(l_{\min}) = -\frac{\xi+1}{l_{\min}}$. The corresponding CDF is

$$F_L(l) = \begin{cases} 0 & \text{if } l \leq l_{\min} \\ 1 - \left(\frac{l}{l_{\min}}\right)^{\xi+1} & \text{if } l > l_{\min}. \end{cases} \tag{6.45}$$

In reality, radio galaxies cannot become arbitrarily long (Hardcastle, 2018): at some distance from the host, all energy carried by the jets will have been radiated away, used to perform work on the IGM, transferred to CMB photons through inverse Compton scattering, or converted into heat. However, at the moment of writing, the implied maximum length $l_{\max}$ remains ill-constrained. As we see throughout Appendix 6.A1, a major advantage of simple model assumptions is that explicit and thus insightful analytic expressions can be derived. Such easy-to-evaluate expressions com-

plement the results of expensive numerical simulations, which aim to maximise realism; our aim to maximise insight is best served by setting $l_{\max} = \infty$. From Eq. 6.45, we see that for $l_{\min} = 0.7$ Mpc and $\xi = -3.5$, the tail of the intrinsic proper length distribution ($l > 5$ Mpc) contains less than 1% of probability. Thus, for most framework applications, our choice $l_{\max} = \infty$ is unproblematic — except when higher powers of L are involved. For example, the volume-filling fraction calculations of Appendix 6.A1.10 necessitate considering L^3; for realistic ξ, results exist only for finite $l_{\max}$.

Upon relabelling $\xi \to -\alpha - 1$, $l_{\min} \to k$, $L \to X$, and $l \to x$, one obtains the literature's most common form of the Pareto Type I PDF:

$$f_X(x) = \begin{cases} \frac{\alpha k^\alpha}{x^{\alpha+1}} & x \geq k, \\ 0 & x < k. \end{cases} \tag{6.46}$$

6.A1.2 Projected proper length

Distribution for RGs

Let $f_{L_\mathrm{p}} : \mathbb{R} \to \mathbb{R}_{\geq 0}$ be the PDF of the distribution of projected proper radio galaxy lengths. The PDF f_{L_p} follows from the associated CDF $F_{L_\mathrm{p}} : \mathbb{R} \to [0,1]$ through differentiation; F_{L_p} and f_L relate through

$$\begin{aligned} F_{L_\mathrm{p}}\left(l_\mathrm{p}\right) &:= \mathbb{P}\left(L_\mathrm{p} \leq l_\mathrm{p}\right) = \int_{\mathbb{R}} \mathbb{P}\left(L \sin\Theta \leq l_\mathrm{p} \mid L = l\right) f_L(l)\ \mathrm{d}l \\ &= \int_{\mathbb{R}} \mathbb{P}\left(\sin\Theta \leq \frac{l_\mathrm{p}}{l}\right) f_L(l)\ \mathrm{d}l. \end{aligned} \tag{6.47}$$

We note that $F_{L_\mathrm{p}}\left(l_\mathrm{p}\right)$ vanishes for $l_\mathrm{p} \leq 0$: $f_L(l)$ vanishes when l is negative, whilst $\mathbb{P}\left(\sin\Theta \leq \frac{l_\mathrm{p}}{l}\right)$ vanishes when l is positive. Clearly, the interesting case is $l_\mathrm{p} > 0$. We can differentiate between two cases: $l_\mathrm{p} < l_{\min}$, and $l_\mathrm{p} \geq l_{\min}$. In the first case, considering that f_L has support from $l_{\min}$ onwards only, we have $\frac{l_\mathrm{p}}{l} < 1$, and

$$
\begin{aligned}
F_{L_\mathrm{p}}\left(l_\mathrm{p}\right) &= \int_{l_\mathrm{min}}^{\infty} \left(\mathbb{P}\left(\Theta \leq \arcsin \frac{l_\mathrm{p}}{l}\right) \right. \\
&\quad \left. + \mathbb{P}\left(\Theta \geq \pi - \arcsin \frac{l_\mathrm{p}}{l}\right)\right) \frac{-(\xi+1)}{l_\mathrm{min}} \left(\frac{l}{l_\mathrm{min}}\right)^{\xi} \mathrm{d}l \\
&= -\frac{\xi+1}{l_\mathrm{min}} \int_{l_\mathrm{min}}^{\infty} 2\mathbb{P}\left(\Theta \leq \arcsin \frac{l_\mathrm{p}}{l}\right) \left(\frac{l}{l_\mathrm{min}}\right)^{\xi} \mathrm{d}l \\
&= -\frac{\xi+1}{l_\mathrm{min}} \int_{l_\mathrm{min}}^{\infty} \left(1 - \sqrt{1 - \left(\frac{l_\mathrm{p}}{l}\right)^2}\right) \left(\frac{l}{l_\mathrm{min}}\right)^{\xi} \mathrm{d}l \\
&= 1 + (\xi+1) \int_{l_\mathrm{min}}^{\infty} \sqrt{\left(\frac{l}{l_\mathrm{min}}\right)^2 - \left(\frac{l_\mathrm{p}}{l_\mathrm{min}}\right)^2} \left(\frac{l}{l_\mathrm{min}}\right)^{\xi-1} \frac{\mathrm{d}l}{l_\mathrm{min}} \\
&= 1 + (\xi+1) \int_{1}^{\infty} \sqrt{\eta^2 - \left(\frac{l_\mathrm{p}}{l_\mathrm{min}}\right)^2} \eta^{\xi-1} \mathrm{d}\eta. \qquad (6.48)
\end{aligned}
$$

In the second case, we split up the integral in two:

$$
\begin{aligned}
F_{L_\mathrm{p}}\left(l_\mathrm{p}\right) &= \int_{l_\mathrm{min}}^{l_\mathrm{p}} \mathbb{P}\left(\sin\Theta \leq \frac{l_\mathrm{p}}{l}\right) f_L\left(l\right)\ \mathrm{d}l \\
&\quad + \int_{l_\mathrm{p}}^{\infty} \mathbb{P}\left(\sin\Theta \leq \frac{l_\mathrm{p}}{l}\right) f_L\left(l\right)\ \mathrm{d}l \\
&= \int_{l_\mathrm{min}}^{l_\mathrm{p}} f_L\left(l\right)\ \mathrm{d}l + \int_{l_\mathrm{p}}^{\infty} \left(1 - \sqrt{1 - \left(\frac{l_\mathrm{p}}{l}\right)^2}\right) f_L\left(l\right)\ \mathrm{d}l \\
&= 1 - \left(\frac{l_\mathrm{p}}{l_\mathrm{min}}\right)^{\xi+1} + \left(\frac{l_\mathrm{p}}{l_\mathrm{min}}\right)^{\xi+1} \\
&\quad + (\xi+1) \int_{l_\mathrm{p}}^{\infty} \sqrt{\left(\frac{l}{l_\mathrm{min}}\right)^2 - \left(\frac{l_\mathrm{p}}{l_\mathrm{min}}\right)^2} \left(\frac{l}{l_\mathrm{min}}\right)^{\xi-1} \frac{\mathrm{d}l}{l_\mathrm{min}} \\
&= 1 + (\xi+1) \int_{\frac{l_\mathrm{p}}{l_\mathrm{min}}}^{\infty} \sqrt{\eta^2 - \left(\frac{l_\mathrm{p}}{l_\mathrm{min}}\right)^2} \eta^{\xi-1} \mathrm{d}\eta \\
&= 1 + (\xi+1) \frac{\sqrt{\pi}}{4} \left(\frac{l_\mathrm{p}}{l_\mathrm{min}}\right)^{\xi+1} \frac{\Gamma\left(-\frac{\xi}{2} - \frac{1}{2}\right)}{\Gamma\left(-\frac{\xi}{2} + 1\right)}. \qquad (6.49)
\end{aligned}
$$

In summary,

$$F_{L_\mathrm{p}}\left(l_\mathrm{p}\right) = \begin{cases} 0 & \text{if } l_\mathrm{p} \leq 0 \\ 1 + (\xi+1)\int_1^\infty \sqrt{\eta^2 - \left(\frac{l_\mathrm{p}}{l_\mathrm{min}}\right)^2}\,\eta^{\xi-1}\,\mathrm{d}\eta & \text{if } 0 < l_\mathrm{p} < l_\mathrm{min} \\ 1 + (\xi+1)\,\frac{\sqrt{\pi}}{4}\left(\frac{l_\mathrm{p}}{l_\mathrm{min}}\right)^{\xi+1}\frac{\Gamma\left(-\frac{\xi}{2}-\frac{1}{2}\right)}{\Gamma\left(-\frac{\xi}{2}+1\right)} & \text{if } l_\mathrm{p} \geq l_\mathrm{min}. \end{cases} \tag{6.50}$$

Through differentiation,

$$f_{L_\mathrm{p}}\left(l_\mathrm{p}\right) = \begin{cases} 0 & \text{if } l_\mathrm{p} \leq 0 \\ -\frac{\xi+1}{l_\mathrm{min}}\frac{l_\mathrm{p}}{l_\mathrm{min}} I\left(\xi-1, \frac{l_\mathrm{p}}{l_\mathrm{min}}\right) & \text{if } 0 < l_\mathrm{p} < l_\mathrm{min} \\ \frac{(\xi+1)^2}{l_\mathrm{min}}\frac{\sqrt{\pi}}{4}\left(\frac{l_\mathrm{p}}{l_\mathrm{min}}\right)^{\xi}\frac{\Gamma\left(-\frac{\xi}{2}-\frac{1}{2}\right)}{\Gamma\left(-\frac{\xi}{2}+1\right)} & \text{if } l_\mathrm{p} \geq l_\mathrm{min}. \end{cases} \tag{6.51}$$

where

$$I\left(a, b\right) := \int_1^\infty \frac{\eta^a\,\mathrm{d}\eta}{\sqrt{\eta^2 - b^2}} \text{ for } a < 0, |b| < 1. \tag{6.52}$$

Significantly, for $l_\mathrm{p} \geq l_\mathrm{min}$, the projected length distribution follows a power law in l_p with the same exponent as the power law for the intrinsic length distribution: $f_{L_\mathrm{p}} \propto l_\mathrm{p}^{\xi}$ (just as $f_L \propto l^{\xi}$).[24] We compare f_L and f_{L_p} in Fig. 6.1.

Distribution for giants

To derive the projected length distribution for giants, we consider the distribution of the conditioned RV $L_\mathrm{p} \mid L_\mathrm{p} > l_\mathrm{p,GRG}$:

$$\begin{aligned} F_{L_\mathrm{p} \mid L_\mathrm{p} > l_\mathrm{p,GRG}}\left(l_\mathrm{p}\right) :=& \mathbb{P}\left(L_\mathrm{p} \leq l_\mathrm{p} \mid L_\mathrm{p} > l_\mathrm{p,GRG}\right) \\ =& 1 - \mathbb{P}\left(L_\mathrm{p} > l_\mathrm{p} \mid L_\mathrm{p} > l_\mathrm{p,GRG}\right) \\ =& 1 - \frac{\mathbb{P}\left(L_\mathrm{p} > l_\mathrm{p}, L_\mathrm{p} > l_\mathrm{p,GRG}\right)}{\mathbb{P}\left(L_\mathrm{p} > l_\mathrm{p,GRG}\right)}. \end{aligned} \tag{6.53}$$

For $l_\mathrm{p} > l_\mathrm{p,GRG}$, this reduces to

$$F_{L_\mathrm{p} \mid L_\mathrm{p} > l_\mathrm{p,GRG}}\left(l_\mathrm{p}\right) = 1 - \frac{1 - F_{L_\mathrm{p}}\left(l_\mathrm{p}\right)}{1 - F_{L_\mathrm{p}}\left(l_\mathrm{p,GRG}\right)}. \tag{6.54}$$

[24] For $l_\mathrm{max} < \infty$, this statement does not hold exactly.

Furthermore assuming $l_{\mathrm{p,GRG}} > l_{\mathrm{min}}$, we twice use the bottom expression of Eq. 6.50 to obtain the final CDF expression. As before, the corresponding PDF follows through differentation. We find

$$F_{L_{\mathrm{p}} \mid L_{\mathrm{p}} > l_{\mathrm{p,GRG}}}\left(l_{\mathrm{p}}\right) = \begin{cases} 0 & \text{if } l_{\mathrm{p}} \leq l_{\mathrm{p,GRG}} \\ 1 - \left(\frac{l_{\mathrm{p}}}{l_{\mathrm{p,GRG}}}\right)^{\xi+1} & \text{if } l_{\mathrm{p}} > l_{\mathrm{p,GRG}}. \end{cases} \tag{6.55}$$

$$f_{L_{\mathrm{p}} \mid L_{\mathrm{p}} > l_{\mathrm{p,GRG}}}\left(l_{\mathrm{p}}\right) = \begin{cases} 0 & \text{if } l_{\mathrm{p}} \leq l_{\mathrm{p,GRG}} \\ -\frac{\xi+1}{l_{\mathrm{p,GRG}}}\left(\frac{l_{\mathrm{p}}}{l_{\mathrm{p,GRG}}}\right)^{\xi} & \text{if } l_{\mathrm{p}} > l_{\mathrm{p,GRG}}. \end{cases} \tag{6.56}$$

Thus, the associated survival function is

$$\mathbb{P}\left(L_{\mathrm{p}} > l_{\mathrm{p}} \mid L_{\mathrm{p}} > l_{\mathrm{p,GRG}}\right) = \left(\frac{l_{\mathrm{p}}}{l_{\mathrm{p,GRG}}}\right)^{\xi+1}. \tag{6.57}$$

The mean projected proper length of giants follows from the PDF by direct computation:

$$\begin{aligned} \mathbb{E}[L_{\mathrm{p}} \mid L_{\mathrm{p}} > l_{\mathrm{p,GRG}}] &:= \int_{-\infty}^{\infty} f_{L_{\mathrm{p}} \mid L_{\mathrm{p}} > l_{\mathrm{p,GRG}}}\left(l_{\mathrm{p}}\right) \cdot l_{\mathrm{p}} \, \mathrm{d}l_{\mathrm{p}} \\ &= l_{\mathrm{p,GRG}} \frac{\xi+1}{\xi+2}. \end{aligned} \tag{6.58}$$

6.A1.3 Deprojection factor

Consider a radio galaxy (RG) with a projected proper length l_{p}. Let the inclination angle θ denote the angle between the RG's central axis and the line of sight. The RG's inclination angle, projected proper length, and intrinsic proper length l relate via $l_{\mathrm{p}} = l \sin\theta$. Switching to random variable (RV) notation by using capital letters, the intrinsic proper length (which is the most physically relevant quantity) follows from the projected proper length (which can be measured) and the inclination angle (which is typically unknown), according to

$$L = \frac{1}{\sin\Theta} L_{\mathrm{p}}. \tag{6.59}$$

Without lobes

Calling the deprojection factor $D := (\sin\Theta)^{-1}$, we now calculate the distribution of D for $f_{\Theta}(\theta) = \frac{1}{2}\sin\theta; \theta \in [0, \pi]$. The result is a continuous univariate distribution

without parameters supported on the semi-infinite interval $(1, \infty)$. Let $F_D : \mathbb{R} \to [0, 1]$ be the cumulative density function (CDF) of D. Then, for $d > 1$,

$$\begin{aligned} F_D(d) &:= \mathbb{P}(D \leq d) = \mathbb{P}\left(\sin\Theta \geq \frac{1}{d}\right) \\ &= \mathbb{P}\left(\arcsin\frac{1}{d} \leq \Theta \leq \pi - \arcsin\frac{1}{d}\right) \\ &= F_\Theta\left(\pi - \arcsin\frac{1}{d}\right) - F_\Theta\left(\arcsin\frac{1}{d}\right) \\ &= \cos\arcsin\frac{1}{d} = \sqrt{1 - \frac{1}{d^2}}. \end{aligned} \tag{6.60}$$

Meanwhile, $F_D(d) = 0$ for $d \leq 1$. The quantile function $F_D^{-1} : [0, 1] \to [1, \infty)$ follows from solving $F_D(d) = p$ for d:

$$F_D^{-1}(p) = \frac{1}{\sqrt{1 - p^2}}. \tag{6.61}$$

Thus, the minimum factor is $F_D^{-1}(p = 0) = 1$, the median $F_D^{-1}\left(p = \frac{1}{2}\right) = \frac{2}{\sqrt{3}}$, $F_D^{-1}\left(p = \frac{1}{2}\sqrt{2}\right) = \sqrt{2}$, $F_D^{-1}\left(p = \frac{1}{2}\sqrt{3}\right) = 2$, and factors can grow arbitrarily large: $F_D^{-1}(p) \to \infty$ as $p \to 1$. We conclude that half of all RGs have an intrinsic proper length more than $\frac{2}{\sqrt{3}}$ their projected proper length.

By differentiating F_D to d we obtain $f_D : \mathbb{R} \to \mathbb{R}_{\geq 0}$, the probability density function (PDF) of D:

$$f_D(d) = \begin{cases} \frac{1}{d^2\sqrt{d^2-1}} & \text{if } d > 1; \\ 0 & \text{if } d \leq 1. \end{cases} \tag{6.62}$$

The mean of D is $\mathbb{E}(D) = \frac{\pi}{2}$; the variance of D is undefined, as the corresponding integral diverges. Because $f_D(d)$ has no maximum, the mode is undefined too. The PDF and CDF of D are shown in the upper left and right panels of Fig. 6.22, respectively.

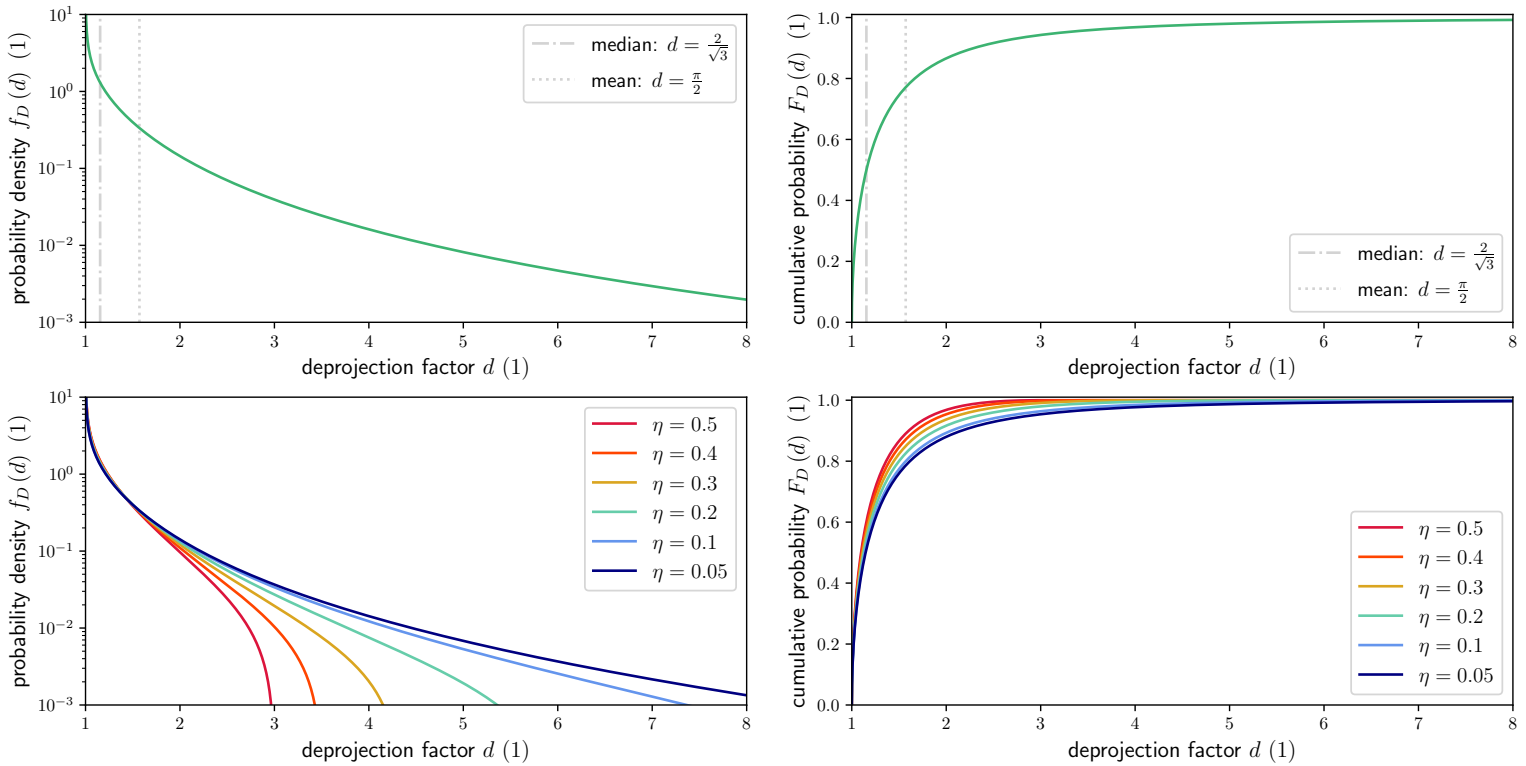


Figure 6.22: PDFs *(left column)* and CDFs *(right column)* of the deprojection factor RV D. These functions quantify how much longer RGs are than their projected lengths suggest. *Top row:* model without lobes. Of the three canonical measures of central tendency, only the median and mean exist (and equal $\frac{2}{\sqrt{3}}$ and $\frac{\pi}{2}$, respectively). *Bottom row:* model with spherical lobes. The distribution of D now depends on a parameter: the ratio η between the lobe diameter $2R$ and the distance between the lobe centres L. The model without lobes is recovered in the limit $\eta \to 0$.

With lobes

The simplest model that includes a pair of lobes approximates them as spheres of radius R, whose centres are connected by a line segment of length L. Regardless of the viewing angle, the spheres retain their size, in contrast to the line segment connecting them. Calling $\eta := \frac{2R}{L} \in \mathbb{R}_{\geq 0}$, we have

$$D := \frac{L + 2R}{L \sin \Theta + 2R} = \frac{1 + \eta}{\sin \Theta + \eta}. \tag{6.63}$$

Again assuming $f_\Theta(\theta) = \frac{1}{2} \sin \theta; \theta \in [0, \pi]$, we repeat the derivation and find

$$F_D(d) = \begin{cases} 0 & \text{if } d \leq 1; \\ \sqrt{1 - \left(\frac{1+\eta}{d} - \eta\right)^2} & \text{if } 1 < d < \frac{1}{\eta} + 1; \\ 1 & \text{if } d \geq \frac{1}{\eta} + 1. \end{cases} \tag{6.64}$$

$$f_D(d) = \begin{cases} 0 & \text{if } d \leq 1; \\ \frac{(1+\eta)(1+\eta-\eta d)}{d^2 \sqrt{d^2 - (1+\eta-\eta d)^2}} & \text{if } 1 < d < \frac{1}{\eta} + 1; \\ 0 & \text{if } d \geq \frac{1}{\eta} + 1. \end{cases} \tag{6.65}$$

The quantile function becomes

$$F_D^{-1}(p) = \frac{1 + \eta}{\sqrt{1 - p^2} + \eta}. \tag{6.66}$$

The minimum factor remains $F_D^{-1}(p = 0) = 1$, but a maximum factor now exists: $F_D^{-1}(p = 1) = \frac{1}{\eta} + 1$. The median of D is

$$F_D^{-1}\left(p = \frac{1}{2}\right) = \frac{1 + \eta}{\frac{1}{2}\sqrt{3} + \eta}, \tag{6.67}$$

which tends to $\frac{2}{\sqrt{3}}$ for $\eta \to 0$ (as before), and to 1 for $\eta \to \infty$: projection ceases to be an appreciable effect when the lobes are much larger than the line segment connecting their centres. If $\eta \neq 0$, D has finite support, and thus the mean, variance, and higher moments exist. The n-th non-central moment is

$$\mathbb{E}\left[D^n\right] = (1 + \eta)^2 I_{n-2}(\eta) - \eta (1 + \eta) I_{n-1}(\eta), \tag{6.68}$$

where

$$I_n(\eta) := \int_1^{\frac{1}{\eta}+1} \frac{d^n \, \mathrm{d}d}{\sqrt{d^2 - (1+\eta-\eta d)^2}}. \tag{6.69}$$

In particular,

$$I_{-1}(\eta) = \frac{\pi}{2(1+\eta)}, \quad I_0(\eta) = \frac{\ln\left(\sqrt{1-\eta^2}+1\right) - \ln\eta}{\sqrt{1-\eta^2}},$$
$$I_1(\eta) = \frac{1}{(1-\eta)\eta}\left(1 - \frac{2\eta^2 \operatorname{arcsinh}\sqrt{\frac{1-\eta}{2\eta}}}{\sqrt{1-\eta^2}}\right). \tag{6.70}$$

The expectation value of D is

$$\begin{aligned}\mathbb{E}[D] &= (1+\eta)^2 I_{-1}(\eta) - \eta(1+\eta) I_0(\eta) \\ &= (1+\eta)\left(\frac{\pi}{2} - \frac{\eta}{\sqrt{1-\eta^2}} \ln\left(\frac{\sqrt{1-\eta^2}}{\eta} + \frac{1}{\eta}\right)\right).\end{aligned} \tag{6.71}$$

The variance of D follows from combining

$$\mathbb{E}\left[D^2\right] = (1+\eta)^2 I_0(\eta) - \eta(1+\eta) I_1(\eta) \tag{6.72}$$

and the identity $\mathbb{V}[D] = \mathbb{E}[D^2] - \mathbb{E}^2[D]$. The mode remains undefined. For typical values of η, we show the PDF and CDF of D in the bottom left and right panels of Fig. 6.22, respectively. Figure 6.23 shows the median, mean, and standard deviation of D as a function of η.

6.A1.4 Intrinsic proper length posterior and its moments

Our next objective is to find the posterior PDF $f_{L|L_\mathrm{p}=l_\mathrm{p}}(l)$ through Bayes' theorem:

$$f_{L|L_\mathrm{p}=l_\mathrm{p}}(l) = \frac{f_{L_\mathrm{p}|L=l}(l_\mathrm{p}) f_L(l)}{f_{L_\mathrm{p}}(l_\mathrm{p})}. \tag{6.73}$$

Our line of attack will be to first calculate the likelihood CDF $F_{L_\mathrm{p}|L=l}(l_\mathrm{p})$, and then through differentiation the likelihood PDF $f_{L_\mathrm{p}|L=l}(l_\mathrm{p})$. Of course, we always consider

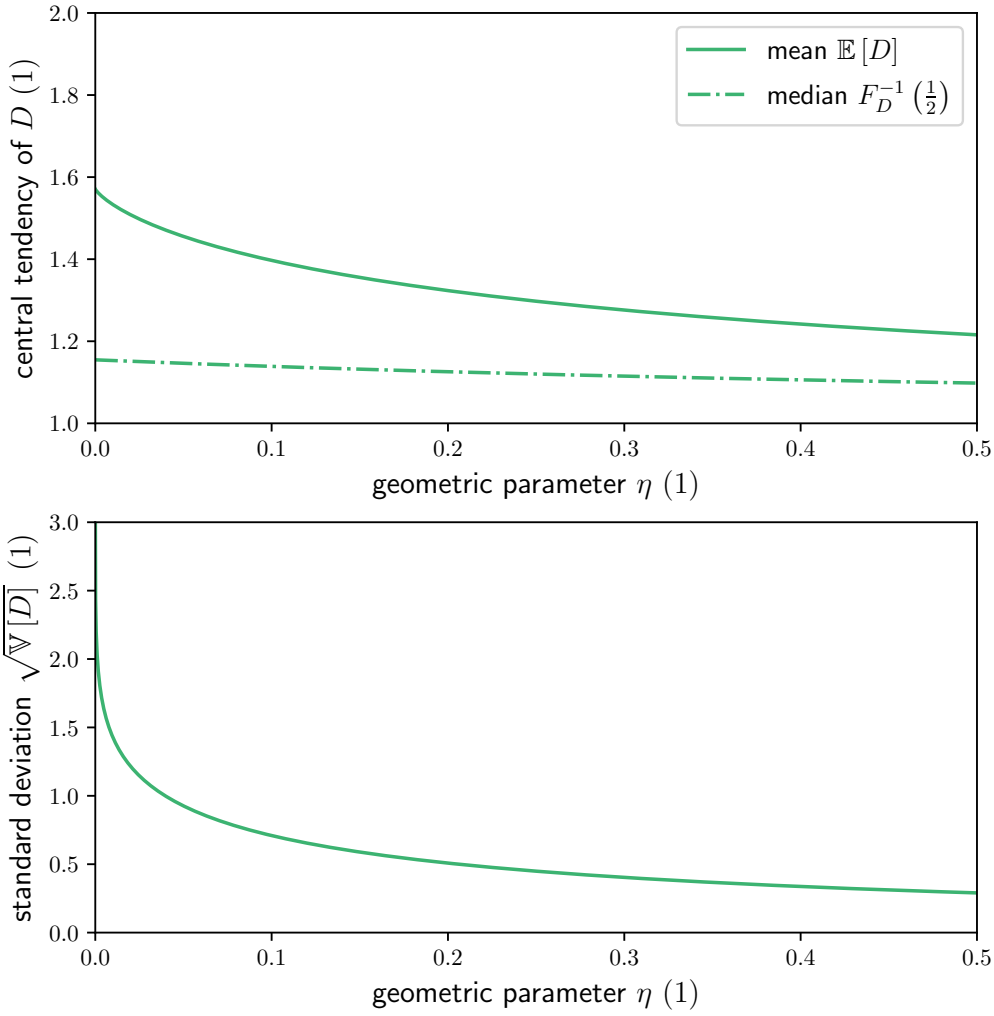


Figure 6.23: Summary statistics for the deprojection factor RV D under the spherical lobe model. *Top:* the η-dependency of two measures of central tendency of D. For $\eta = 0$, the mean $\mathbb{E}[D] = \frac{\pi}{2}$ and the median $F_D^{-1}\left(\frac{1}{2}\right) = \frac{2}{\sqrt{3}}$. *Bottom:* the η-dependency of the standard deviation of D. As $\eta \to 0+$, $\sqrt{\mathbb{V}[D]} \to \infty$.

$l > 0$, and

$$
\begin{aligned}
F_{L_\mathrm{p}|L=l}\left(l_\mathrm{p}\right) &:= \mathbb{P}\left(L_\mathrm{p} \leq l_\mathrm{p} \mid L = l\right) = \mathbb{P}\left(l \sin\Theta \leq l_\mathrm{p}\right) \\
&= \begin{cases} 0 & \text{if } l_\mathrm{p} \leq 0 \\ 2\,\mathbb{P}\left(\Theta \leq \arcsin\frac{l_\mathrm{p}}{l}\right) & \text{if } 0 < l_\mathrm{p} \leq l \\ 1 & \text{if } l_\mathrm{p} > l, \end{cases}
\end{aligned}
\tag{6.74}
$$

where we have invoked Eq. 6.59 and the fact that $f_\Theta(\theta)$ is symmetric in $\theta = \frac{\pi}{2}$ for the case $0 < l_\mathrm{p} < l$. Concretising this case further yields

$$\begin{aligned} F_{L_\mathrm{p}|L=l}\left(l_\mathrm{p}\right) &= 2\,F_\Theta\left(\arcsin\frac{l_\mathrm{p}}{l}\right) = 1 - \cos\arcsin\frac{l_\mathrm{p}}{l} \\ &= 1 - \sqrt{1 - \left(\frac{l_\mathrm{p}}{l}\right)^2}. \end{aligned} \tag{6.75}$$

Through differentiation,

$$f_{L_\mathrm{p}|L=l}\left(l_\mathrm{p}\right) = \begin{cases} 0 & \text{if } l_\mathrm{p} \leq 0 \\ \frac{l_\mathrm{p}}{l^2}\frac{1}{\sqrt{1-\left(\frac{l_\mathrm{p}}{l}\right)^2}} & \text{if } 0 < l_\mathrm{p} < l \\ 0 & \text{if } l_\mathrm{p} \geq l. \end{cases} \tag{6.76}$$

Having found the likelihood, the posterior PDF follows directly through Bayes' theorem and the PDFs computed hitherto. In concreto, when $0 < l_\mathrm{p} \leq l_\mathrm{min}$,

$$f_{L|L_\mathrm{p}=l_\mathrm{p}}(l) = \begin{cases} 0 & \text{if } l \leq l_\mathrm{min} \\ \frac{1}{I\left(\xi-1,\frac{l_\mathrm{p}}{l_\mathrm{min}}\right)}\frac{1}{l_\mathrm{p}}\frac{1}{\sqrt{\left(\frac{l}{l_\mathrm{p}}\right)^2-1}}\left(\frac{l}{l_\mathrm{min}}\right)^{\xi-1} & \text{if } l > l_\mathrm{min}. \end{cases} \tag{6.77}$$

whereas for $l_\mathrm{p} > l_\mathrm{min}$,

$$f_{L|L_\mathrm{p}=l_\mathrm{p}}(l) = \begin{cases} 0 & \text{if } l \leq l_\mathrm{p} \\ -\frac{\xi}{2^{1+\xi}\pi}\frac{\Gamma^2\left(-\frac{\xi}{2}\right)}{\Gamma(-\xi)}\frac{1}{l_\mathrm{p}}\frac{1}{\sqrt{\left(\frac{l}{l_\mathrm{p}}\right)^2-1}}\left(\frac{l}{l_\mathrm{p}}\right)^{\xi-1} & \text{if } l > l_\mathrm{p}. \end{cases} \tag{6.78}$$

We note that in both cases, for $l \gg l_\mathrm{p}, f_{L|L_\mathrm{p}=l_\mathrm{p}}(l) \propto \left(\frac{l}{l_\mathrm{p}}\right)^{\xi-2}$: the posterior probability density follows a power law in l with exponent $\xi - 2$.

For $l_\mathrm{p} \leq l_\mathrm{min}$, the posterior mean is

$$\mathbb{E}\left[L \mid L_\mathrm{p} = l_\mathrm{p}\right] = l_\mathrm{min}\frac{I\left(\xi, \frac{l_\mathrm{p}}{l_\mathrm{min}}\right)}{I\left(\xi-1, \frac{l_\mathrm{p}}{l_\mathrm{min}}\right)}. \tag{6.79}$$

The posterior variance $\mathbb{V}\left[L \mid L_\mathrm{p} = l_\mathrm{p}\right]$ follows from considering the second non-central moment:

$$\mathbb{E}\left[L^2 \mid L_\mathrm{p} = l_\mathrm{p}\right] = l_\mathrm{min}^2 \frac{I\left(\xi+1, \frac{l_\mathrm{p}}{l_\mathrm{min}}\right)}{I\left(\xi-1, \frac{l_\mathrm{p}}{l_\mathrm{min}}\right)}, \tag{6.80}$$

so that

$$\begin{aligned}\mathbb{V}\left[L \mid L_\mathrm{p} = l_\mathrm{p}\right] &= \mathbb{E}\left[L^2 \mid L_\mathrm{p} = l_\mathrm{p}\right] - \mathbb{E}^2\left[L \mid L_\mathrm{p} = l_\mathrm{p}\right] \\ &= l_\mathrm{min}^2 \left(\frac{I\left(\xi+1, \frac{l_\mathrm{p}}{l_\mathrm{min}}\right)}{I\left(\xi-1, \frac{l_\mathrm{p}}{l_\mathrm{min}}\right)} - \frac{I^2\left(\xi, \frac{l_\mathrm{p}}{l_\mathrm{min}}\right)}{I^2\left(\xi-1, \frac{l_\mathrm{p}}{l_\mathrm{min}}\right)}\right).\end{aligned} \tag{6.81}$$

For $l_\mathrm{p} > l_\mathrm{min}$,

$$\begin{aligned}\mathbb{E}\left[L \mid L_\mathrm{p} = l_\mathrm{p}\right] &= l_\mathrm{p} \cdot \frac{-\xi}{2^{\xi+1}\pi} \frac{\Gamma^2\left(-\frac{\xi}{2}\right)}{\Gamma\left(-\xi\right)} \int_1^\infty \frac{\eta^\xi \, \mathrm{d}\eta}{\sqrt{\eta^2-1}} \\ &= l_\mathrm{p} \cdot \frac{-\xi}{2^{2\xi+3}\pi} \frac{\Gamma^4\left(-\frac{\xi}{2}\right)}{\Gamma^2\left(-\xi\right)}.\end{aligned} \tag{6.82}$$

Proceeding analogously,

$$\mathbb{E}\left[L^2 \mid L_\mathrm{p} = l_\mathrm{p}\right] = l_\mathrm{p}^2 \frac{\xi}{\xi+1}, \tag{6.83}$$

so that

$$\mathbb{V}\left[L \mid L_\mathrm{p} = l_\mathrm{p}\right] = l_\mathrm{p}^2 \left(\frac{\xi}{\xi+1} - \frac{\xi^2}{2^{4\xi+6}\pi^2} \frac{\Gamma^8\left(-\frac{\xi}{2}\right)}{\Gamma^4\left(-\xi\right)}\right). \tag{6.84}$$

In this case, both the mean and standard deviation of $L \mid L_\mathrm{p} = l_\mathrm{p}$ are proportional to l_p. In the table below, we list the mean and standard deviation in multiples of l_p for several values of ξ. Since we assume $\xi < -1$, the mean and variance are guaranteed to exist. The existence of higher-order moments is ξ-dependent; in concreto, the highest defined order is $\lceil -\xi \rceil$.

We prove that L_p and D are not independent by contradiction. If we assume that L_p and D are independent, then $\mathbb{E}[L_\mathrm{p} D] = \mathbb{E}[L_\mathrm{p}]\mathbb{E}[D]$, or $\mathbb{E}[L] = \mathbb{E}[L]\mathbb{E}[\sin\Theta]\mathbb{E}[D]$

by the independence of L and $\sin\Theta$. In other words, if L_p and D are independent, then $1 = \mathbb{E}[\sin\Theta]\mathbb{E}[D]$. However, $\mathbb{E}[\sin\Theta] = \frac{\pi}{4}$ and $\mathbb{E}[D] = \frac{\pi}{2}$; because $1 \neq \frac{\pi^2}{8}$, the initial assertion must be wrong.

6.A1.5 GRG inclination angle

We derive the inclination angle distribution for giants. The probability that a GRG has inclination angle θ is

$$\begin{aligned} f_{\Theta \mid L_\mathrm{p} > l_\mathrm{p,GRG}}(\theta)\,\mathrm{d}\theta &= \frac{\mathbb{P}\left(\Theta = \theta, L\sin\Theta > l_\mathrm{p,GRG}\right)}{\mathbb{P}\left(L_\mathrm{p} > l_\mathrm{p,GRG}\right)} \\ &= \frac{f_\Theta(\theta)\,\mathrm{d}\theta \cdot \mathbb{P}\left(L > \frac{l_\mathrm{p,GRG}}{\sin\theta}\right)}{\mathbb{P}\left(L_\mathrm{p} > l_\mathrm{p,GRG}\right)}, \end{aligned} \tag{6.85}$$

where we make use of the fact that the numerator's joint probability factorises because L and Θ are independent RVs. We thus find

$$f_{\Theta \mid L_\mathrm{p} > l_\mathrm{p,GRG}}(\theta) = \frac{\left(1 - F_L\left(\frac{l_\mathrm{p,GRG}}{\sin\theta}\right)\right) f_\Theta(\theta)}{1 - F_{L_\mathrm{p}}(l_\mathrm{p,GRG})}. \tag{6.86}$$

Under the Paretian assumption for L, we have

$$f_{\Theta \mid L_\mathrm{p} > l_\mathrm{p,GRG}}(\theta) = \frac{2}{-(\xi+1)\sqrt{\pi}} \frac{\Gamma\left(-\frac{\xi}{2}+1\right)}{\Gamma\left(-\frac{\xi}{2}-\frac{1}{2}\right)} \sin^{-\xi}\theta. \tag{6.87}$$

6.A1.6 GRG angular length

The GRG angular length RV $\Phi \mid L_\mathrm{p} > l_\mathrm{p,GRG}$ relates to the GRG projected proper length RV $L_\mathrm{p} \mid L_\mathrm{p} > l_\mathrm{p,GRG}$ and the comoving distance RV R as

$$\Phi \mid L_\mathrm{p} > l_\mathrm{p,GRG} = L_\mathrm{p} \mid L_\mathrm{p} > l_\mathrm{p,GRG} \cdot \frac{1 + z(R)}{R}. \tag{6.88}$$

The model predicts the distribution of GRG angular lengths in the Local Universe up to comoving distance r_max. If the GRG number density is constant in the Local

Universe,

$$F_R(r) = \begin{cases} 0 & \text{if } r \leq 0 \\ \left(\frac{r}{r_{\max}}\right)^3 & \text{if } 0 < r < r_{\max} \\ 1 & \text{if } r \geq r_{\max} \end{cases} \tag{6.89}$$

$$f_R(r) = \begin{cases} 0 & \text{if } r \leq 0 \\ \frac{3}{r_{\max}}\left(\frac{r}{r_{\max}}\right)^2 & \text{if } 0 < r < r_{\max} \\ 0 & \text{if } r \geq r_{\max}. \end{cases} \tag{6.90}$$

Because GRG life cycles are shorter than the age of the Universe, $L_\mathrm{p} \mid L_\mathrm{p} > l_{\mathrm{p,GRG}}$ and R are independent RVs. In a non-expanding universe, $z(R) = 0$, and the distribution of $\Phi \mid L_\mathrm{p} > l_{\mathrm{p,GRG}}$ can be calculated analytically. From a well-known ratio distribution identity,

$$\begin{aligned} f_{\Phi \mid L_\mathrm{p} > l_{\mathrm{p,GRG}}}(\varphi) &= \int_{-\infty}^{\infty} |r| \cdot f_{L_\mathrm{p} \mid L_\mathrm{p} > l_{\mathrm{p,GRG}}}(\varphi r) \cdot f_R(r)\, \mathrm{d}r \\ &= \int_0^{r_{\max}} f_{L_\mathrm{p} \mid L_\mathrm{p} > l_{\mathrm{p,GRG}}}(\varphi r) \cdot 3\left(\frac{r}{r_{\max}}\right)^3 \mathrm{d}r. \end{aligned} \tag{6.91}$$

The integrand is non-zero only when $\varphi r \geq l_{\mathrm{p,GRG}}$, suggesting a lower integration limit of $\frac{l_{\mathrm{p,GRG}}}{\varphi}$. The integral vanishes altogether when $\frac{l_{\mathrm{p,GRG}}}{\varphi} \geq r_{\max}$. Calling $\varphi_{\mathrm{GRG}} := \frac{l_{\mathrm{p,GRG}}}{r_{\max}}$, we have $f_{\Phi \mid L_\mathrm{p} > l_{\mathrm{p,GRG}}}(\varphi) = 0$ for $\varphi \leq \varphi_{\mathrm{GRG}}$. For $\varphi > \varphi_{\mathrm{GRG}}$,

$$\begin{aligned} f_{\Phi \mid L_\mathrm{p} > l_{\mathrm{p,GRG}}}(\varphi) &= \int_{\frac{l_{\mathrm{p,GRG}}}{\varphi}}^{r_{\max}} -\frac{\xi+1}{l_{\mathrm{p,GRG}}}\left(\frac{\varphi r}{l_{\mathrm{p,GRG}}}\right)^{\xi} \cdot 3\left(\frac{r}{r_{\max}}\right)^3 \mathrm{d}r \\ &= -3\frac{\xi+1}{\xi+4} \cdot \frac{1}{\varphi_{\mathrm{GRG}}}\left(\left(\frac{\varphi}{\varphi_{\mathrm{GRG}}}\right)^{\xi} - \left(\frac{\varphi}{\varphi_{\mathrm{GRG}}}\right)^{-4}\right), \end{aligned} \tag{6.92}$$

where we assume $\xi \neq -4$. This PDF depends on ξ, $l_{\mathrm{p,GRG}}$, and $r_{\max}$ only. The CDF follows from direct integration:

$$F_{\Phi \mid L_\mathrm{p} > l_{\mathrm{p,GRG}}}(\varphi) = \begin{cases} 0 & \text{if } \varphi \leq \varphi_{\mathrm{GRG}} \\ 1 + \frac{1}{\xi+4}\left(-3\left(\frac{\varphi}{\varphi_{\mathrm{GRG}}}\right)^{\xi+1} - (\xi+1)\left(\frac{\varphi}{\varphi_{\mathrm{GRG}}}\right)^{-3}\right) & \text{if } \varphi > \varphi_{\mathrm{GRG}}, \end{cases} \tag{6.93}$$

just as the mean:

$$\mathbb{E}\left[\Phi \mid L_{\mathrm{p}} > l_{\mathrm{p,GRG}}\right] = 3\frac{\xi+1}{\xi+4}\left(\frac{1}{\xi+2}+\frac{1}{2}\right)\cdot\varphi_{\mathrm{GRG}}, \tag{6.94}$$

which exists only for $\xi < -2$. By solving

$$\frac{\mathrm{d}f_{\Phi} \mid L_{\mathrm{p}} > l_{\mathrm{p,GRG}}}{\mathrm{d}\varphi}\left(\varphi_{\mathrm{mode}}\right) = 0 \tag{6.95}$$

for the GRG angular length mode φ_{mode}, we find

$$\varphi_{\mathrm{mode}} = \left(\frac{-4}{\xi}\right)^{\frac{1}{\xi+4}}\cdot\varphi_{\mathrm{GRG}}. \tag{6.96}$$

There is no explicit expression for the associated median.

For a general non-Euclidean universe, no simple analytic form for the GRG angular length PDF appears to exist. We find an approximation valid at low redshifts by considering the Maclaurin polynomial of degree 1 for $z(r)$, the relation between comoving distance and cosmological redshift. One finds $z(r) \approx \frac{H_0}{c}r$. As a result, for the Local Universe, Eq. 6.88 becomes

$$\begin{aligned}\Phi \mid L_{\mathrm{p}} > l_{\mathrm{p,GRG}} &\approx L_{\mathrm{p}} \mid L_{\mathrm{p}} > l_{\mathrm{p,GRG}}\cdot\frac{1+\frac{H_0}{c}R}{R}\\ &= L_{\mathrm{p}} \mid L_{\mathrm{p}} > l_{\mathrm{p,GRG}}\cdot\left(\frac{1}{R}+\frac{1}{d_H}\right),\end{aligned} \tag{6.97}$$

where we use that the Hubble distance $d_H := \frac{c}{H_0}$. From a well-known inverse distribution identity,

$$f_{\frac{1}{R}}(k) = \begin{cases} 0 & \text{if } k \leq \frac{1}{r_{\max}} \\ \frac{3}{r_{\max}^3 k^4} & \text{if } k > \frac{1}{r_{\max}}, \end{cases} \tag{6.98}$$

so that

$$f_{\frac{1}{R}+\frac{1}{d_H}}(k) = \begin{cases} 0 & \text{if } k \leq \frac{1}{r_{\max}}+\frac{1}{d_H} \\ \frac{3}{r_{\max}^3\left(k-\frac{1}{d_H}\right)^4} & \text{if } k > \frac{1}{r_{\max}}+\frac{1}{d_H}. \end{cases} \tag{6.99}$$

Combining a well-known product distribution identity with the fact that $L_{\mathrm{p}} \mid L_{\mathrm{p}} >$

$l_{\mathrm{p,GRG}}$ and $\frac{1}{R}+\frac{1}{d_H}$ are independent RVs, we find

$$f_{\Phi \mid L_{\mathrm{p}}>l_{\mathrm{p,GRG}}}(\varphi)=\frac{-3\left(\xi+1\right)\varphi^{\xi}}{r_{\max}^{3}l_{\mathrm{p,GRG}}^{\xi+1}}\int_{\frac{1}{r_{\max}}+\frac{1}{d_H}}^{\frac{\varphi}{l_{\mathrm{p,GRG}}}}\frac{\mathrm{d}k}{k^{\xi+1}\left(k-\frac{1}{d_H}\right)^{4}} \tag{6.100}$$

when $\varphi>\varphi_{\mathrm{GRG}}+\frac{l_{\mathrm{p,GRG}}}{d_H}$. When $\varphi\leq\varphi_{\mathrm{GRG}}+\frac{l_{\mathrm{p,GRG}}}{d_H}$,$f_{\Phi \mid L_{\mathrm{p}}>l_{\mathrm{p,GRG}}}(\varphi)=0$. It is easy to verify that in the Euclidean limit, $\frac{1}{d_H}\to 0$, this expression reduces to that of Eq. 6.92.

6.A1.7 Maximum likelihood estimation of the tail index

How can we estimate ξ from a sample of N giants? Let ξ_{MLE} be the maximum likelihood estimate (MLE) of ξ. This RV is a function of N IID RVs $\{L_{\mathrm{p},1},\ldots,L_{\mathrm{p},N}\}\sim L_{\mathrm{p}}\mid L_{\mathrm{p}}>l_{\mathrm{p,GRG}}$. Define the following likelihood and log-likelihood functions:

$$\mathcal{L}\left(\xi\right):=\prod_{i=1}^{N}f_{L_{\mathrm{p}}\mid L_{\mathrm{p}}>l_{\mathrm{p,GRG}}}\left(L_{\mathrm{p,i}}\right) \tag{6.101}$$

$$\begin{aligned}\tilde{\mathcal{L}}\left(\xi\right)&:=\ln\left(\mathcal{L}\left(\xi\right)l_{\mathrm{p,GRG}}^{N}\right)\\&=N\ln\left(-\left(\xi+1\right)\right)+\xi\sum_{i=1}^{N}\ln\frac{L_{\mathrm{p,i}}}{l_{\mathrm{p,GRG}}}.\end{aligned} \tag{6.102}$$

We note the necessity to include a factor $l_{\mathrm{p,GRG}}^{N}$ in the definition of the log-likelihood to avoid a dimensionality error. The second derivative of $\tilde{\mathcal{L}}$ to ξ is

$$\frac{\mathrm{d}^2\tilde{\mathcal{L}}}{\mathrm{d}\xi^2}=-\frac{N}{\left(\xi+1\right)^2}<0. \tag{6.103}$$

Thus, if there exists a solution to the equation $\frac{\mathrm{d}\tilde{\mathcal{L}}}{\mathrm{d}\xi}=0$, it must correspond to a global *maximum* of the likelihood and log-likelihood functions:

$$\frac{\mathrm{d}\tilde{\mathcal{L}}}{\mathrm{d}\xi}\left(\xi=\xi_{\mathrm{MLE}}\right)=0,\text{ or }\xi_{\mathrm{MLE}}=-\frac{N}{\sum_{i=1}^{N}\ln\frac{L_{\mathrm{p,i}}}{l_{\mathrm{p,GRG}}}}-1. \tag{6.104}$$

6.A1.8 Observed projected proper length

General considerations

The model can be extended to incorporate observational selection effects. The relevant effects to consider vary from RG search campaign to RG search campaign, although some formulae apply in all cases. We derive these here.

To keep our extensions simple, we assume that the survey sensitivity is sufficient to detect RGs up to some redshift $z_{\max}$ only. In addition we assume that the projected proper length distribution does not evolve between $z = z_{\max}$ and $z = 0$. We let $p_{\text{obs}}(l_{\text{p}}, z)$ denote the probability that an RG of projected proper length l_{p} at cosmological redshift z is detected during the campaign. Also, r is the radial comoving distance and n is the total RG number density, counting the intrinsic number of RGs (irrespective of length) per unit of comoving volume. The observed number of RGs with projected proper length between l_{p} and $l_{\text{p}} + \mathrm{d}l_{\text{p}}$ throughout a survey covering a solid angle Ω is $\mathrm{d}N_{L_{\text{p}},\text{obs}}\left(l_{\text{p}}, \Omega\right)$, where

$$\begin{aligned}&\mathrm{d}N_{L_{\text{p}},\text{obs}}\left(l_{\text{p}}, \Omega\right)\\ &= \frac{\Omega}{4\pi}\int_0^{z_{\max}} n \cdot f_{L_{\text{p}}}\left(l_{\text{p}}\right)\mathrm{d}l_{\text{p}} \cdot p_{\text{obs}}\left(l_{\text{p}}, z\right) 4\pi r^2\left(z\right)\frac{\mathrm{d}r}{\mathrm{d}z}\,\mathrm{d}z.\end{aligned} \tag{6.105}$$

The total number of RGs with projected proper length between l_{p} and $l_{\text{p}} + \mathrm{d}l_{\text{p}}$ throughout a survey with solid angle Ω is $\mathrm{d}N_{L_{\text{p}}}\left(l_{\text{p}}, \Omega\right)$, where

$$\mathrm{d}N_{L_{\text{p}}}\left(l_{\text{p}}, \Omega\right) = \frac{\Omega}{4\pi}\int_0^{z_{\max}} n \cdot f_{L_{\text{p}}}\left(l_{\text{p}}\right)\mathrm{d}l_{\text{p}} \cdot 4\pi r^2\left(z\right)\frac{\mathrm{d}r}{\mathrm{d}z}\,\mathrm{d}z. \tag{6.106}$$

We define the completeness $C(l_{\text{p}}, z_{\max})$ to be

$$\begin{aligned}C(l_{\text{p}}, z_{\max}) &:= \frac{\mathrm{d}N_{L_{\text{p}},\text{obs}}\left(l_{\text{p}}, \Omega\right)}{\mathrm{d}N_{L_{\text{p}}}\left(l_{\text{p}}, \Omega\right)}\\ &= \frac{\int_0^{z_{\max}} p_{\text{obs}}\left(l_{\text{p}}, z\right) r^2\left(z\right) E^{-1}\left(z\right)\,\mathrm{d}z}{\int_0^{z_{\max}} r^2\left(z\right) E^{-1}\left(z\right)\,\mathrm{d}z}.\end{aligned} \tag{6.107}$$

The completeness only depends on the function $p_{\text{obs}}(l_{\text{p}}, z)$ and our choice of $z_{\max}$.

Let the RV $\mathcal{O}$ denote whether an RG picked at random within $z < z_{\max}$ is detected during the search campaign. We have $\mathcal{O} \sim \text{Bernoulli}\left(C(L_{\text{p}}, z_{\max})\right)$; the parameter that determines the distribution of $\mathcal{O}$ is itself an RV. It immediately follows that $\mathbb{P}\left(\mathcal{O} = 1 \mid L_{\text{p}} = l_{\text{p}}\right) = C(l_{\text{p}}, z_{\max})$.

If L_p would be discrete,

$$\begin{aligned}\mathbb{P}\left(L_\mathrm{p} = l_\mathrm{p} \mid \mathcal{O} = 1\right) &= \frac{\mathbb{P}\left(\mathcal{O} = 1\,, L_\mathrm{p} = l_\mathrm{p}\right)}{\mathbb{P}\left(\mathcal{O} = 1\right)} \\ &= \frac{\mathbb{P}\left(\mathcal{O} = 1 \mid L_\mathrm{p} = l_\mathrm{p}\right)\mathbb{P}\left(L_\mathrm{p} = l_\mathrm{p}\right)}{\sum_{l_\mathrm{p}} \mathbb{P}\left(\mathcal{O} = 1 \mid L_\mathrm{p} = l_\mathrm{p}\right)\mathbb{P}\left(L_\mathrm{p} = l_\mathrm{p}\right)}. \end{aligned} \tag{6.108}$$

Let $L_\mathrm{p,obs}$ be the observed projected proper length RV; that is $L_\mathrm{p,obs} := L_\mathrm{p} \mid \mathcal{O} = 1$. Its PDF is given by the continuous analogon of the preceding equation:

$$f_{L_\mathrm{p,obs}}\left(l_\mathrm{p}\right) = \frac{C\left(l_\mathrm{p}, z_\mathrm{max}\right) f_{L_\mathrm{p}}\left(l_\mathrm{p}\right)}{\int_0^\infty C\left(l_\mathrm{p}, z_\mathrm{max}\right) f_{L_\mathrm{p}}\left(l_\mathrm{p}\right)\,\mathrm{d}l_\mathrm{p}}. \tag{6.109}$$

Obviously, $F_{L_\mathrm{p,obs}}\left(l_\mathrm{p}\right) = \int_0^{l_\mathrm{p}} f_{L_\mathrm{p,obs}}\left(l'_\mathrm{p}\right)\,\mathrm{d}l'_\mathrm{p}$. The CDF $F_{L_\mathrm{p,obs}|L_\mathrm{p,obs}>l_\mathrm{p,GRG}}\left(l_\mathrm{p}\right)$ follows from an analogon of Eq. 6.54.

We note that multiplying $p_\mathrm{obs}(l_\mathrm{p}, z)$ with an l_p- and z-independent factor changes $C(l_\mathrm{p}, z_\mathrm{max})$ by the same factor, but leaves $f_{L_\mathrm{p,obs}}$ unaltered.

Fuzzy angular length threshold

Here we illustrate a simple extension. When performing visual searches for GRG candidates, a natural criterion is to only inspect sources with an angular length that is larger than some threshold. Researchers determine the threshold based on the amount of time available to them: lower thresholds will lead to more complete samples, but will take more time to collect. For humans, it is hard to estimate a source's angular length precisely by eye; as a result, some of the GRG candidates included in the project's GRG candidate catalogue will be sources with an angular length below the threshold, whilst others will be sources with an angular length above the threshold. We idealise this situation by asserting that sources with angular length φ_min or below are included in the catalogue with probability 0 (i.e. never), and that sources with an angular length φ_max or above are included in the catalogue with probability 1 (i.e. always). We assume a linear increase in probability as a function of φ for intermediate angular lengths: a source with angular length $\varphi_\mathrm{min} < \varphi < \varphi_\mathrm{max}$ is included in the catalogue with probability

$$p_\mathrm{obs}\left(l_\mathrm{p}, z\right) = \min\left\{\max\left\{\frac{\varphi\left(l_\mathrm{p}, z\right) - \varphi_\mathrm{min}}{\varphi_\mathrm{max} - \varphi_\mathrm{min}}, 0\right\}, 1\right\}. \tag{6.110}$$

In a flat Friedmann–Lemaître–Robertson–Walker (FLRW) universe, the angular length of an RG with projected proper length l_{p} at cosmological redshift z is

$$\varphi\left(l_{\mathrm{p}}, z\right) = \frac{l_{\mathrm{p}}\left(1+z\right)}{r\left(z\right)}. \tag{6.111}$$

Surface brightness limitations

Fanaroff–Riley class II If the surface brightness B_ν of the lobes is proportional to some negative power ζ of the GRG's proper length L, that is

$$B_\nu = b_{\nu,\mathrm{ref}}\left(\frac{L}{l_{\mathrm{ref}}}\right)^{\zeta}, \tag{6.112}$$

then B_ν is Pareto distributed — just like L:

$$F_{B_\nu}\left(b\right) = \begin{cases} 0 & \text{if } b \leq 0 \\ \left(\frac{b}{b_{\nu,\mathrm{ref}}}\right)^{\frac{\xi+1}{\zeta}} & \text{if } 0 < b < b_{\nu,\mathrm{ref}} \\ 1 & \text{if } b \geq b_{\nu,\mathrm{ref}}, \end{cases} \tag{6.113}$$

$$f_{B_\nu}\left(b\right) = \begin{cases} 0 & \text{if } b \leq 0 \\ \frac{\xi+1}{\zeta}\frac{1}{b_{\nu,\mathrm{ref}}}\left(\frac{b}{b_{\nu,\mathrm{ref}}}\right)^{\frac{\xi+1}{\zeta}-1} & \text{if } 0 < b < b_{\nu,\mathrm{ref}} \\ 0 & \text{if } b \geq b_{\nu,\mathrm{ref}}. \end{cases} \tag{6.114}$$

If RGs are self-similar in shape, then their lobe volumes are proportional to L^3. RGs appear to retain constant lobe luminosity density over most of their lifetime, so that their lobe monochromatic emission coefficients (Rybicki & Lightman, 1986) are proportional to L^{-3}. As line-of-sight lengths through the lobes are proportional to L, surface brightness is proportional to L^{-2}. These arguments thus suggest $\zeta = -2$.

It is a poor approximation to assume that all giants of proper length l_{ref} have the same surface brightness $b_{\nu,\mathrm{ref}}$: observations suggest a variability of several orders of magnitude. A better description is

$$B_\nu = b_{\nu,\mathrm{ref}}\left(\frac{L}{l_{\mathrm{ref}}}\right)^{\zeta} S, \tag{6.115}$$

where S is a lognormal RV whose median is 1. The PDF of S is then determined by

parameter σ_{ref}:

$$f_S(s) = \frac{1}{\sqrt{2\pi}\sigma_{\text{ref}} s} \exp\left(-\frac{\ln^2 s}{2\sigma_{\text{ref}}^2}\right). \tag{6.116}$$

The surface brightness CDF and PDF now become

$$F_{B_\nu}(b) = \begin{cases} 0 & \text{if } b \leq 0 \\ F_S\left(\frac{b}{b_{\nu,\text{ref}}}\right) + \left(\frac{b}{b_{\nu,\text{ref}}}\right)^{\frac{\xi+1}{\zeta}} \int_{\frac{b}{b_{\nu,\text{ref}}}}^{\infty} s^{-\frac{\xi+1}{\zeta}} f_S(s)\,\mathrm{d}s & \text{if } b > 0, \end{cases}$$

$$f_{B_\nu}(b) = \begin{cases} 0 & \text{if } b \leq 0 \\ \frac{\xi+1}{\zeta} \frac{1}{b_{\nu,\text{ref}}} \left(\frac{b}{b_{\nu,\text{ref}}}\right)^{\frac{\xi+1}{\zeta}-1} \int_{\frac{b}{b_{\nu,\text{ref}}}}^{\infty} s^{-\frac{\xi+1}{\zeta}} f_S(s)\,\mathrm{d}s & \text{if } b > 0. \end{cases} \tag{6.117}$$

We note that B_ν is not exactly Pareto distributed anymore.

In a relativistic rather than Euclidean universe, surface brightness is not constant with distance. To describe RGs beyond $z = 0$, we introduce a final model refinement:

$$B_\nu = \frac{b_{\nu,\text{ref}} \cdot S}{(1+Z)^{3-\alpha}} \left(\frac{L}{l_{\text{ref}}}\right)^{\zeta}, \tag{6.118}$$

where the RV Z denotes cosmological redshift and α is the spectral index of the lobes. We interpret $b_{\nu,\text{ref}} \cdot S$ as the (lognormally distributed) lobe surface brightness for RGs of intrinsic proper length l_{ref} at $z = 0$. In this case, the CDF and PDF of B_ν are most easily determined through sampling. To sample Z, we can first sample the comoving distance RV R instead, and subsequently use $Z = z_{\mathfrak{M}}(R)$. We stress that this conversion depends on cosmological parameters $\mathfrak{M}$.

To forward model a survey's surface brightness selection effect, we must compute

$$\begin{aligned} p_{\text{obs}}(l_{\text{p}}, z) &= \mathbb{P}\left(B_\nu > b_{\nu,\text{th}} \mid L_{\text{p}} = l_{\text{p}},\, Z = z\right) \\ &= 1 - F_{B_\nu \mid L_{\text{p}} = l_{\text{p}},\, Z = z}\left(b_{\nu,\text{th}}\right), \end{aligned} \tag{6.119}$$

where $b_{\nu,\text{th}} > 0$ is the surface brightness threshold. Typically, $b_{\nu,\text{th}}$ is comparable to the survey's RMS noise. What is $F_{B_\nu \mid L_{\text{p}} = l_{\text{p}},\, Z = z}(b)$? In the simplest case, devoid of S-

and Z-dependence,

$$
\begin{aligned}
F_{B_\nu \mid L_\mathrm{p}=l_\mathrm{p},\, Z=z}\left(b\right) &= \mathbb{P}\left(b_{\nu,\mathrm{ref}}\left(\frac{L_\mathrm{p}}{\sin\Theta\cdot l_\mathrm{ref}}\right)^{\zeta} \leq b \mid L_\mathrm{p} = l_\mathrm{p}\right) \\
&= \mathbb{P}\left(\sin^{-\zeta}\Theta \leq \frac{b}{b_{\nu,\mathrm{ref}}}\left(\frac{l_\mathrm{p}}{l_\mathrm{ref}}\right)^{-\zeta}\right) \\
&= \mathbb{P}\left(\sin\Theta \leq \tilde{s}^{-\frac{1}{\zeta}}\right) \\
&= \begin{cases} 0 & \text{if } b \leq 0 \\ 2F_\Theta\left(\arcsin\left(\tilde{s}^{-\frac{1}{\zeta}}\right)\right) & \text{if } 0 < b < b_{\nu,\mathrm{ref}}\left(\frac{l_\mathrm{p}}{l_\mathrm{ref}}\right)^{\zeta} \\ 1 & \text{if } b \geq b_{\nu,\mathrm{ref}}\left(\frac{l_\mathrm{p}}{l_\mathrm{ref}}\right)^{\zeta} \end{cases} \\
&= 1 - \sqrt{1 - \min\left\{\tilde{s}^{-\frac{2}{\zeta}}\left(\max\left\{b, 0\right\}\right), 1\right\}},
\end{aligned}
\tag{6.120}
$$

where

$$
\tilde{s} = \tilde{s}\left(b\right) := \frac{b}{b_{\nu,\mathrm{ref}}}\left(\frac{l_\mathrm{p}}{l_\mathrm{ref}}\right)^{-\zeta}. \tag{6.121}
$$

Therefore,

$$
p_\mathrm{obs}\left(l_\mathrm{p}, z\right) = \sqrt{1 - \min\left\{\tilde{s}^{-\frac{2}{\zeta}}\left(b_{\nu,\mathrm{th}}\right), 1\right\}}. \tag{6.122}
$$

In the most refined case, for $b > 0$,

$$
\begin{aligned}
&F_{B_\nu \mid L_\mathrm{p}=l_\mathrm{p},\, Z=z}\left(b\right) \\
&= \mathbb{P}\left(\frac{b_{\nu,\mathrm{ref}} \cdot S}{\left(1+Z\right)^{3-\alpha}} \left(\frac{L_\mathrm{p}}{\sin\Theta \cdot l_\mathrm{ref}}\right)^{\zeta} \leq b \mid L_\mathrm{p} = l_\mathrm{p},\, Z = z\right) \\
&= \mathbb{P}\left(\sin^{-\zeta}\Theta \cdot S \leq \frac{b}{b_{\nu,\mathrm{ref}}} \left(\frac{l_\mathrm{p}}{l_\mathrm{ref}}\right)^{-\zeta} \left(1+z\right)^{3-\alpha}\right) \\
&= \int_0^\infty \mathbb{P}\left(\sin^{-\zeta}\Theta \leq \frac{b}{b_{\nu,\mathrm{ref}}} \frac{1}{s} \left(\frac{l_\mathrm{p}}{l_\mathrm{ref}}\right)^{-\zeta} \left(1+z\right)^{3-\alpha}\right) f_S\left(s\right) \mathrm{d}s \\
&= \int_0^\infty \mathbb{P}\left(\sin\Theta \leq \left(\frac{\tilde{s}}{s}\right)^{-\frac{1}{\zeta}}\right) f_S\left(s\right) \mathrm{d}s \\
&= \int_0^{\tilde{s}} f_S\left(s\right) \mathrm{d}s + \int_{\tilde{s}}^\infty 2F_\Theta\left(\arcsin\left(\left(\frac{\tilde{s}}{s}\right)^{-\frac{1}{\zeta}}\right)\right) f_S\left(s\right) \mathrm{d}s \\
&= \int_0^{\tilde{s}} f_S\left(s\right) \mathrm{d}s + \int_{\tilde{s}}^\infty f_S\left(s\right) \mathrm{d}s \\
&- \int_{\tilde{s}}^\infty \cos\arcsin\left(\left(\frac{\tilde{s}}{s}\right)^{-\frac{1}{\zeta}}\right) f_S\left(s\right) \mathrm{d}s \\
&= 1 - \int_{\tilde{s}}^\infty \sqrt{1 - \left(\frac{\tilde{s}}{s}\right)^{-\frac{2}{\zeta}}} f_S\left(s\right) \mathrm{d}s,
\end{aligned}
\tag{6.123}
$$

where

$$
\tilde{s} = \tilde{s}\left(b\right) := \frac{b}{b_{\nu,\mathrm{ref}}} \left(\frac{l_\mathrm{p}}{l_\mathrm{ref}}\right)^{-\zeta} \left(1+z\right)^{3-\alpha}. \tag{6.124}
$$

The minimum value of S for which an RG of projected length l_p at redshift z is detectable, is $s_\mathrm{min}(l_\mathrm{p}, z) := \tilde{s}(b_{\nu,\mathrm{th}})$. For brevity, we simply write s_min instead. We find

$$
p_\mathrm{obs}\left(l_\mathrm{p}, z\right) = \int_{s_\mathrm{min}}^\infty \sqrt{1 - \left(\frac{s_\mathrm{min}}{s}\right)^{-\frac{2}{\zeta}}} f_S\left(s\right) \mathrm{d}s. \tag{6.125}
$$

The following approximation might facilitate the numerical evaluation of this integral. We note that for $s \gg s_\mathrm{min}$, the square root factor in the integral approaches 1. Now split up the original integral in two parts, where η governs the approximation's

accuracy:

$$
\begin{aligned}
p_{\text{obs}} &\approx \hat{p}_{\text{obs}} \\
&= \int_{s_{\min}}^{\eta s_{\min}} \sqrt{1-\left(\frac{s_{\min}}{s}\right)^{-\frac{2}{\zeta}}} f_S(s)\,\mathrm{d}s + \int_{\eta s_{\min}}^{\infty} f_S(s)\,\mathrm{d}s \\
&= \int_{s_{\min}}^{\eta s_{\min}} \sqrt{1-\left(\frac{s_{\min}}{s}\right)^{-\frac{2}{\zeta}}} f_S(s)\,\mathrm{d}s + 1 - F_S(\eta s_{\min}),
\end{aligned}
\tag{6.126}
$$

where we substitute numerically integrating to infinity for an evaluation of the CDF of the lognormally distributed RV S. The approximation error is bounded from above:

$$
\begin{aligned}
\hat{p}_{\text{obs}} - p_{\text{obs}} &= \int_{\eta s_{\min}}^{\infty} \left(1 - \sqrt{1-\left(\frac{s_{\min}}{s}\right)^{-\frac{2}{\zeta}}}\right) f_S(s)\,\mathrm{d}s \\
&< \left(1 - \sqrt{1-\left(\frac{s_{\min}}{\eta s_{\min}}\right)^{-\frac{2}{\zeta}}}\right) \int_{\eta s_{\min}}^{\infty} f_S(s)\,\mathrm{d}s \\
&= \left(1 - \sqrt{1-\eta^{\frac{2}{\zeta}}}\right)\left(1 - F_S(\eta s_{\min})\right) \\
&< 1 - \sqrt{1-\eta^{\frac{2}{\zeta}}}.
\end{aligned}
\tag{6.127}
$$

Let us assume $\zeta = -2$. For $\eta = 100$, the approximation error is at most 0.005, and for $\eta = 1000$, the approximation error is at most 0.0005.

Fanaroff–Riley class I The simplest correction in which FRI RGs retain a well-defined notion of length assumes a linearly decreasing surface brightness, from some value $b_\nu(0)$ at the core to zero at the RG's two endpoints. For a symmetric FRI RG, the surface brightness profile along one of the jets is $b_\nu : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, and depends on the projected proper distance from the core r_{p} as

$$
b_\nu(r_{\text{p}}) = \begin{cases} b_\nu(0)\left(1 - \frac{2r_{\text{p}}}{l_{\text{p}}}\right) & \text{if } 0 \leq r_{\text{p}} < \frac{l_{\text{p}}}{2} \\ 0 & \text{if } r_{\text{p}} \geq \frac{l_{\text{p}}}{2}. \end{cases}
\tag{6.128}
$$

Now we consider an FRI GRG, whose projected proper length $l_{\text{p}} > l_{\text{p,GRG}}$ would only be observed in full in the absence of noise. In actual observations, this GRG is

detected as a GRG if and only if

$$b_\nu\left(\frac{l_{\mathrm{p,GRG}}}{2}\right) > b_{\nu,\mathrm{th}}, \quad \text{or } b_\nu(0)\left(1-\frac{l_{\mathrm{p,GRG}}}{l_{\mathrm{p}}}\right) > b_{\nu,\mathrm{th}}. \tag{6.129}$$

Under our assumption of a linear surface brightness profile, the mean surface brightness along the jet axis $\langle b_\nu \rangle = b_\nu\left(\frac{1}{2}\cdot\frac{l_{\mathrm{p}}}{2}\right)$. As this must be half of the surface brightness at the core,

$$\langle b_\nu \rangle = b_\nu\left(\frac{l_{\mathrm{p}}}{4}\right) = \frac{b_\nu(0)}{2}. \tag{6.130}$$

Combining Eqs. 6.129 and 6.130, we find that the GRG is detected as such if

$$\langle b_\nu \rangle > \frac{b_{\nu,\mathrm{th}}}{2\left(1-\frac{l_{\mathrm{p,GRG}}}{l_{\mathrm{p}}}\right)}. \tag{6.131}$$

Now regarding $\langle b_\nu \rangle$ as an RV and recognising that it might behave exactly as in Eq. 6.118, we find that the surface brightness selection effect for FRI giants may be modelled as

$$p_{\mathrm{obs}}(l_{\mathrm{p}}, z) = \mathbb{P}\left(B_\nu > \frac{b_{\nu,\mathrm{th}}}{2\left(1-\frac{l_{\mathrm{p,GRG}}}{l_{\mathrm{p}}}\right)} \mid L_{\mathrm{p}} = l_{\mathrm{p}}, Z = z\right). \tag{6.132}$$

We see that the full formulaic structure of $p_{\mathrm{obs}}(l_{\mathrm{p}}, z)$ is the same for FRI and FRII giants, except that for FRI giants a change

$$b_{\nu,\mathrm{th}} \to \frac{b_{\nu,\mathrm{th}}}{2\left(1-\frac{l_{\mathrm{p,GRG}}}{l_{\mathrm{p}}}\right)} \tag{6.133}$$

is necessary. There is no change for $l_{\mathrm{p}} = 2l_{\mathrm{p,GRG}}$.

6.A1.9 GRG number density

A statistic of major interest is the number density of giants in the contemporary Universe. Let n_{GRG} be the comoving GRG number density, so that

$$n_{\mathrm{GRG}} = n\,\mathbb{P}\left(L_{\mathrm{p}} > l_{\mathrm{p,GRG}}\right) = n\left(1 - F_{L_{\mathrm{p}}}\left(l_{\mathrm{p,GRG}}\right)\right). \tag{6.134}$$

$$N_{\mathrm{GRG,obs}}\left(\Omega, z_{\max}\right) := \int_{l_{\mathrm{p,GRG}}}^{\infty} \mathrm{d}N_{L_{\mathrm{p,obs}}}\left(l_{\mathrm{p}}, \Omega\right). \tag{6.135}$$

After invoking Eq. 6.105 and isolating n, we obtain

$$n = \frac{\frac{4\pi}{\Omega} N_{\mathrm{GRG,obs}}\left(\Omega, z_{\max}\right)}{\int_{l_{\mathrm{p,GRG}}}^{\infty} \int_0^{z_{\max}} f_{L_{\mathrm{p}}}\left(l_{\mathrm{p}}\right) p_{\mathrm{obs}}\left(l_{\mathrm{p}}, z\right) 4\pi r^2\left(z\right) \frac{\mathrm{d}r}{\mathrm{d}z}\, \mathrm{d}z\, \mathrm{d}l_{\mathrm{p}}}. \tag{6.136}$$

Combining Eqs. 6.50, 6.51, 6.134, and 6.136 for $l_{\mathrm{p,GRG}} > l_{\min}$, we arrive at

$$\begin{aligned} & n_{\mathrm{GRG}} \\ & = -\frac{l_{\mathrm{p,GRG}}^{\xi+1}}{\xi+1} \cdot \frac{\frac{4\pi}{\Omega} N_{\mathrm{GRG,obs}}\left(\Omega, z_{\max}\right)}{\int_{l_{\mathrm{p,GRG}}}^{\infty} l_{\mathrm{p}}^{\xi} \int_0^{z_{\max}} p_{\mathrm{obs}}\left(l_{\mathrm{p}}, z\right) 4\pi r^2\left(z\right) \frac{\mathrm{d}r}{\mathrm{d}z}\, \mathrm{d}z\, \mathrm{d}l_{\mathrm{p}}}. \end{aligned} \tag{6.137}$$

From observations, we know Ω and can — for a given $z_{\max}$ — simply count $N_{\mathrm{GRG,obs}}$ $(\Omega, z_{\max})$. Moreover, we can fit ξ and the parameters that occur in $p_{\mathrm{obs}}(l_{\mathrm{p}}, z)$ (e.g. $\frac{1}{2}(\varphi_{\max} - \varphi_{\min})$, $b_{\nu,\mathrm{ref}}$, and σ_{ref}) to the ECDF of $L_{\mathrm{p,obs}} \mid L_{\mathrm{p,obs}} > l_{\mathrm{p,GRG}}$. We note that n_{GRG} does not depend on $l_{\min}$, which drops out through the division. However, n_{GRG} does depend on cosmological parameters through the relation between cosmological redshift z and radial comoving distance r.

6.A1.10 GRG lobe volume-filling fraction

Assuming self-similar growth, the combined proper volume V of an RG's lobes and its intrinsic proper length l obey $V \propto l^3$. The constant of proportionality varies per RG and depends on the shape of the lobes; we treat it as an RV $\Upsilon := \frac{V}{L^3}$. Then the proper VFF of GRG lobes $\mathrm{VFF}_{\mathrm{GRG}}(z) = \mathrm{VFF}_{\mathrm{GRG}}(z=0) \cdot (1+z)^3$, where

$$\begin{aligned} \mathrm{VFF}_{\mathrm{GRG}}(z=0) &:= n_{\mathrm{GRG}} \cdot \mathbb{E}[V \mid L_{\mathrm{p}} > l_{\mathrm{p,GRG}}] \\ &= n_{\mathrm{GRG}} \cdot \mathbb{E}[\frac{V}{L^3} \cdot L^3 \mid L_{\mathrm{p}} > l_{\mathrm{p,GRG}}] \\ &= n_{\mathrm{GRG}} \cdot \mathbb{E}[\Upsilon \cdot L^3 \mid L_{\mathrm{p}} > l_{\mathrm{p,GRG}}]. \end{aligned} \tag{6.138}$$

Assuming that RGs grow self-similarly, so that shape does not reveal length, Υ and L^3 are conditionally independent given $L_{\mathrm{p}} > l_{\mathrm{p,GRG}}$: $\Upsilon \perp\!\!\!\perp L^3 \mid L_{\mathrm{p}} > l_{\mathrm{p,GRG}}$. As a

result,

$$\begin{aligned}\mathrm{VFF}_{\mathrm{GRG}}(z=0) &= n_{\mathrm{GRG}} \cdot \mathbb{E}[\Upsilon \mid L_{\mathrm{p}} > l_{\mathrm{p,GRG}}] \cdot \mathbb{E}[L^3 \mid L_{\mathrm{p}} > l_{\mathrm{p,GRG}}] \\ &= n_{\mathrm{GRG}} \cdot \mathbb{E}[\Upsilon] \cdot \mathbb{E}[L^3 \mid L_{\mathrm{p}} > l_{\mathrm{p,GRG}}]. \end{aligned} \tag{6.139}$$

To obtain this last line, we once more exploit self-similarity: $\Upsilon \mid (L_{\mathrm{p}} > l_{\mathrm{p,GRG}}) = \Upsilon$. We can approximate $\mathbb{E}[\Upsilon]$ by taking the mean of some $\frac{V}{l^3}$ deduced from observations. A technical complication arises from the fact that $\mathbb{E}[L^3 \mid L_{\mathrm{p}} > l_{\mathrm{p,GRG}}]$ does not exist for $\xi \geq -4$ under our model. This is an artefact of the Pareto distribution assumption for L, which unrealistically features support over an infinitely long part of the real line: $\{l \in \mathbb{R} \mid l > l_{\mathrm{min}}\}$. This causes the expectation value integral to diverge for $\xi \geq -4$. An approximation to $\mathrm{VFF}_{\mathrm{GRG}}(z=0)$ that works for all ξ is the lower bound

$$\mathrm{VFF}_{\mathrm{GRG}}(z=0) > n_{\mathrm{GRG}} \cdot \mathbb{E}[\Upsilon] \cdot \mathbb{E}^3[L \mid L_{\mathrm{p}} > l_{\mathrm{p,GRG}}], \tag{6.140}$$

which follows from Jensen's inequality. Here

$$\mathbb{E}[L \mid L_{\mathrm{p}} > l_{\mathrm{p,GRG}}] = l_{\mathrm{p,GRG}} \frac{\Gamma\left(-\frac{\xi}{2}-1\right)\Gamma\left(-\frac{\xi}{2}+1\right)}{\Gamma\left(-\frac{\xi}{2}-\frac{1}{2}\right)\Gamma\left(-\frac{\xi}{2}+\frac{1}{2}\right)}. \tag{6.141}$$

Alternative approximation formulae, which use $\Upsilon_{\mathrm{p}} := \frac{V}{L_{\mathrm{p}}^3}$, are

$$\mathrm{VFF}_{\mathrm{GRG}}(z=0) > n_{\mathrm{GRG}} \cdot \mathbb{E}[\Upsilon_{\mathrm{p}}] \cdot \mathbb{E}^3[L_{\mathrm{p}} \mid L_{\mathrm{p}} > l_{\mathrm{p,GRG}}] \tag{6.142}$$

and

$$\mathrm{VFF}_{\mathrm{GRG}}(z=0) > n_{\mathrm{GRG}} \cdot \mathbb{E}[\Upsilon_{\mathrm{p}}] \cdot m_{L_{\mathrm{p}}^3 \mid L_{\mathrm{p}} > l_{\mathrm{p,GRG}}}, \tag{6.143}$$

where

$$m_{L_{\mathrm{p}}^3 \mid L_{\mathrm{p}} > l_{\mathrm{p,GRG}}} = l_{\mathrm{p,GRG}}^3 \cdot 2^{\frac{3}{-(\xi+1)}} \tag{6.144}$$

is the median of the cubed projected proper length for giants. An advantage of these latter expressions is that there are more data available to estimate $\mathbb{E}[\Upsilon_{\mathrm{p}}]$ than there are to estimate $\mathbb{E}[\Upsilon]$.

6.A1.11 Unification model constraints from quasar and non-quasar giants

The quasar GRG probability p_{Q} is

$$\begin{aligned}
p_{\mathrm{Q}} :=& \frac{\mathbb{P}\left(L_{\mathrm{p,obs}} \geq l_{\mathrm{p,GRG}},\ \sin\Theta \leq \sin\theta_{\max}\right)}{\mathbb{P}\left(L_{\mathrm{p,obs}} \geq l_{\mathrm{p,GRG}}\right)} \\
=& \mathbb{P}\left(\sin\Theta \leq \sin\theta_{\max} \mid L_{\mathrm{p,obs}} \geq l_{\mathrm{p,GRG}}\right) \\
=& \int_0^{\sin\theta_{\max}} \frac{\mathbb{P}\left(L_{\mathrm{p,obs}} \geq l_{\mathrm{p,GRG}} \mid \sin\Theta = x\right) f_{\sin\Theta}(x)\,\mathrm{d}x}{\mathbb{P}\left(L_{\mathrm{p,obs}} \geq l_{\mathrm{p,GRG}}\right)} \\
=& \int_0^{\sin\theta_{\max}} \frac{\mathbb{P}\left(L_{\mathrm{p}} \geq l_{\mathrm{p,GRG}} \mid \sin\Theta = x\right) f_{\sin\Theta}(x)\,\mathrm{d}x}{\mathbb{P}\left(L_{\mathrm{p}} \geq l_{\mathrm{p,GRG}}\right)} \\
=& \int_0^{\sin\theta_{\max}} \frac{\left(1 - F_L\left(\frac{l_{\mathrm{p,GRG}}}{x}\right)\right) f_{\sin\Theta}(x)\,\mathrm{d}x}{1 - F_{L_{\mathrm{p}}}(l_{\mathrm{p,GRG}})}.
\end{aligned} \tag{6.145}$$

Now using Eqs. 6.45 and 6.50 and $f_{\sin\Theta}(x) = \frac{x}{\sqrt{1-x^2}}$ over the domain of integration,

$$p_{\mathrm{Q}} = \frac{4\Gamma\left(-\frac{\xi}{2}+1\right)}{-(\xi+1)\sqrt{\pi}\,\Gamma\left(-\frac{\xi}{2}-\frac{1}{2}\right)} \int_0^{\sin\theta_{\max}} \frac{x^{-\xi}\,\mathrm{d}x}{\sqrt{1-x^2}}. \tag{6.146}$$

6.A1.12 Extreme giants in a sample

An interesting feature of the model is its ability to predict the occurrence of giants with extreme projected proper lengths in a GRG sample of, say, size N. Now consider some $l_{\mathrm{p}} > l_{\mathrm{p,GRG}}$ — what is the probability $p_{>l_{\mathrm{p}}}$ that an observed GRG will have a projected proper length exceeding l_{p}? Proceeding as in the derivation of Eq. 6.54, we find

$$\begin{aligned}
p_{>l_{\mathrm{p}}}\left(l_{\mathrm{p}}\right) :=& \mathbb{P}\left(L_{\mathrm{p,obs}} > l_{\mathrm{p}} \mid L_{\mathrm{p,obs}} > l_{\mathrm{p,GRG}}\right) \\
=& \frac{1 - F_{L_{\mathrm{p,obs}}}\left(l_{\mathrm{p}}\right)}{1 - F_{L_{\mathrm{p,obs}}}\left(l_{\mathrm{p,GRG}}\right)}.
\end{aligned} \tag{6.147}$$

In the absence of selection effects, $p_{>l_{\mathrm{p}}}\left(l_{\mathrm{p}}\right)$ is given by Eq. 6.57. The number of giants with extreme projected proper lengths $N_{>l_{\mathrm{p}}} \sim \mathrm{Binom}(N, p_{>l_{\mathrm{p}}}(l_{\mathrm{p}}))$.

Interesting questions can be answered readily. For example, the probability that

the sample contains at least one GRG with projected proper length l_p or larger, is

$$\mathbb{P}\left(N_{>l_\mathrm{p}} \geq 1\right) = 1 - \mathbb{P}\left(N_{>l_\mathrm{p}} = 0\right) = 1 - \left(1 - p_{>l_\mathrm{p}}\left(l_\mathrm{p}\right)\right)^N. \tag{6.148}$$

6.A2 Additional images

In this appendix, as in Fig. 6.12, we show newly discovered giants. These cover the projected length range $l_\mathrm{p} \in [0.7\ \mathrm{Mpc}, 4\ \mathrm{Mpc})$.

6.A3 Stellar and supermassive black hole masses

In Fig. 6.27, we present the SDSS-derived relations between host stellar mass and projected proper length, and between host supermassive black hole mass and projected proper length, for all giants in our final catalogue. We obtained the data as in Section 3.7 of Oei et al. (2022a).

6.A4 Surface brightness prior

Consider a radio galaxy at cosmological redshift z of intrinsic proper length l bounded by spherical lobes of radius R and spectral index α. If a fraction f_l of the radio galaxy's central axis lies within the lobes, then

$$R = \frac{1}{4} f_l \cdot l. \tag{6.149}$$

If a fraction f_{L_ν} of the radio galaxy's total luminosity density L_ν comes from the lobes, then the monochromatic emission coefficient

$$j_\nu = \frac{L_\nu}{4\pi} \frac{f_{L_\nu}}{2} \frac{1}{V}, \tag{6.150}$$

where the lobe volume $V = \frac{4}{3}\pi R^3$. A formula of practical value features projected proper length l_p instead of l. Given the approximate nature of our approach, we therefore simply assume $l \approx \mathbb{E}[D](\eta(f_l)) \cdot l_\mathrm{p}$, with the deprojection factor expectation $\mathbb{E}[D]$ given in Eq. 6.71; $\eta(f_l) = \frac{f_l}{2-f_l}$. The maximum surface brightness of the lobes as seen by an observer is

$$b_{\nu,\mathrm{max}} = \frac{3 f_{L_\nu} \cdot L_\nu}{\pi^2 \cdot \mathbb{E}[D](\eta(f_l)) \cdot f_l^2 \cdot l_\mathrm{p}^2 \cdot (1+z)^{3-\alpha}}, \tag{6.151}$$

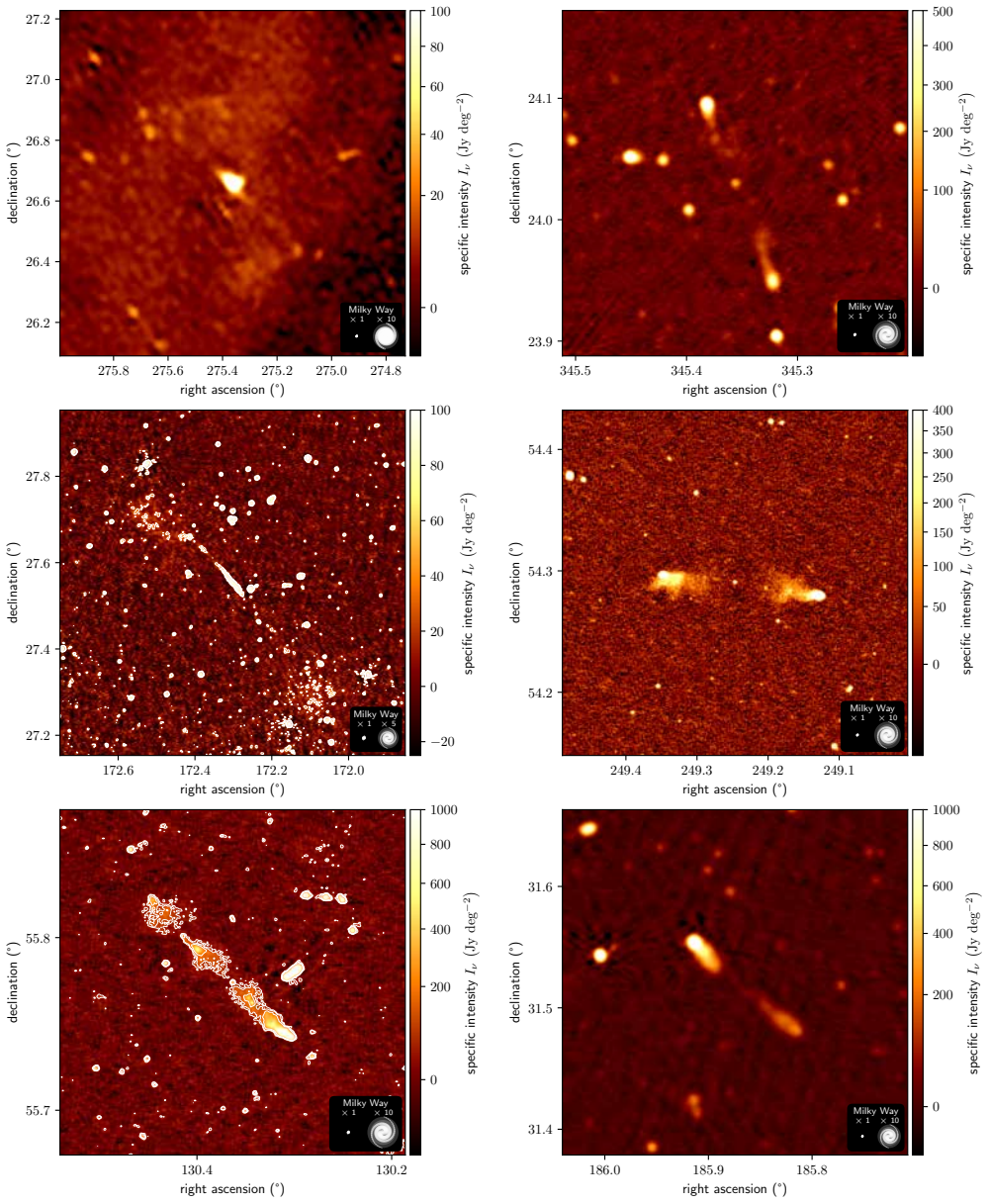


Figure 6.24: Details of the LoTSS DR2–estimated specific intensity function $I_\nu(\hat{r})$ at central observing frequency $\nu_{\rm obs} = 144$ MHz and resolutions $\theta_{\rm FWHM} \in \{6'', 20'' \; 90''\}$, centred around the hosts of newly discovered giants. Row-wise from left to right, from top to bottom, the projected proper length $l_{\rm p}$ is 3.9 Mpc, 3.5 Mpc, 3.3 Mpc, 3.3 Mpc, 3.3 Mpc, and 3.2 Mpc; in the same order, $\theta_{\rm FWHM}$ is $90''$, $20''$, $20''$, $6''$, $6''$, and $20''$. The giants in the top-left and middle-left panels appear larger in the sky than the Moon. Contours signify 3, 5, and 10 sigma-clipped standard deviations above the sigma-clipped median. For scale, we show the stellar Milky Way disk (with a diameter of 50 kpc) generated using the Ringermacher & Mead (2009) formula, alongside a 5 or 10 times inflated version.

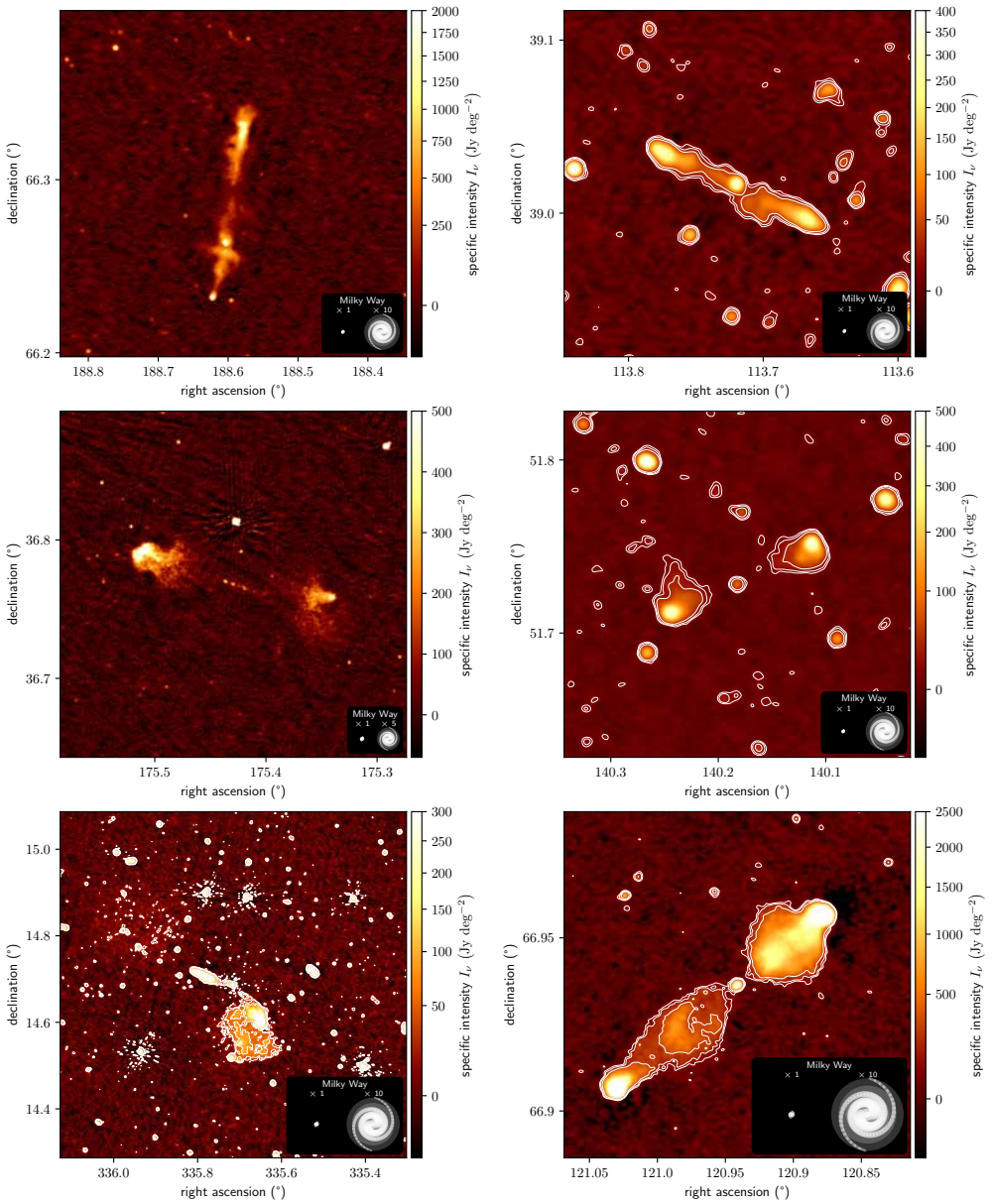


Figure 6.25: Details of the LoTSS DR2–estimated specific intensity function $I_\nu(\hat{r})$ at central observing frequency $\nu_{\rm obs} = 144$ MHz and resolutions $\theta_{\rm FWHM} \in \{6'', 20''\}$, centred around the hosts of newly discovered giants. Row-wise from left to right, from top to bottom, the projected proper length $l_{\rm p}$ is 2.8 Mpc, 2.6 Mpc, 2.6 Mpc, 2.2 Mpc, 2.1 Mpc, and 2.1 Mpc; in the same order, $\theta_{\rm FWHM}$ is $6''$, $20''$, $6''$, $20''$, $20''$, and $6''$. The GRG in the bottom-left panel appears larger in the sky than the Moon. Contours signify 3, 5, and 10 sigma-clipped standard deviations above the sigma-clipped median. For scale, we show the stellar Milky Way disk (with a diameter of 50 kpc) generated using the Ringermacher & Mead (2009) formula, alongside a 5 or 10 times inflated version.

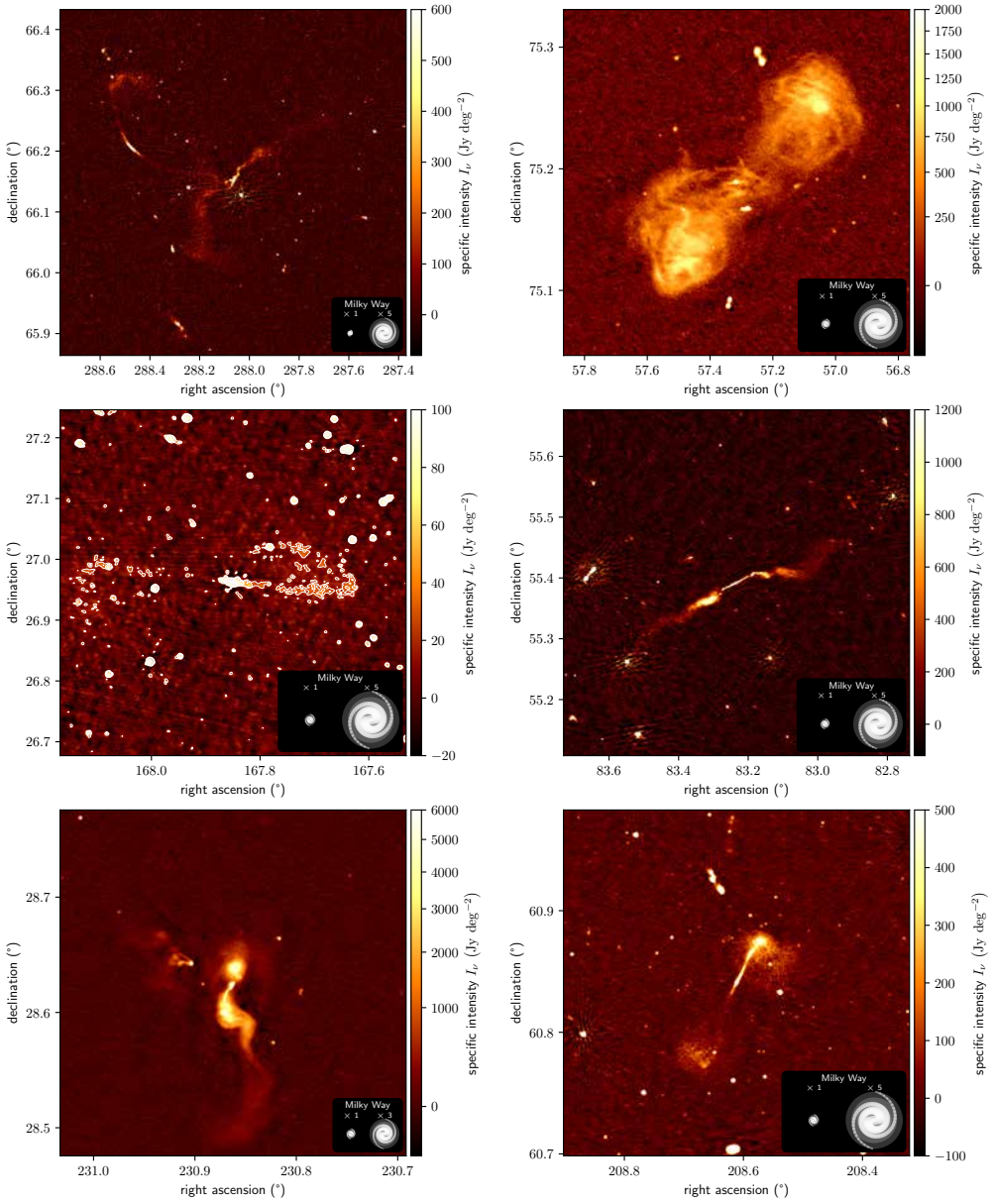


Figure 6.26: Details of the LoTSS DR2–estimated specific intensity function $I_\nu(\hat{r})$ at central observing frequency $\nu_{\rm obs} = 144$ MHz and resolutions $\theta_{\rm FWHM} \in \{6'', 20''\}$, centred around the hosts of newly discovered giants. Row-wise from left to right, from top to bottom, the projected proper length $l_{\rm p}$ is 1.6 Mpc, 1.5 Mpc, 1.3 Mpc, 1.2 Mpc, 1.1 Mpc, and 0.7 Mpc; in the same order, $\theta_{\rm FWHM}$ is $6''$, $6''$, $20''$, $6''$, $6''$, and $6''$. Contours signify 3, 5, and 10 sigma-clipped standard deviations above the sigma-clipped median. For scale, we show the stellar Milky Way disk (with a diameter of 50 kpc) generated using the Ringermacher & Mead (2009) formula, alongside a 3 or 5 times inflated version.

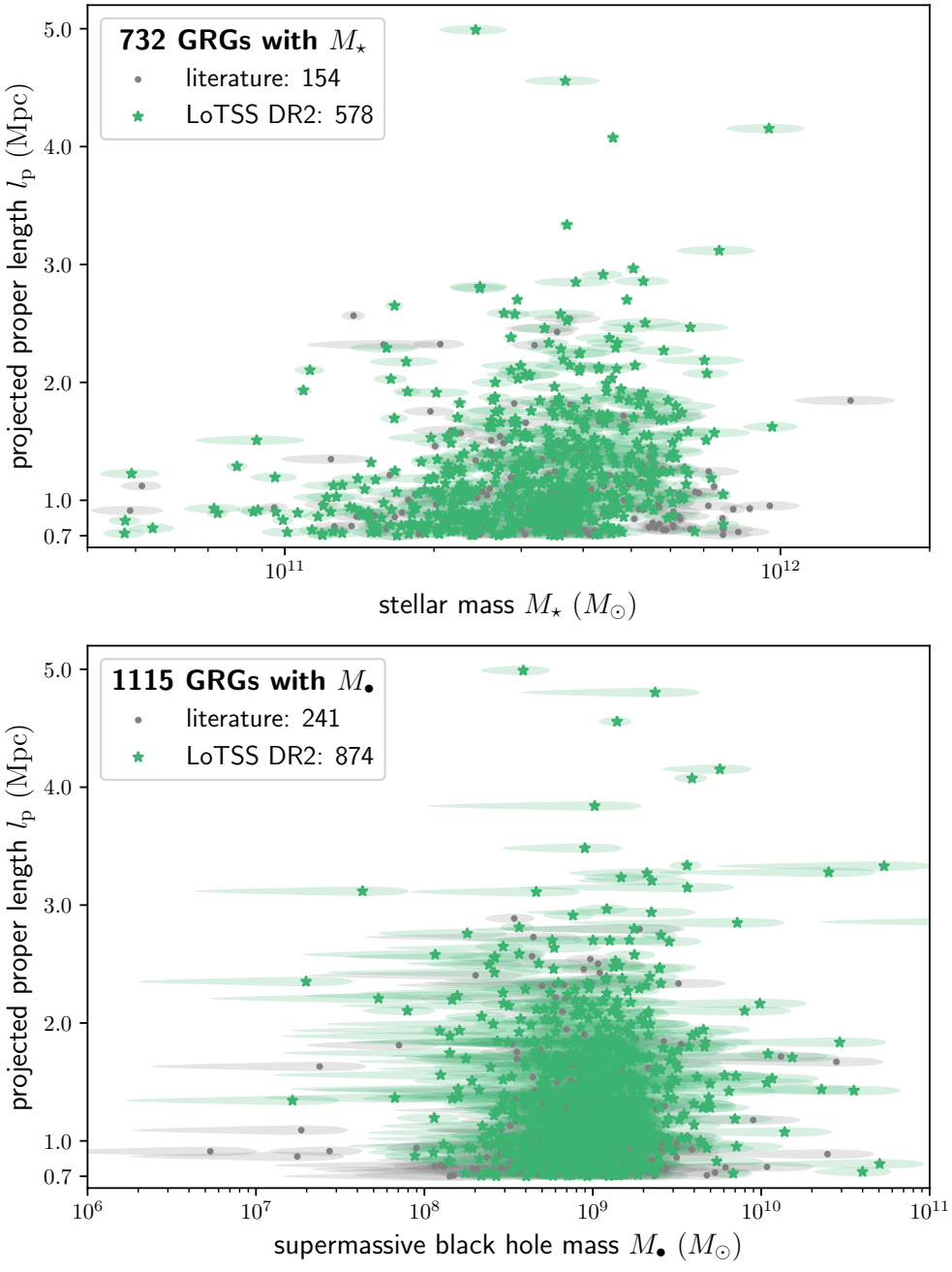


Figure 6.27: Observed relations between host stellar mass $M_\star$ and GRG projected length l_p *(top)* and between host SMBH mass $M_\bullet$ and GRG projected length l_p *(bottom)*. Our LoTSS DR2 sample confirms that luminous giants typically have $M_\star \in 10^{11}$–10^{12} $M_\odot$ and $M_\bullet \in 10^{8}$–10^{10} $M_\odot$. The sample effects an almost fivefold increase in the number of giants with SDSS-derived host stellar masses and SMBH masses. We do not show or count giants for which only a nearest host candidate could be determined.

valid for the line of sight that pierces through a lobe along a diameter. The average line of sight length d within a lobe is smaller than $d_{\max} = 2R$, though, and given by

$$\langle d \rangle := \frac{\int_0^R 2\sqrt{R^2 - x^2} 2\pi x \, \mathrm{d}x}{\pi R^2} = \frac{4}{3} R. \tag{6.152}$$

Therefore, the mean surface brightness of the lobes as seen by an observer is

$$\langle b_\nu \rangle = \frac{\langle d \rangle}{d_{\max}} b_{\nu,\max} = \frac{2}{3} b_{\nu,\max}. \tag{6.153}$$

6.A5 Likelihood function

In Table 6.5, we present maximum likelihood and likelihood mean and standard deviation estimates for the inference described in Section 6.4.1; one may compare the results to those in Table 6.4. In Fig. 6.28, we visually summarise the likelihood function; one may compare to Fig. 6.16.

The strong degeneracy between ξ and $b_{\nu,\mathrm{ref}}$, directly apparent from the central two-parameter marginal in the leftmost column of Fig. 6.28, translates to a ridge of essentially constant likelihood that extends from $\xi \approx -4$ to $\xi \approx -2$. Compared to a non-degenerate case, this makes the maximum likelihood parameters both intrinsically less meaningful and more prone to numerical approximation errors.

Table 6.5: Maximum likelihood estimate (MLE) and likelihood mean and standard deviation (SD) estimates of the free parameters in intrinsic GRG length distribution inference.[25]

$z_{\max} = 0.5$:

parameter	MLE	likelihood mean and SD
ξ	-2.15	-3.4 ± 0.5
$\frac{1}{2}(\varphi_{\max} - \varphi_{\min})$	$1.7'$	$1.9 \pm 0.3'$
$b_{\nu,\mathrm{ref}}$	$90\,\mathrm{Jy\,deg^{-2}}$	$660 \pm 400\,\mathrm{Jy\,deg^{-2}}$
σ_{ref}	1.35	1.2 ± 0.2

$z_{\max} = 0.25$:

parameter	MLE	likelihood mean and SD
ξ	-2.2	-3.5 ± 0.4
$\frac{1}{2}(\varphi_{\max} - \varphi_{\min})$	$2.25'$	$1.7 \pm 0.3'$
$b_{\nu,\mathrm{ref}}$	$150\,\mathrm{Jy\,deg^{-2}}$	$1050 \pm 560\,\mathrm{Jy\,deg^{-2}}$
σ_{ref}	1.25	1.4 ± 0.4

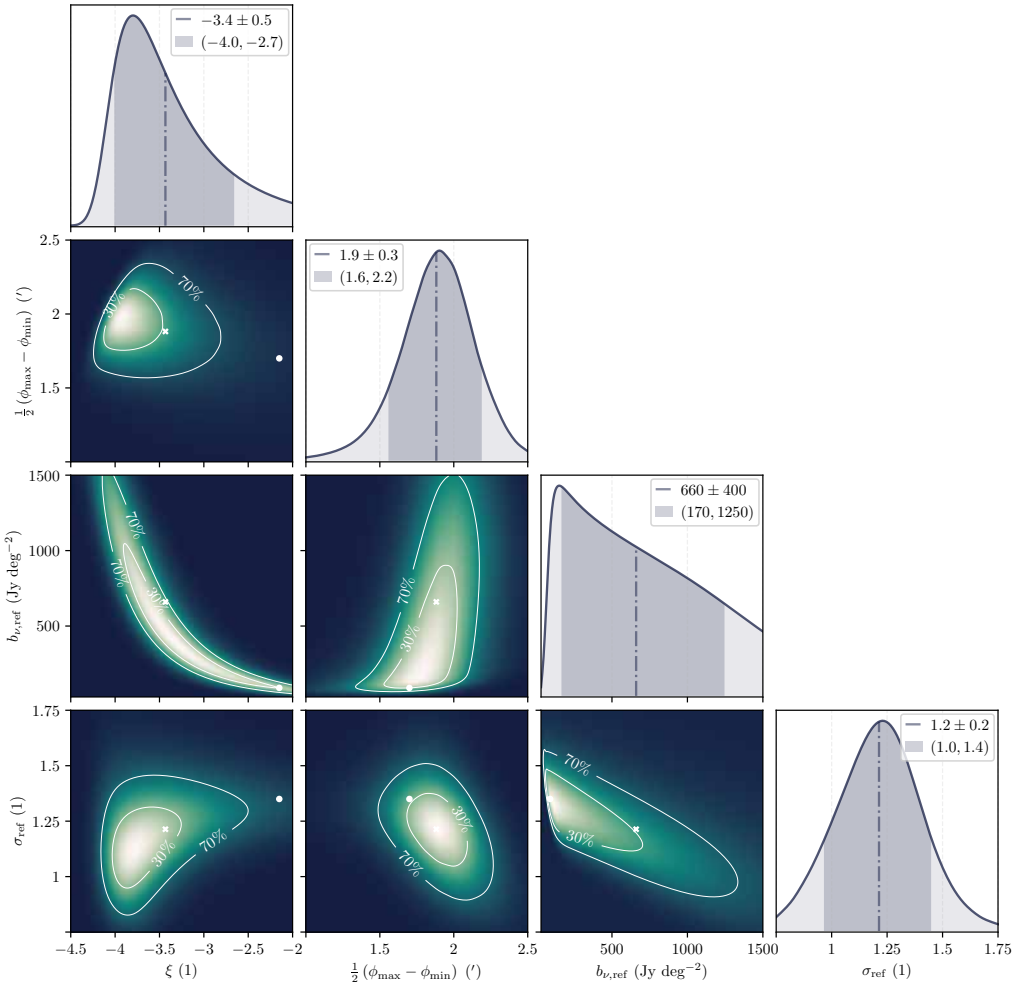


Figure 6.28: Joint likelihood function over ξ — the parameter of interest — and $\frac{1}{2}(\varphi_{\max} - \varphi_{\min})$, $b_{\nu,\mathrm{ref}}$ and σ_{ref} — the selection effect parameters, based on 1473 projected lengths of LoTSS DR2 giants up to $z_{\max} = 0.5$. We show all two-parameter marginals of the likelihood function, with contours enclosing 30% and 70% of total probability. We mark the maximum likelihood parameters (white circle) and the likelihood mean parameters (white cross). The single-parameter marginals again show the estimated mean, now marked by a vertical line, alongside shaded median-centred 80% credible intervals. The likelihood function is the posterior for a uniform prior. To compare the likelihood function to the posterior actually chosen, see Fig. 6.16.

[25] The model assumes ξ is constant for $z \in [0, z_{\max}]$. We determined the likelihood function twice: for $z_{\max} = 0.5$, using 1473 giants, and for $z_{\max} = 0.25$, using 811 giants. We caution that, in this case, the MLE parameters are a poor measure of central tendency.

6.A6 Properties of newly discovered giants

Table 6.2 provides properties of the 50 projectively longest giants discovered during this work's LoTSS DR2 search campaign. We share these data, alongside those for the other 2010 (98%) giants in our sample, in Flexible Image Transport System (FITS) format through the Centre de Données astronomiques de Strasbourg (CDS). One can either use anonymous File Transfer Protocol (FTP) to

`ftp://cdsarc.cds.unistra.fr` (130.79.128.5)

or visit

`https://cdsarc.cds.unistra.fr/cgi-bin/qcat?J/A+A/`.

For our final catalogue, which also includes literature giants, please contact the authors.

They are ill discoverers that think there is no land, when they can see nothing but sea.

Francis Bacon, English philosopher, *The Advancement of Learning (1605)*

7

Do luminous giants populate special large-scale environments? Or: the radio luminosity–Cosmic Web density relation for radio galaxies

M. S. S. L. Oei, R. J. van Weeren, M. J. Hardcastle, A. R. D. J. G. I. B. Gast, F. Leclercq, H. J. A. Röttgering, P. Dabhade, T. W. Shimwell, A. Botteon — *Astronomy & Astrophysics*, accepted

Abstract

Context Giant radio galaxies (GRGs, giant RGs, or giants) are megaparsec-scale, jet-driven outflows from accretion disks of supermassive black holes, and represent the most extreme pathway by which galaxies can impact the Cosmic Web around them. A long-standing but unresolved question is why giants are so much larger than other radio galaxies.

Aims It has been proposed that, in addition to having higher jet powers than most RGs, giants might live in especially low-density Cosmic Web environments. In this work, we aim to test this hypothesis by pinpointing Local Universe giants and other RGs in physically principled, Bayesian large-scale structure reconstructions.

METHODS More specifically, we localise a LOFAR Two-metre Sky Survey (LoTSS) DR2–dominated sample of luminous ($l_\nu(\nu = 150\ \mathrm{MHz}) \geq 10^{24}\ \mathrm{W\,Hz^{-1}}$) giants and a control sample of LoTSS DR1 RGs, both with spectroscopic redshifts up to $z_{\mathrm{max}} = 0.16$, in the BORG SDSS Cosmic Web reconstructions. We measure the Cosmic Web density on a smoothing scale of $\sim$2.9 Mpc h^{-1} for each RG; for the control sample, we then quantify the relation between RG radio luminosity and Cosmic Web density. With the BORG SDSS tidal tensor, we also measure for each RG whether the gravitational dynamics of its Cosmic Web environment resemble those of clusters, filaments, sheets, or voids.

RESULTS For both luminous giants and general RGs, the Cosmic Web density distribution is gamma distribution–like. Luminous giants populate large-scale environments that tend to be denser than those of general RGs. This result is corroborated by gravitational dynamics classification and a cluster catalogue cross-matching analysis. We find that the Cosmic Web density around RGs with 150 MHz radio luminosity l_ν is distributed as $1+\Delta_{\mathrm{RG}} \,|\, L_\nu = l_\nu \sim \Gamma(k, \theta)$, where $k = 4.8+0.2\cdot\lambda$, $\theta = 1.4+0.02\cdot\lambda$, and $\lambda := \log_{10}(l_\nu\,(10^{23}\ \mathrm{W\,Hz^{-1}})^{-1})$.

CONCLUSIONS This work presents more than a thousand inferred Mpc-scale densities around radio galaxies, which may be correct up to a factor of order unity — except in clusters of galaxies, where the densities can be more than an order of magnitude too low. We pave the way to a future in which Mpc-scale densities around RGs are common inferred quantities, which help to better understand their dynamics, morphology, and interaction with the enveloping Cosmic Web. Our data demonstrate that luminous giants inhabit denser environments than general RGs. This underlines that — at least at high jet powers — low-density environments are no prerequisite for giant growth. Using general RGs, we quantify the relation between radio luminosity at 150 MHz and Cosmic Web density on a smoothing scale of $\sim$2.9 Mpc h^{-1}. This positive relation, combined with the discrepancy in radio luminosity between known giants and general RGs, reproduces the discrepancy in Cosmic Web density between known giants and general RGs. Our findings are consistent with the view that giants are regular, rather than mechanistically special, members of the radio galaxy population.

Key words: radio continuum: galaxies – galaxies: active – jets – inter-galactic medium – large-scale structure of Universe

7.1 Introduction

Supermassive Kerr black holes are key building blocks of the Universe, on galactic and cosmological scales alike. During episodes of baryon accretion, they turn into active galactic nuclei (AGN), launching winds and jets that warm and rarefy the interstellar medium (e.g. Fabian, 2012; King & Pounds, 2015). These energy flows generally suppress the formation of new stars, although local star formation enhancement can occur within expanding kiloparsec-radius rings (e.g. Dugan et al., 2017). Meanwhile, jet-mediated AGN outflows — also known as radio galaxies (RGs) — can have a vast, megaparsec-scale reach, protruding from both the galaxy and its circumgalactic medium. The synchrotron radiation from RGs illuminates their jets, lobes, cocoons, and threads (e.g. Ramatsoku et al., 2020), and dominates the extragalactic radio sky. Because the behaviour of RGs links to the physics of black hole accretion and galactic winds, the pressure field and magnetisation history (e.g. Vazza et al., 2017) of the warm–hot intergalactic medium (warm–hot IGM, or WHIM), and shocks (e.g. Nolting et al., 2019a), vorticity (e.g. Nolting et al., 2019b), and cooling flows (Fabian et al., 1984) in the intracluster medium, a precise understanding of the RG phenomenon is indispensable to modern astrophysics.

A key goal of the study of RGs is to identify the main factors that determine their dynamics and to formulate models (e.g. Scheuer, 1974; Turner & Shabala, 2015; Hardcastle, 2018) that describe them quantitatively. One way to investigate the growth of RGs is to search for particularly large examples and to analyse what internal or external traits set them apart from the rest. Following this logic, Oei et al. (2022a) presented and studied *Alcyoneus*, a giant radio galaxy (GRG, giant RG, or simply *giant*) whose proper length component in the plane of the sky $l_\mathrm{p} = 4.99 \pm 0.04$ Mpc. In general, giants are members of the RG population for which $l_\mathrm{p} \geq l_\mathrm{p,GRG}$ — where the latter is some fixed megaparsec-scale threshold — and that therefore rank among the largest RGs in existence. However, despite being one of the largest RGs known, Alcyoneus is not particularly luminous, and its host galaxy does not feature a particularly massive central black hole or stellar population — at least, when compared to other giants and their hosts (Oei et al., 2022a). The question naturally arises whether external properties, rather than those internal to the host galaxy, are the most important drivers of RG growth.

It is well established that the IGM resists the growth of RGs by forcing jets to convert a part of their kinetic energy into work spent to form lobe cavities (e.g. Hardcastle & Croston, 2020). The work needed to free up a cavity is the product of its volume and the local pressure.[1] Following this line of reasoning, an RG with fixed intrin-

[1] As the pressure field in modern large-scale structure is not constant, as predicted by approximate

sic properties should reach a larger end-of-life extent in more tenuous (and colder) Cosmic Web (CW) environments. It has therefore been proposed that the astonishing growth of giants might be explained by their presumptive tendency to reside in tenuous parts of the Cosmic Web.

Several previous works have investigated the role of the enveloping Cosmic Web density field on GRG growth. Notably, in all cases, the authors traced Cosmic Web environments through three-dimensional galaxy positions estimated with photometric or spectroscopic redshifts. The pioneering work of Subrahmanyan et al. (2008) presented a case study of GRG MSH 05-22 and its Cosmic Web environment as traced by 6dF data (Jones et al., 2004). Using a sample of 12 giants and environments traced by 2dF/AAOmega data (Sharp et al., 2006), Malarecki et al. (2015) concluded that the lobes of giants grow in directions that avoid denser regions of the Cosmic Web. In the most comprehensive study yet, Lan & Prochaska (2021) used a sample of 110 giants and environments traced by DESI Legacy Imaging Surveys DR9 data (Dey et al., 2019), and did not find evidence that giants occur in more dilute environments than non-giant RGs. A major limitation of this latest study is the use of photometric redshifts as a probe of the Cosmic Web density field.

In order to determine decisively how GRG growth and the Cosmic Web relate, analyses with both more giants and more accurate Cosmic Web reconstructions seem necessary. In this work, we intend to present major developments in both regards. Firstly, the Low-Frequency Array (LOFAR; van Haarlem et al., 2013) Two-metre Sky Survey (LoTSS; Shimwell et al., 2017) has made possible the discovery of thousands of previously unknown giants in its Northern Sky imagery at observing frequency $\nu_{\rm obs} = 144$ MHz and at resolutions $\theta_{\rm FWHM} \in \{6'', 20'', 60'', 90''\}$. In particular, the joint search efforts of Dabhade et al. (2020b) in the LoTSS DR1 (Shimwell et al., 2019), those of Simonte et al. (2022) in the LoTSS Boötes Deep Field, and those of Oei et al. (2023a) in the LoTSS DR2 (Shimwell et al., 2022) have tripled the total number of known giants, which now stands at $\sim 3 \cdot 10^3$. In particular, the number of known giants in the part of the Local Universe covered by the Sloan Digital Sky Survey (SDSS; York et al., 2000) DR7 (Abazajian et al., 2009) has quintupled — a fact whose relevance will become clear in Sect. 7.2.3. Secondly, the last two decades have seen the development of principled, physics-based Bayesian inference techniques through which the three-dimensional total (i.e. baryonic *plus* dark) matter density field of the Cosmic Web can be reconstructed (e.g. Wandelt et al., 2004; Kitaura & Enßlin, 2008; Jasche et al., 2010b; Jasche & Kitaura, 2010b; Jasche & Wandelt, 2013; Jasche & Lavaux, 2019). These techniques make use of the fact that the statistical

hydrostatic equilibrium, this expression only holds for sufficiently small volumes.

behaviour of the Early Universe's total matter density field is known, as is the dominant process by which this field has evolved over cosmic time: gravity. By simulating gravity acting on Early Universe density fields and comparing the evolved fields to the observed spatial distribution of galaxies, the late-time density field can be inferred. These late-time density fields subsequently enable megaparsec-scale density measurements, including uncertainties, around individual (giant) radio galaxies. We shall use these measurements as an improved probe of the Cosmic Web density field.

Section 7.2 presents the data we use: radio galaxy observables and late-time density field reconstructions via which we probe the influence of the Cosmic Web on the growth of giants. In Sect. 7.3, we combine these data to determine megaparsec-scale densities and large-scale structure type probability distributions for giants and general RGs. Section 7.4 presents these results, alongside a quantification of the relation between RG radio luminosity and Cosmic Web density. We finally evaluate evidence for the claim that the Cosmic Web affects RG growth. In Sect. 7.5, we test the reliabilty of RG Cosmic Web density measurements and dynamical classifications and discuss caveats of and promising future extensions to the current work, before we present conclusions in Sect. 7.6.

For consistency with Oei et al. (2022a), Oei et al. (2023a), and Mostert & Oei (2023), we assume a flat, inflationary ΛCDM model with parameters from Planck Collaboration et al. (2020): $h = 0.6766$, $\Omega_{\mathrm{BM},0} = 0.0490$, $\Omega_{\mathrm{M},0} = 0.3111$, and $\Omega_{\Lambda,0} = 0.6889$, where $\Omega_{\mathrm{DM},0} := \Omega_{\mathrm{M},0} - \Omega_{\mathrm{BM},0} = 0.2621$ and $H_0 := h \cdot 100\ \mathrm{km\ s^{-1}\ Mpc^{-1}}$. With 'Local Universe', we refer to the spherical region of space observed to have redshift $z < z_{\mathrm{max}} := 0.16$. All reported redshifts are heliocentric. Terminology-wise, we strictly distinguish an RG (a radio-bright structure of plasma and magnetic fields, consisting of a core, jets, hotspots, lobes, a cocoon, and collimated synchrotron threads) from the host galaxy that has generated it. As in our previous work, we define giants to be RGs with projected proper[2] lengths $l_{\mathrm{p}} \geq l_{\mathrm{p,GRG}} := 0.7\ \mathrm{Mpc}$. We define the spectral index α so that it relates to flux density F_ν at frequency ν as $F_\nu \propto \nu^\alpha$; under this convention, radio spectral indices are typically negative.

[2] In Cosmic Web filament environments, where giants appear most common (Sect. 7.4), lobes may expand along the Hubble flow, rendering their proper and comoving extents different. To avoid ambiguity, we stress that our projected lengths are proper instead of comoving. A less precise synonym for 'projected proper length' often found in the literature is 'largest linear size' (LLS).

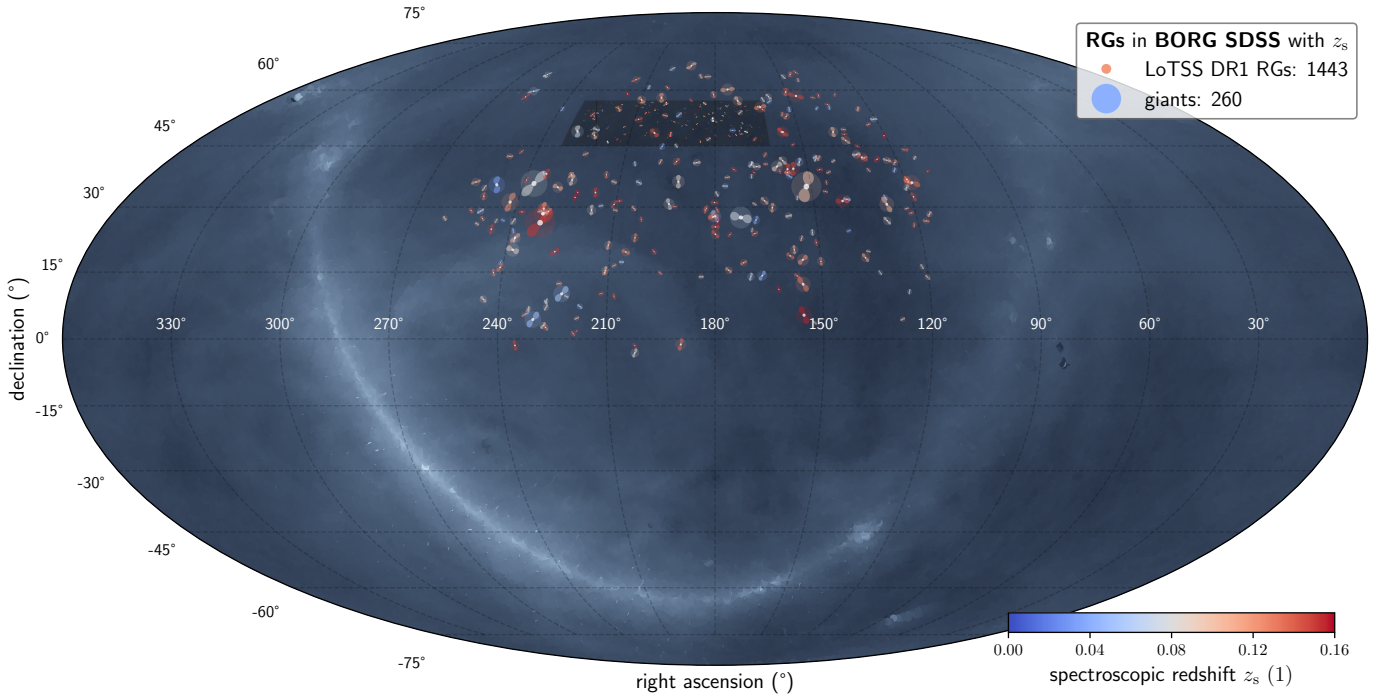


Figure 7.1: Mollweide view of the sky showing the locations of all giants and LoTSS DR1 RGs in the Local Universe for which we infer Cosmic Web densities and dynamical states. The background shows the Milky Way at 150 MHz (Zheng et al., 2017), on which we overlay the LoTSS DR1 footprint (hatched dark rectangle). RGs are drawn as lemniscates of Bernoulli; we do not attempt to portray realistic morphologies or position angles. The colours represent redshifts $z \in (0, z_{\mathrm{max}} := 0.16)$, whilst the diameters are proportional to projected proper lengths $l_{\mathrm{p}} \in (\sim 1\,\mathrm{kpc}, 4.6\,\mathrm{Mpc})$. Giants are translucent. Upon zooming in, the reader can appreciate the wide variety of sizes that radio galaxies can attain.

7.2 Data

To compare the Cosmic Web environments of Local Universe giants to those of surrounding smaller radio galaxies, we combined a GRG catalogue, a general RG catalogue, and a Cosmic Web density field reconstruction. All three data sets are publicly available.

7.2.1 Giant radio galaxies

As our source of giants we used the catalogue aggregated by Oei et al. (2023a), which contains 3341 giants with $l_{\mathrm{p}} \geq l_{\mathrm{p,GRG}} := 0.7$ Mpc.[3] This catalogue is intended to be complete up to (and including) September 2022. Thanks to the steradian-scale Northern Sky coverage of the LoTSS, combined with its arcsecond-scale resolution and sensitivity up to degree scales, LoTSS-discovered giants (Dabhade et al., 2020b; Simonte et al., 2022; Oei et al., 2023a) dominate the catalogue.

In their LoTSS DR2 manual visual search for giants, Oei et al. (2023a) used an angular length threshold of $\frac{1}{2}\left(\varphi_{\mathrm{max}} - \varphi_{\mathrm{min}}\right) = 5'$ to limit the duration of their search to a manageable few hundred hours. At the same time, this $5'$-threshold ensured that the search would yield most giants with sufficient surface brightness in the Local Universe.[4] This has been by design: Oei et al. (2023a) aimed to build a surface brightness–limited, but otherwise complete census of Local Universe giants in the LoTSS DR2 footprint with the intent of localising them within the Cosmic Web density field reconstructions presented in Sect. 7.2.3.

As elaborated upon in Sect. 7.3.1, we retained 281 giants in the part of the Local Universe where Cosmic Web analysis is possible. Of these, 260 have spectroscopic redshifts; only these giants could be reliably localised. We show their sky locations in Fig. 7.1. LoTSS DR2 discoveries make up 208 of the final 260 giants (80%). Through six example LoTSS DR2 giants, Fig. 7.2 provides the reader a sense of the quality of the radio imagery underpinning this work, from which angular lengths φ have been inferred, a sense of the reliability of our host galaxy identification, from which spectroscopic redshifts z_{s} have inferred, and a sense of the morphological and surface brightness diversity of the objects under consideration in this study. This publication is the first to contain images of the objects shown.

[3] Oei et al. (2023a) provide references to all discovery articles.

[4] Oei et al. (2023a) illustrate this point in their Fig. 9.

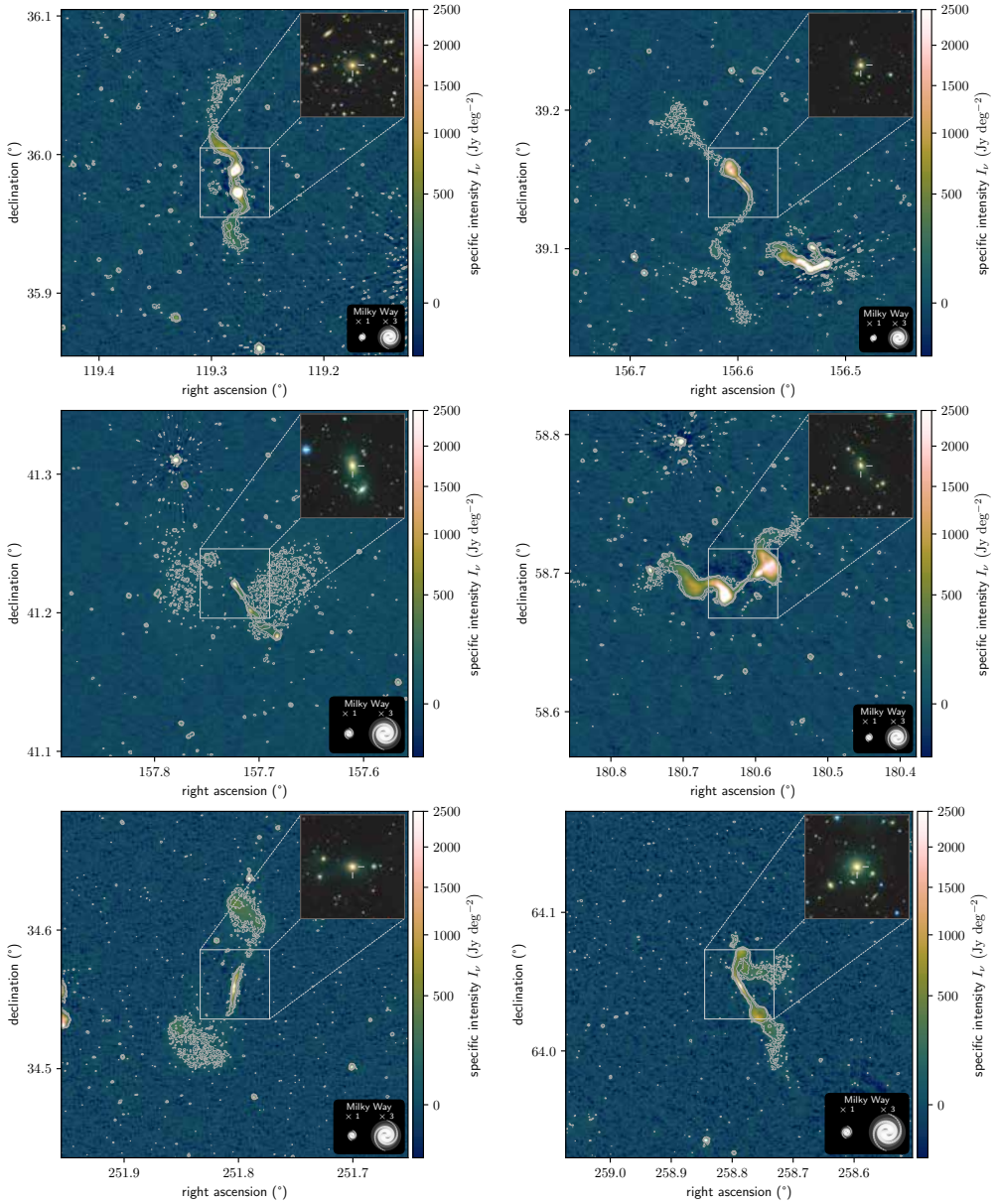


Figure 7.2: LoTSS DR2 cutouts at central observing frequency $\nu_{\rm obs} = 144$ MHz and resolution $\theta_{\rm FWHM} = 6''$, centred around giant-generating BCGs of Local Universe clusters. All giants shown are discoveries of Oei et al. (2023a) for which no previous images have been published. Each cutout covers a solid angle of $15' \times 15'$. Contours signify 3, 5, and 10 sigma-clipped standard deviations above the sigma-clipped median. For scale, we show the stellar Milky Way disk (with a diameter of 50 kpc) generated using the Ringermacher & Mead (2009) formula, alongside a 3 times inflated version. Each DESI Legacy Imaging Surveys DR9 (g, r, z) inset shows the central $3' \times 3'$ region.

7.2.2 General radio galaxies

In order to determine whether the Cosmic Web environments of giants are exceptional in any way, we must create a sample of reference Cosmic Web environments. For this reason, we also localised within the Cosmic Web a sample of general RGs, selected without regard for their length.

In particular, our starting point for this sample is the radio-bright active galactic nucleus (RLAGN) sub-sample described by Hardcastle et al. (2019), which contains 23,344 of the 318,520 sources (7%) in the LoTSS DR1 value-added catalogue of Williams et al. (2019). To construct their sub-sample, Hardcastle et al. (2019) combined multiple criteria that separate RLAGN from star-forming galaxies (SFGs). At the low redshifts considered in this work ($z < z_{\mathrm{max}} := 0.16$), most RLAGN are identified through SDSS spectroscopy, and we therefore expect high RLAGN completeness and minimal contamination from SFGs.

As elaborated upon in Sect. 7.3.1, we retained 1870 LoTSS DR1 RGs in the part of the Local Universe where Cosmic Web analysis is possible. Of these, 1443 have spectroscopic redshifts; only these RGs could be reliably localised. We show their sky locations in Fig. 7.1. The LoTSS DR1 image quality is very similar to the LoTSS DR2 image quality, which Fig. 7.2 illustrates.

7.2.3 Cosmic Web late-time density field

To localise RGs within the Cosmic Web, we used data products from the Bayesian Origin Reconstruction from Galaxies (BORG; Jasche & Wandelt, 2013) SDSS run (Jasche et al., 2015). The BORG SDSS uses second-order Lagrangian perturbation theory (2LPT; Bouchet et al., 1995) to forward model structure formation and evaluates the plausibility of a proposed structure formation history by comparing its late-time density field to the three-dimensional positions of galaxies in the SDSS DR7 Main Galaxy Sample (MGS; Abazajian et al., 2009). As a result, the BORG SDSS provides a late-time total[5] matter density field posterior for the part of the Local Universe covered by the SDSS DR7 footprint. The posterior is represented by a Hamiltonian Monte Carlo (HMC; Duane et al., 1987) Markov chain of approximately ten thousand samples after the burn-in phase. Each sample covers the same volume of $(750\ \mathrm{Mpc}\ h^{-1})^3$ extent with a cubical grid of 256^3 voxels. Thus, the side length of a BORG SDSS voxel $L = \frac{1}{256} \cdot 750\ \mathrm{Mpc}\ h^{-1} \approx 2.9\ \mathrm{Mpc}\ h^{-1}$. The late-time posterior can be compactly summarised by taking the mean and standard deviation (SD) of the samples on a per-voxel basis. In this work, we will use both the individual samples and

[5] The BORG algorithm does not differentiate between baryonic and dark matter, assuming identical behaviour on the multi-megaparsec scale.

the summary cubes, the latter of which we shall call the BORG SDSS mean and SD. However, by using these summary cubes, information contained in higher moments of single-voxel posteriors and inter-voxel correlations remains unused.

Leclercq et al. (2015) have extended the BORG SDSS by calculating, for each sample, the Cosmic Web classification as stipulated by the T-web definition (Hahn et al., 2007). The result is a probabilistic classification with a marginal distribution for each voxel. As one aspect of a broader information theoretic analysis, Leclercq et al. (2016) demonstrated how these classifications can be used to predict galaxy properties, such as their $g - r$ colours. We now use these classifications to characterise the dynamical environments of luminous giants and general RGs.

Both the BORG SDSS posterior and a posterior from another BORG run, the BORG 2M++ (Jasche & Lavaux, 2019), have been used before to relate properties of active galaxies to the density of the enveloping Cosmic Web (Frank et al., 2016; Porqueres et al., 2018). Our work is the first to relate properties of radio galaxies to BORG (or BORG-like) Cosmic Web reconstructions.

7.3 Methods

As pointed out in Sect. 7.1, the number of known giants has increased substantially in recent years. For example, within the Local Universe covered by the SDSS DR7, the manual visual search of Oei et al. (2023a) alone has quintupled the number of known giants with spectroscopic redshifts — from 52 to 260. Meanwhile, the BORG SDSS now offers the first physically principled, probabilistic reconstruction of the total matter density field over this entire volume. In this section, we combine both advances by describing the localisation of giants and general RGs within the BORG SDSS.

7.3.1 BORG SDSS localisation procedure

For each RG in our two samples (be it a giant or a general RG), we first transformed right ascension, declination, and redshift into a vector with comoving coordinates $\mathbf{r}_i := [x_i, y_i, z_i]^\top$ following the coordinate system convention of Jasche et al. (2015). We note that this transformation is cosmology-dependent. The BORG SDSS adopts the cosmological parameters of Jasche et al. (2015) to convert SDSS DR7 MGS redshifts into radial comoving distances and subsequently infer the structure formation history of the Local Universe. In order to obtain valid localisations, it is imperative that we use the same conversion between redshift and radial comoving distance for

our RGs. We therefore adopted the Jasche et al. (2015) cosmology for this particular procedure. Afterwards, we simply associated each RG to the voxel nearest to $\mathbf{r}_i$.

After localising an RG to a voxel, we verified whether the late-time density field at this location is sufficiently constrained by the BORG SDSS. More precisely, we evaluated whether — at the RG's voxel — the survey response operator of Jasche et al. (2015)'s lowest r-band absolute magnitude bin ($-21.00 < M_r < -20.33$), R^0, equals or exceeds some threshold $R^0_{\rm min}$. If indeed $R^0(\mathbf{r}_i) \geq R^0_{\rm min}$, we retained the RG for Cosmic Web analysis; if not, we discarded it. In this work, we chose $R^0_{\rm min} = 0.1$. Higher choices for $R^0_{\rm min}$ rid the sample of comparatively uninformative (i.e. more prior-dominated) density and dynamical state measurements, but come at the cost of a reduced sample size.

Next, in order to obtain density distributions for each retained RG, we explored two methods.

Fixed voxel method

In the 'fixed voxel method', we considered for each RG only the *marginal* posterior density distribution at its voxel — i.e. the distribution for that voxel in isolation — despite the fact that the BORG SDSS provides reconstructions with complex inter-voxel density correlations. Moreover, we summarised the voxel's marginal distribution through the mean and SD only. This method is the simplest of the two methods we have used.

In Fig. 7.3, we show six example localisations of giants in the Local Universe. The giants in the upper two panels are pinpointed to galaxy clusters, the giants in the middle two panels are pinpointed to filaments, and the giants in the bottom two panels are pinpointed to sheets.

Flexible voxel method

There are at least two issues associated with the fixed voxel method, though. One issue is the fact that the BORG SDSS has inferred the Cosmic Web with a limited set of bright SDSS DR7 galaxies, which (amongst other factors) causes reconstruction uncertainty. In practice, this means that a given cluster or filament may morph and wiggle around in different BORG SDSS samples. If one does not wiggle around the voxel to sample from accordingly (but sticks with the same voxel all the time), one regularly samples outside of the cluster or filament in which the RG of interest resides. This, of course, biases the inferred densities low.

On top of Cosmic Web reconstruction uncertainty, host galaxy peculiar motion uncertainty leads to additional difficulty in the determination of an RG's Cosmic

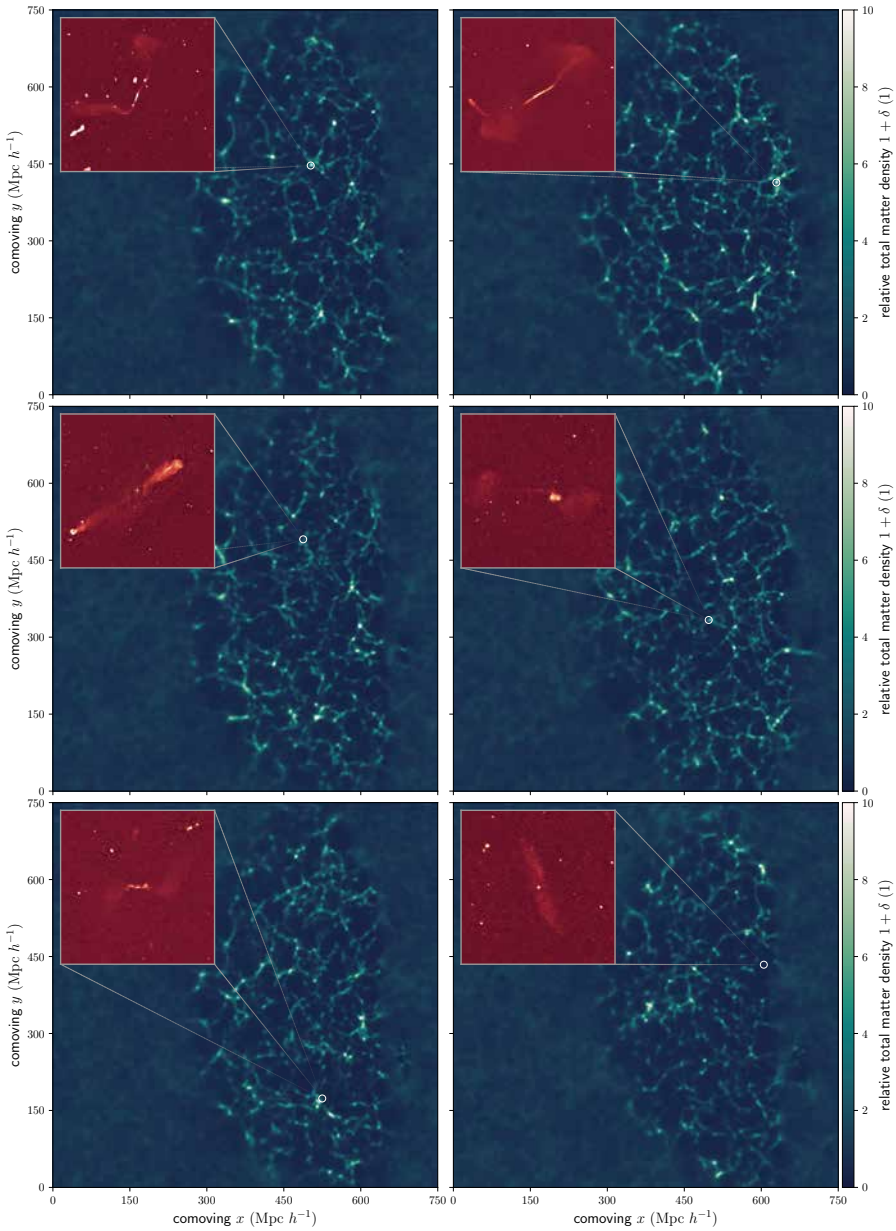


Figure 7.3: Example localisations of giant radio galaxies within the large-scale structure of the Local Universe. The top row shows two cluster giants, the middle row shows two filament giants, and the bottom row shows two sheet giants. For each giant, we show a slice of constant Cartesian comoving z through the late-time BORG SDSS posterior mean total matter density field and a LoTSS DR2 $6''$ image at $\nu_{\mathrm{obs}} = 144$ MHz (inset). Outside of the SDSS DR7–constrained volume, the posterior mean tends to the Universe's late-time mean total matter density $\bar{\rho}_0$. The locations of the giants are marked by white circles.

Web density. In Appendix 7.A1, we show that redshift uncertainties generally lead to sub-voxel localisation uncertainty when spectroscopic redshifts are used, even after taking low-mass galaxy cluster–like peculiar motion into account. For high-mass galaxy clusters however, peculiar motion can cause localisation errors of several voxels. Even more alarmingly, we show that photometric redshift uncertainties cause localisation uncertainties of 10^1–10^2 Mpc, or up to tens of voxels, that are unworkably large. Cosmic Web localisation therefore only seems possible for RGs with spectroscopically detected hosts.

To counteract the fixed voxel method's tendency to sample densities outside of the clusters and filaments in which our RGs truly reside, we propose a flexible voxel method. In this method, the voxel considered hitherto serves as a reference voxel, around which we search for the most likely correct voxel to sample from. More precisely, we iterate over 1000 BORG SDSS samples equally spaced within the MCMC, and consider for each sample all voxels in a sphere of radius 5 Mpc h^{-1}. We then simply adopt the voxel with the highest density as the most likely correct voxel for that sample. This procedure encapsulates our prior knowledge that the massive ellipticals which give rise to radio galaxies are more likely to form in a high-density region than in a low-density region of the same extent.

The search radius of 5 Mpc h^{-1} chosen here is arbitrary to some degree. Clearly, for the flexible voxel method to be any different from the fixed voxel method, this radius must exceed the side length of a single voxel, 2.9 Mpc h^{-1}. For larger radii, we are able to correct for larger peculiar motion errors, and thus provide more accurate densities for RGs located in massive galaxy clusters. At the same time, for larger radii, we are at risk of straying too far from the reference voxel; for example, this could lead to sampling cluster-like densities for RGs that actually reside in an adjacent filament. Because the majority of RGs appears to reside in filaments, and because the gravity solver of the BORG SDSS significantly limits the usefulness of cluster densities anyways (see Sect. 7.5), we chose a relatively 'small' search radius of 5 Mpc h^{-1} — less than two voxels in each direction.

In Fig. 7.4, we compare the mean Cosmic Web densities inferred for all RGs considered in this work, with fixed voxel method densities on the horizontal axis, and flexible voxel method densities on the vertical axis. Flexible voxel–based relative densities are typically a factor two higher. The methods agree most often at the lowest redshifts, where the BORG SDSS Cosmic Web reconstructions are least uncertain.

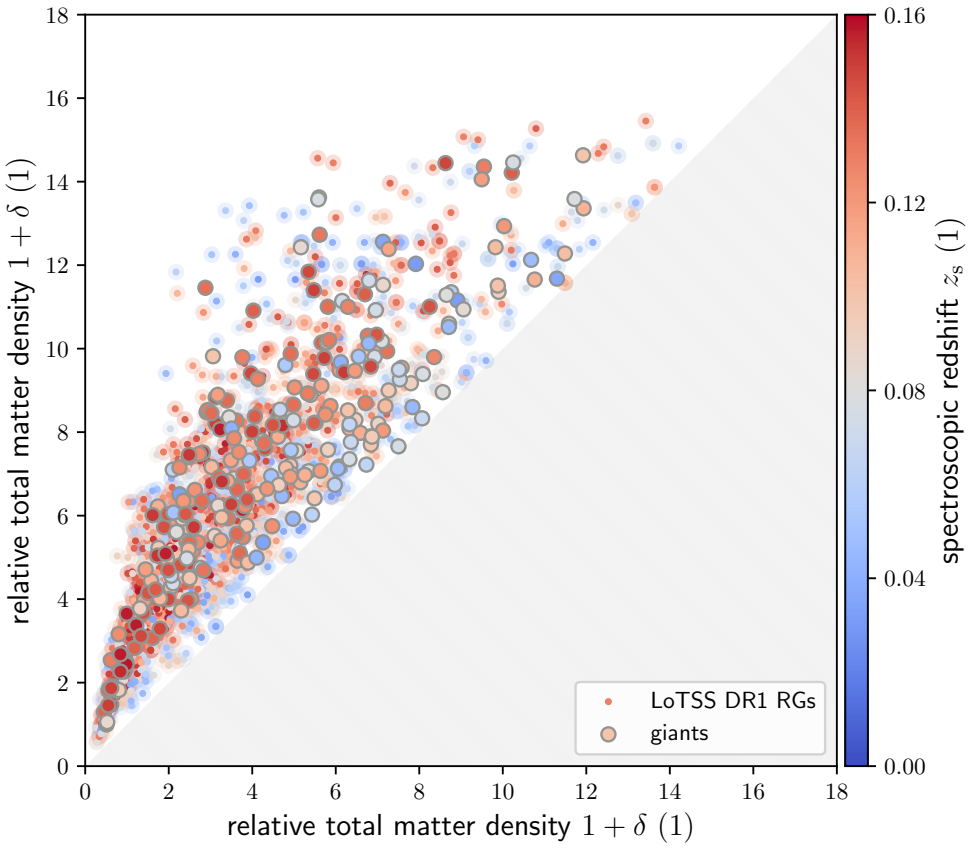


Figure 7.4: Comparison between Cosmic Web densities of giants and LoTSS DR1 RGs in the Local Universe, inferred via two variations of our BORG SDSS–based method. The horizontal axis represents densities as measured through the fixed voxel approach of Sect. 7.3.1, while the vertical axis represents densities as measured through the flexible voxel approach of Sect. 7.3.1.

7.3.2 BORG SDSS localisation in practice

As described in Sect. 7.3.1, for each giant and LoTSS DR1 RG in our samples, we measured the mean and SD of its marginal posterior density RV. Of 281 giants that lie within the part of the BORG SDSS volume where $R^0(\mathbf{r}) \geq R^0_{\text{min}}$, there are 260 with a spectroscopic redshift (93%), which were therefore suitable for Cosmic Web analysis. Of these, 208 (80%) are LoTSS DR2 discoveries (Oei et al., 2023a). In exactly the same way, of the 1870 LoTSS DR1 RGs that lie within the constrained BORG SDSS volume, we retained 1443 specimina with spectroscopic redshifts (77%), which we selected for Cosmic Web analysis.

At the BORG SDSS resolution of 2.9 Mpc h^{-1} per voxel side, the baryonic matter density field approximately equals the dark matter density field scaled down by a fac-

tor $\frac{\Omega_{\mathrm{BM},0}}{\Omega_{\mathrm{DM},0}}$. The BORG SDSS does not distinguish between these two fields. Instead, it provides the sum of the baryonic and dark matter density field: the total matter density field. In this article, we shall exclusively mention relative total matter densities — total matter densities divided by today's cosmic mean total matter density $\Omega_{\mathrm{M},0}\rho_{\mathrm{c},0}$, where $\rho_{\mathrm{c},0}$ is today's critical density. To be less verbose, we shall just refer to 'relative densities', and write $1 + \delta$, where δ denotes overdensity in the usual sense.

7.4 Results

In this section, we present the first empirical distributions of the Cosmic Web total matter density and dynamical state around giants and general RGs. We also determine the radio luminosity–Cosmic Web density relation, and explore whether it can cause the density distribution discrepancy between observed giants and general RGs.

7.4.1 Cosmic Web density distributions

The marginal relative density distributions of individual RGs resemble lognormal distributions, irrespective of whether the fixed or flexible voxel method is used to determine them. We demonstrate this in Fig. 7.5 through example marginals for a typical giant and a typical LoTSS DR1 RG.

Indeed modelling $1 + \Delta_{\mathrm{GRG,obs}} \mid 1 + \Delta_{\mathrm{GRG}} = 1 + \delta_i \sim \mathrm{Lognormal}(\mu, \sigma^2)$, we can succinctly summarise each giant's measured density distribution with two parameters. We provide these, for 50 out of 260 giants, in Table 7.1. (For access to such data for all giants, and for similar data on LoTSS DR1 giants, see its footnote.)

By aggregating just the means of these marginal relative density distributions, we can analyse the distributions for our observed populations as a whole. In green, Fig. 7.6 shows kernel density estimated (KDE) relative density distributions for both giants (top panels) and the broader population of RGs (bottom panels) in the Local Universe. The panels in the left column represent the fixed voxel method, whilst the panels in the right column represent the flexible voxel method. More precisely, these KDE distributions approximate the distributions of the observed GRG relative density RV $1 + \Delta_{\mathrm{GRG,obs}}$ and the observed RG relative density RV $1 + \Delta_{\mathrm{RG,obs}}$.

We sought to summarise these distributions parametrically. After testing various two-parameter distributions for continuous, non-negative RVs (such as the gamma distribution and the lognormal distribution), the KDE distributions of $1 + \Delta_{\mathrm{RG,obs}}$ in the bottom panels — which are based on all 1443 selected LoTSS DR1 RGs — appeared best approximated by a gamma distribution.[6] Maximum likelihood estima-

[6] In Appendix 7.A2, we provide an astrophysical–statistical argument that could explain the

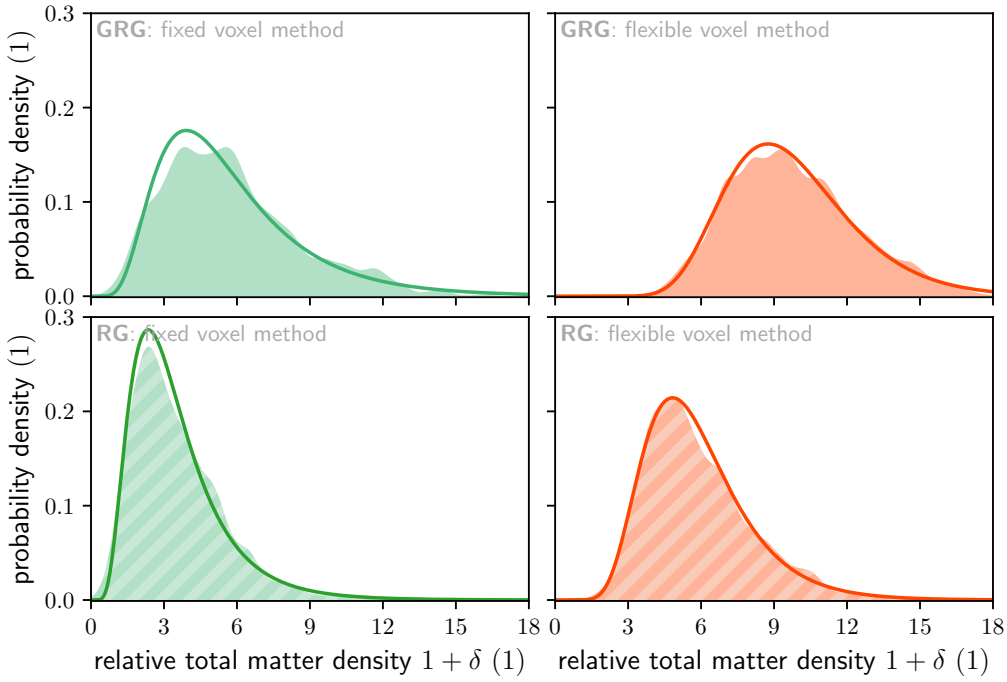


Figure 7.5: Distributions for the measured relative total matter density RVs $1+\Delta_{\mathrm{GRG,obs}} \,|\, 1+\Delta_{\mathrm{GRG}} = 1+\delta_i$ and $1+\Delta_{\mathrm{RG,obs}} \,|\, 1+\Delta_{\mathrm{RG}} = 1+\delta_i$ for an individual giant (top row; solid), and for an individual LoTSS DR1 RG (bottom row; hatched). Fixed voxel method densities (left column; green) are lower than flexible voxel method densities (right column; orange). The MLE-fitted lognormal PDFs (solid curves) demonstrate that the distributions are approximately lognormal.

tion (MLE) suggested $1+\Delta_{\mathrm{RG,obs}} \sim \Gamma(k=2.1, \theta=1.8)$ for the fixed voxel method and $1+\Delta_{\mathrm{RG,obs}} \sim \Gamma(k=3.6, \theta=1.8)$ for the flexible voxel method. Similarly, fitting a gamma distribution to $1+\Delta_{\mathrm{GRG,obs}}$ through MLE gave $1+\Delta_{\mathrm{GRG,obs}} \sim \Gamma(k=2.5, \theta=1.9)$ for the fixed voxel method and $1+\Delta_{\mathrm{GRG,obs}} \sim \Gamma(k=4.7, \theta=1.6)$ for the flexible voxel method, although the latter does not provide a tight fit. These gamma distributions constitute practical, two-parameter representations of the underlying data and are drawn as solid lines in Fig. 7.6.

The mean of a gamma-distributed RV $1+\Delta \sim \Gamma(k,\theta)$ is $\mathbb{E}[1+\Delta] = k\theta$. Thus, for the fixed voxel method, $\mathbb{E}[1+\Delta_{\mathrm{GRG,obs}}] = 4.8$ and $\mathbb{E}[1+\Delta_{\mathrm{RG,obs}}] = 3.8$. Similarly, for the flexible voxel method, $\mathbb{E}[1+\Delta_{\mathrm{GRG,obs}}] = 7.5$ and $\mathbb{E}[1+\Delta_{\mathrm{RG,obs}}] = 6.5$. Naively, it appears that we can conclude that, in a statistical sense, giants occupy denser regions of the Cosmic Web than radio galaxies in general. Using a two-sample Kolmogorov–Smirnov (KS) test, we formally tested the null hypothesis that the observed giant and general RG relative density distributions of Fig. 7.6 share a common underlying dis-

gamma distribution's emergence here.

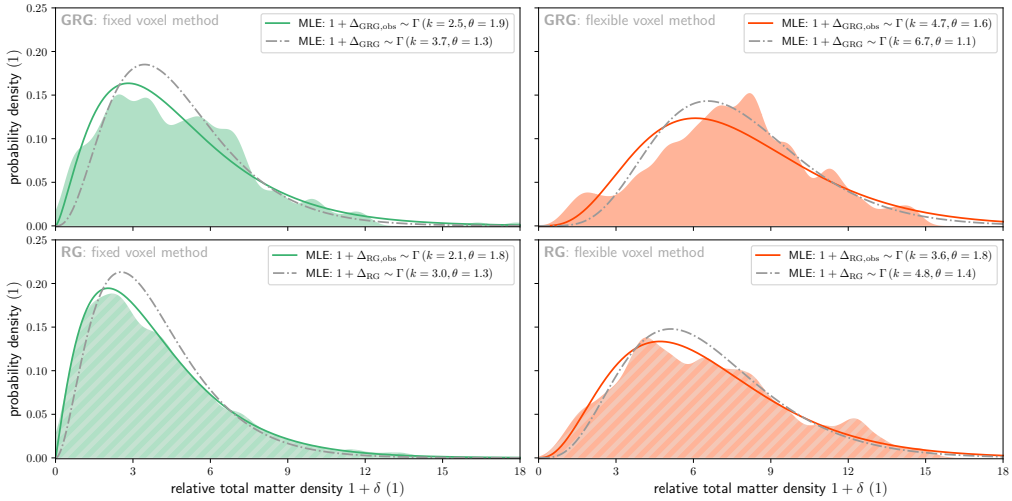


Figure 7.6: Probability density functions (PDFs) of the relative total matter density RV of 260 Local Universe giants (top row), and of 1443 Local Universe RGs (bottom row), determined through the fixed voxel method (left column) and flexible voxel method (right column). These RVs correspond to the density field smoothed to a scale of 2.9 Mpc h^{-1}. We also show PDFs (solid lines) of gamma-distributed RVs with parameters obtained via maximum likelihood estimation (MLE), alongside PDFs (dash-dotted lines) of gamma-distributed RVs with parameters obtained via MLE with a heteroskedasticity correction. We warn that these distributions are affected by surface brightness selection, and thus represent observed populations only.

tribution. For both the fixed and flexible voxel methods, the p-value $p \lesssim \sim 10^{-6}$. The null hypothesis is thus rejected (for typical significance levels).

However, these distributions come with two major caveats.

Heteroskedasticity

Firstly, BORG SDSS relative density measurements are heteroskedastic: measurements of higher densities have higher errors than measurements of lower densities (see e.g. Fig. 6 of Jasche et al., 2015). Heteroskedasticity causes the observed distributions to differ from the intrinsic distributions; not only by widening them, as also occurs in the more familiar homoskedastic setting, but also by systematically shifting the distributions towards lower densities. Therefore, in Appendix 7.A3, we propose a simple method to infer distributions corrected for this effect. As a result, we distin-

[7] We share a table with all 260 entries, alongside the analogous table for our 1443 selected LoTSS DR1 RGs, in Flexible Image Transport System (FITS) format through the Centre de Données astronomiques de Strasbourg (CDS). One can either use anonymous File Transfer Protocol (FTP) to `ftp://cdsarc.cds.unistra.fr` (130.79.128.5) or visit `https://cdsarc.cds.unistra.fr/cgi-bin/qcat?J/A+A/`.

Table 7.1: Cosmic Web properties for 50 of 260 BORG SDSS–constrained giants, sorted by right ascension. We provide MLE parameters of lognormal fits to total matter (i.e. both baryonic and dark matter) density distributions. We report parameters for both the fixed and flexible voxel methods. Densities span the BORG SDSS comoving voxel volume of $(2.9\ \mathrm{Mpc}\ h^{-1})^3$ and are relative to the Universe's mean late-time total matter density. The mean and variance of a relative density distribution are $\mathbb{E}[1+\Delta_{\mathrm{GRG,obs}} \,|\, 1+\Delta_{\mathrm{GRG}} = 1+\delta_i] = \exp\left(\mu + \frac{1}{2}\sigma^2\right)$ and $\mathbb{V}[1+\Delta_{\mathrm{GRG,obs}} \,|\, 1+\Delta_{\mathrm{GRG}} = 1+\delta_i] = (\exp(\sigma^2) - 1)\exp(2\mu + \sigma^2)$. The components of the T-web probability vector $\mathbf{p} = (p_1, p_2, p_3, p_4)$ correspond to clusters, filaments, sheets, and voids, respectively; due to rounding, their sum can appear to deviate from 1. Cluster masses are given only for giants generated by BCGs, and stem from cross-matching with Wen & Han (2015).[7]

rank ↓	host coordinates equatorial J2000 (°)	spectroscopic redshift z (1)	Cosmic Web density fixed voxel μ, σ^2 (1)	Cosmic Web density flexible voxel μ, σ^2 (1)	Cosmic Web T-web probabilities $\mathbf{p}$ (1)	cluster mass ($10^{14}\ M_\odot$)
1	111.5768, 38.6332	$0.1539 \pm 3 \cdot 10^{-5}$	−0.49, 0.60	0.60, 0.45	(0.0, 0.2, 0.7, 0.1)	-
2	113.7719, 41.9743	$0.0873 \pm 2 \cdot 10^{-5}$	1.94, 0.06	2.43, 0.03	(0.1, 0.9, 0.0, 0.0)	-
3	114.9887, 43.9833	$0.1489 \pm 2 \cdot 10^{-5}$	−0.73, 0.52	0.42, 0.42	(0.0, 0.2, 0.6, 0.2)	-
4	115.9636, 28.3578	$0.1063 \pm 2 \cdot 10^{-5}$	0.88, 0.53	1.87, 0.20	(0.1, 0.8, 0.2, 0.0)	-
5	116.0338, 43.9917	$0.1348 \pm 2 \cdot 10^{-5}$	0.95, 0.23	2.41, 0.05	(0.0, 0.9, 0.1, 0.0)	-
6	117.5396, 26.7355	$0.1304 \pm 2 \cdot 10^{-5}$	1.59, 0.17	2.16, 0.06	(0.2, 0.8, 0.0, 0.0)	-
7	118.1447, 35.8398	$0.1365 \pm 3 \cdot 10^{-5}$	0.95, 0.25	2.09, 0.10	(0.0, 0.8, 0.2, 0.0)	-
8	119.1594, 32.4631	$0.1462 \pm 2 \cdot 10^{-5}$	0.47, 0.50	1.28, 0.23	(0.1, 0.5, 0.4, 0.0)	-
9	119.2623, 36.6224	$0.1395 \pm 3 \cdot 10^{-5}$	−0.02, 0.42	0.94, 0.22	(0.0, 0.2, 0.8, 0.0)	-
10	119.2794, 35.9797	$0.1330 \pm 2 \cdot 10^{-5}$	1.15, 0.28	2.08, 0.07	(0.1, 0.8, 0.1, 0.0)	1.3
11	119.4716, 36.6728	$0.1284 \pm 2 \cdot 10^{-5}$	1.26, 0.15	1.66, 0.08	(0.1, 0.8, 0.1, 0.0)	0.8
12	120.2556, 13.8312	$0.1087 \pm 2 \cdot 10^{-5}$	2.46, 0.03	2.58, 0.02	(0.7, 0.3, 0.0, 0.0)	3.2
13	120.3054, 34.6752	$0.0827 \pm 2 \cdot 10^{-5}$	1.56, 0.08	1.93, 0.04	(0.1, 0.8, 0.0, 0.0)	-
14	120.3832, 47.6045	$0.1568 \pm 3 \cdot 10^{-5}$	−0.30, 0.56	1.14, 0.32	(0.0, 0.4, 0.6, 0.0)	-
15	120.8052, 51.9326	$0.0695 \pm 2 \cdot 10^{-5}$	2.00, 0.04	2.21, 0.03	(0.3, 0.7, 0.0, 0.0)	-
16	121.0142, 40.8026	$0.1262 \pm 1 \cdot 10^{-5}$	0.85, 0.20	1.59, 0.11	(0.1, 0.8, 0.1, 0.0)	2.5
17	121.3279, 28.6242	$0.1426 \pm 2 \cdot 10^{-5}$	1.18, 0.39	1.78, 0.17	(0.1, 0.6, 0.3, 0.0)	-
18	121.3804, 25.8032	$0.1370 \pm 2 \cdot 10^{-5}$	1.02, 0.31	1.89, 0.09	(0.1, 0.8, 0.2, 0.0)	-
19	121.4295, 16.2322	$0.1000 \pm 2 \cdot 10^{-5}$	2.46, 0.03	2.67, 0.02	(0.6, 0.4, 0.0, 0.0)	0.9
20	122.1485, 38.9145	$0.0408 \pm 1 \cdot 10^{-5}$	1.79, 0.04	2.26, 0.02	(0.1, 0.9, 0.0, 0.0)	-
21	122.2863, 29.6790	$0.1257 \pm 2 \cdot 10^{-5}$	1.18, 0.18	2.03, 0.06	(0.1, 0.8, 0.1, 0.0)	-
22	122.3128, 41.2890	$0.1334 \pm 3 \cdot 10^{-5}$	0.86, 0.38	1.48, 0.14	(0.2, 0.6, 0.2, 0.0)	-
23	123.4121, 41.3650	$0.0998 \pm 2 \cdot 10^{-5}$	1.05, 0.15	2.27, 0.04	(0.0, 0.7, 0.2, 0.0)	0.8
24	123.9160, 50.5406	$0.1380 \pm 3 \cdot 10^{-5}$	1.16, 0.29	1.85, 0.08	(0.1, 0.6, 0.2, 0.0)	-
25	124.4446, 54.7009	$0.1187 \pm 2 \cdot 10^{-5}$	2.22, 0.07	2.63, 0.03	(0.2, 0.8, 0.0, 0.0)	-
26	125.7805, 10.5983	0.0660 ± 0.0001	0.95, 0.14	1.53, 0.05	(0.2, 0.6, 0.2, 0.0)	-
27	126.4742, 41.6025	$0.1536 \pm 2 \cdot 10^{-5}$	1.38, 0.38	2.03, 0.15	(0.2, 0.7, 0.1, 0.0)	-
28	127.8646, 32.3241	$0.0512 \pm 1 \cdot 10^{-5}$	0.69, 0.12	1.78, 0.04	(0.0, 1.0, 0.0, 0.0)	-
29	127.9987, 30.6585	$0.1070 \pm 2 \cdot 10^{-5}$	1.80, 0.08	2.07, 0.04	(0.5, 0.6, 0.0, 0.0)	-
30	128.1421, 4.4100	$0.1060 \pm 1 \cdot 10^{-5}$	0.47, 0.24	1.34, 0.10	(0.0, 0.2, 0.7, 0.0)	-
31	129.0326, 26.8121	$0.0878 \pm 2 \cdot 10^{-5}$	1.22, 0.12	1.91, 0.04	(0.1, 0.7, 0.2, 0.0)	-
32	130.1852, 58.6971	$0.1439 \pm 3 \cdot 10^{-5}$	0.15, 0.40	1.21, 0.19	(0.0, 0.4, 0.6, 0.0)	1.2
33	130.5014, 38.9375	$0.1198 \pm 2 \cdot 10^{-5}$	1.82, 0.09	2.23, 0.04	(0.2, 0.8, 0.0, 0.0)	-
34	130.5398, 41.9404	$0.1257 \pm 2 \cdot 10^{-5}$	0.67, 0.31	1.93, 0.08	(0.0, 0.8, 0.2, 0.0)	-
35	131.2433, 42.0774	$0.1493 \pm 2 \cdot 10^{-5}$	1.57, 0.30	2.40, 0.08	(0.1, 0.8, 0.1, 0.0)	-
36	131.3629, 44.9240	$0.1506 \pm 3 \cdot 10^{-5}$	1.70, 0.27	2.19, 0.10	(0.3, 0.6, 0.1, 0.0)	-
37	132.7367, 42.8046	$0.0924 \pm 1 \cdot 10^{-5}$	1.05, 0.17	1.68, 0.05	(0.2, 0.7, 0.0, 0.0)	1.3
38	133.4574, 14.8739	$0.0693 \pm 2 \cdot 10^{-5}$	2.00, 0.04	2.24, 0.02	(0.8, 0.2, 0.0, 0.0)	0.7
39	133.8097, 49.1933	$0.1177 \pm 1 \cdot 10^{-5}$	0.99, 0.36	2.14, 0.09	(0.1, 0.8, 0.2, 0.0)	-
40	134.2352, 47.9559	$0.1493 \pm 2 \cdot 10^{-5}$	1.20, 0.44	2.00, 0.17	(0.1, 0.7, 0.2, 0.0)	-
41	134.5815, 46.3709	$0.1171 \pm 1 \cdot 10^{-5}$	1.19, 0.24	1.54, 0.11	(0.2, 0.6, 0.2, 0.0)	-
42	135.3674, 55.0446	0.0460	2.15, 0.03	2.35, 0.01	(0.4, 0.6, 0.0, 0.0)	-
43	135.4596, 55.9243	$0.1409 \pm 2 \cdot 10^{-5}$	1.60, 0.21	2.06, 0.09	(0.4, 0.6, 0.0, 0.0)	-
44	136.0915, 19.7245	$0.0995 \pm 2 \cdot 10^{-5}$	1.30, 0.10	1.72, 0.06	(0.1, 0.8, 0.0, 0.0)	-
45	137.1353, 18.2803	$0.1155 \pm 3 \cdot 10^{-5}$	1.17, 0.17	1.96, 0.06	(0.1, 0.8, 0.1, 0.0)	-
46	137.2374, 58.3841	$0.1436 \pm 3 \cdot 10^{-5}$	0.17, 0.52	1.33, 0.20	(0.0, 0.6, 0.4, 0.0)	-
47	137.8783, 48.5475	$0.1086 \pm 2 \cdot 10^{-5}$	0.32, 0.30	1.45, 0.09	(0.0, 0.6, 0.4, 0.0)	-
48	139.7475, 31.8613	$0.0619 \pm 1 \cdot 10^{-5}$	1.65, 0.08	1.78, 0.04	(0.6, 0.4, 0.0, 0.0)	-
49	139.9519, 57.8489	$0.1369 \pm 3 \cdot 10^{-5}$	1.21, 0.17	1.67, 0.09	(0.1, 0.8, 0.1, 0.0)	-
50	140.3421, 54.8650	$0.0447 \pm 1 \cdot 10^{-5}$	1.90, 0.02	2.31, 0.02	(0.0, 0.6, 0.4, 0.0)	-

guish between $1 + \Delta_{\mathrm{GRG,obs}}$ and $1 + \Delta_{\mathrm{GRG}}$, and between $1 + \Delta_{\mathrm{RG,obs}}$ and $1 + \Delta_{\mathrm{RG}}$ — in both cases, the latter RV has been corrected for heteroskedasticity.

As the effect appears to induce minor shifts only, we assumed that the heteroskedasticity-free distributions are also well approximated by gamma distributions. We then applied the method of Appendix 7.A3 to both our GRG and RG data. Having no closed-form MLE parameter expressions at our disposal, we obtained the MLE parameters of the heteroskedasticity-free gamma distributions by evaluating the likelihood function over an exhaustive grid of parameter values (k, θ). For the fixed voxel method, this yielded $1+\Delta_{\mathrm{GRG}} \sim \Gamma(k = 3.7, \theta = 1.3)$ and $1+\Delta_{\mathrm{RG}} \sim \Gamma(k = 3.0, \theta = 1.3)$; $\mathbb{E}[1 + \Delta_{\mathrm{GRG}}] = 4.8$ and $\mathbb{E}[1 + \Delta_{\mathrm{RG}}] = 3.9$. For the flexible voxel method, this yielded $1 + \Delta_{\mathrm{GRG}} \sim \Gamma(k = 6.7, \theta = 1.1)$ and $1 + \Delta_{\mathrm{RG}} \sim \Gamma(k = 4.8, \theta = 1.4)$; $\mathbb{E}[1 + \Delta_{\mathrm{GRG}}] = 7.4$ and $\mathbb{E}[1 + \Delta_{\mathrm{RG}}] = 6.7$. We draw the corresponding PDFs as dash-dotted lines in Fig. 7.6.

Surface brightness selection

The second — and plausibly most important — caveat is the imprint of selection effects, and in particular the surface brightness selection effect induced by the LoTSS noise level $\sigma_{I_\nu} = 25\ \mathrm{Jy\ deg^{-2}}$ at $\theta_{\mathrm{FWHM}} = 6''$. This effect causes RGs whose lobe surface brightnesses are below the noise level to evade sample inclusion. Lobe surface brightness depends on the RG's projected proper length l_{p}, morphological parameters F_l and F_{l_ν}, radio luminosity l_ν, redshift z, and lobe spectral index α. Equation 40 of Oei et al. (2023a) presents an approximate formula for the mean lobe surface brightness B_ν:

$$B_\nu = \frac{2 \cdot \mathsf{F}_{l_\nu} \cdot l_\nu}{\pi^2 \cdot \mathbb{E}[D](\eta(\mathsf{F}_l)) \cdot \mathsf{F}_l^2 \cdot l_{\mathrm{p}}^2 \cdot (1+z)^{3-\alpha}}, \tag{7.1}$$

where F_{l_ν} is the fraction of the total radio luminosity that belongs to the lobes, F_l is the fraction of the RG's axis that lies inside the lobes, $\eta = \frac{\mathsf{F}_l}{2-\mathsf{F}_l}$, and $\mathbb{E}[D]$ is the mean deprojection factor as given by Eq. A.29 of Oei et al. (2023a):

$$\mathbb{E}[D] = (1+\eta)\left(\frac{\pi}{2} - \frac{\eta}{\sqrt{1-\eta^2}} \ln\left(\frac{\sqrt{1-\eta^2}}{\eta} + \frac{1}{\eta}\right)\right). \tag{7.2}$$

To explicitly demonstrate that surface brightness selection biases our Local Universe GRG and RG samples towards short lengths and high radio luminosities, we approximated B_ν for all LoTSS DR1 RGs with $z < 0.2$, assuming $\mathsf{F}_l = \mathsf{F}_{l_\nu} = 0.3$ (Oei et al., 2022a) and $\alpha = -0.7$.

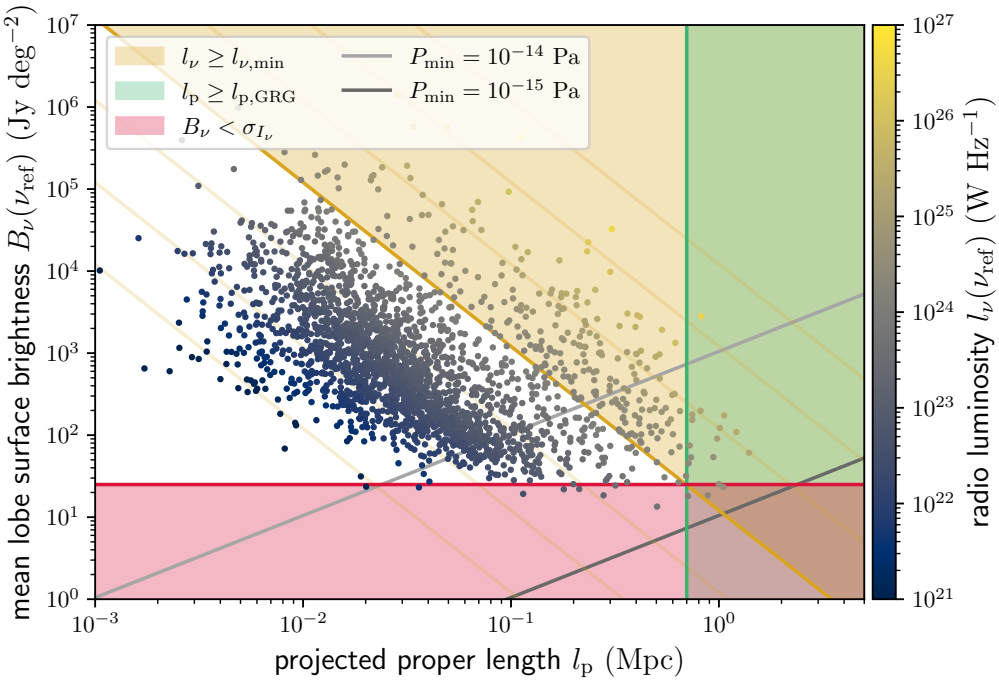


Figure 7.7: Mean lobe surface brightness estimates B_ν at $\nu_\mathrm{ref} = 150$ MHz for LoTSS DR1 RGs at $z < 0.2$, shown as a function of projected proper length l_p. During the majority ($\sim$90%) of their lifetime, RGs grow while maintaining a roughly constant radio luminosity. If — additionally — growth is shape-preserving, surface brightness decreases quadratically with l_p. The seven golden lines therefore denote approximate evolutionary tracks of RGs at $z = 0$ with end-of-life radio luminosities $l_\nu \in \{10^{21}\ \mathrm{W\,Hz^{-1}}, \ldots, 10^{27}\ \mathrm{W\,Hz^{-1}}\}$. The region of parameter space where RGs can become detectable giants, $l_\nu \geq l_{\nu,\mathrm{min}} := 10^{24}\ \mathrm{W\,Hz^{-1}}$, is shaded gold. The region of parameter space occupied by giants, $l_\mathrm{p} \geq l_\mathrm{p,GRG} := 0.7$ Mpc, is shaded green. The region of parameter space inaccessible to the LoTSS, $B_\nu < \sigma_{I_\nu} = 25$ Jy deg^{-2}, is shaded red. The grey lines indicate levels of constant minimum-energy lobe pressure P_min.

Figure 7.7 shows the results. As mature RGs grow, they trace out a curve approximately parallel to the golden lines (which represent constant radio luminosities), directed towards the bottom-right of the plot. If their lives last long enough, they therefore reach a length — which depends on their radio luminosity — beyond which they disappear into the noise of the LoTSS. In particular, our calculations suggest that the only giants that can be detected through the LoTSS are those with $l_\nu \geq 10^{24}\ \mathrm{W\,Hz^{-1}}$.[8] Indeed, the lowest–radio luminosity giant in the LoTSS DR1 GRG sample of Dabhade et al. (2020b) — which consists of 239 giants — has $l_\nu \sim$

[8] In Fig. 7.7, the approximate evolutionary track for $l_\nu = 10^{24}\ \mathrm{W\,Hz^{-1}}$ shows that when RGs with this radio luminosity turn into giants (i.e. when the fourth golden line intersects the green line), they disappear in the LoTSS noise (this golden line intersects the red line).

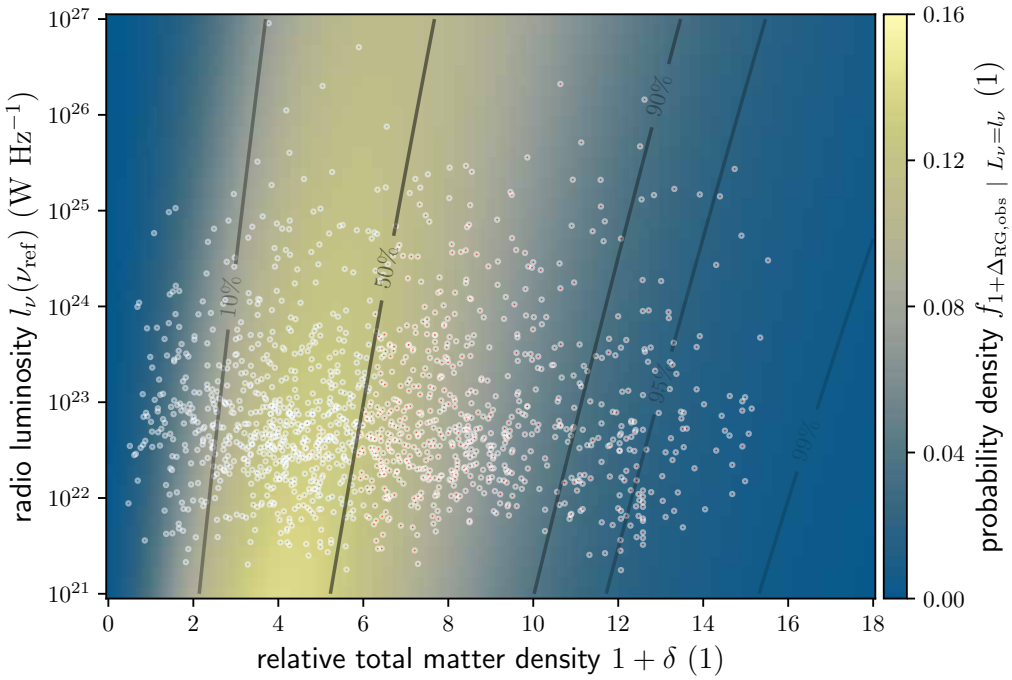


Figure 7.8: PDFs of the observed RG relative total matter density RV given a radio luminosity at $\nu_{\rm ref} = 150$ MHz, using MLE parameter values for the model described in Sect. 7.4.2. The black contours denote CDF values. We overplot all 1443 selected LoTSS DR1 RGs (dots), with those above the empirical median density coloured red, and those below coloured blue. We use flexible voxel method densities here. For fixed voxel method densities, see Fig. 7.19.

$1 \cdot 10^{24}$ W Hz^{-1}. Clearly, if RGs with $l_\nu \sim 10^{23}$ W Hz^{-1} can also form giants, then this subpopulation will be missing in LoTSS GRG samples. Equivalently, the giants that *are* included in our LoTSS GRG sample necessarily have radio luminosities $l_\nu \geq 10^{24}$ W Hz^{-1}. It is overwhelmingly likely that the sample's radio luminosities are biased high.

Concerningly, Croston et al. (2019) have shown that, for RGs at $z < 0.4$, radio luminosity and environmental richness correlate positively. This gives rise to the possibility that the trend shown in Fig. 7.6, where giants appear to reside in denser Cosmic Web environments than RGs in general, is a result of surface brightness selection: a GRG sample biased high in radio luminosity will be biased high in Cosmic Web density.

7.4.2 Radio luminosity–Cosmic Web density relation

Interestingly, our data allow us to revisit the radio luminosity–environmental richness correlation that Croston et al. (2019) have found. In Fig. 7.8, we plot all 1443

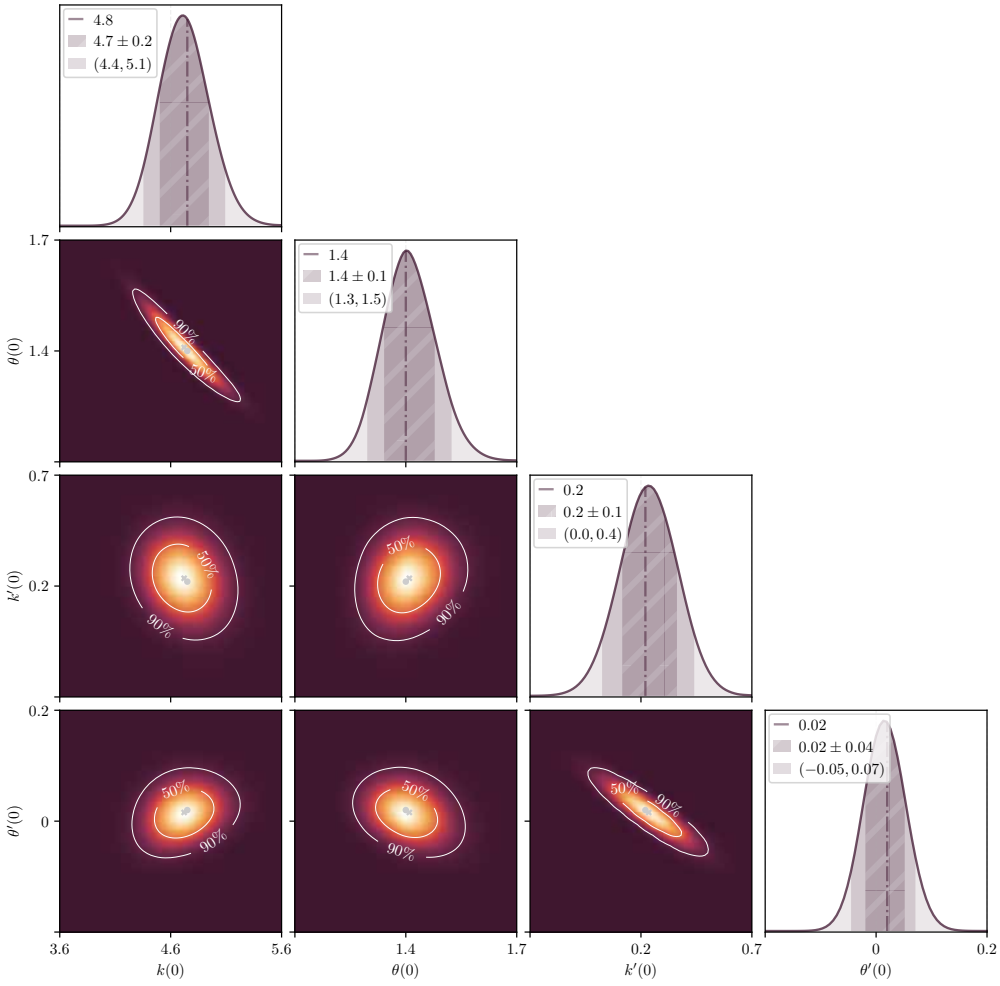


Figure 7.9: Posterior distribution over $k(0)$, $\theta(0)$, $k'(0)$, and $\theta'(0)$, based on selected Local Universe LoTSS DR1 RGs. We show all two-parameter marginals of the likelihood function, with contours enclosing 50% and 90% of total probability. We mark the maximum likelihood estimate (MLE) values (grey dot) and the maximum a posteriori (MAP) values (grey cross). The one-parameter marginals again show the MLE (dash-dotted line), a mean-centred interval of standard deviation–sized half-width (hashed region), and a median-centred 90% credible interval (shaded region). We use flexible voxel method densities here. For fixed voxel method densities, see Fig. 7.20.

selected LoTSS DR1 RGs in Cosmic Web density–radio luminosity parameter space. By eye, it appears that for higher radio luminosities, the Cosmic Web density distribution shifts upwards. To quantify the apparent tendency that more luminous radio galaxies occupy denser regions of the Cosmic Web, we extend the gamma distribution fitting discussed in Sect. 7.4.1. In particular, we assume that the RG relative density

RV at some fixed radio luminosity, $1 + \Delta_{\mathrm{RG}} \mid L_\nu = l_\nu$, is gamma distributed:

$$1 + \Delta_{\mathrm{RG}} \mid L_\nu = l_\nu \sim \Gamma(k(\lrcorner), \theta(\lrcorner)). \tag{7.3}$$

Instead of assuming constant parameters k and θ, we now assume that k and θ depend on radio luminosity:

$$\lrcorner = \lrcorner(l_\nu) := \log_{10}\left(\frac{l_\nu}{l_{\nu,\mathrm{ref}}}\right). \tag{7.4}$$

The natural minimal extension to assuming that k and θ remain constant as $\lrcorner$ changes, is to consider Taylor polynomial approximations of degree one (around $\lrcorner = 0$ — i.e. Maclaurin polynomials):

$$k(\lrcorner) \approx k(0) + k'(0) \cdot \lrcorner, \qquad \theta(\lrcorner) \approx \theta(0) + \theta'(0) \cdot \lrcorner. \tag{7.5}$$

However, we note that the current model does not allow for a direct comparison to our data, which are samples from $1 + \Delta_{\mathrm{RG,obs}} \mid L_\nu = l_\nu$ rather than of $1 + \Delta_{\mathrm{RG}} \mid L_\nu = l_\nu$. Therefore, we must extend the model by applying heteroskedastic measurement errors. In line with Appendix 7.A3, we assume

$$1 + \Delta_{\mathrm{RG,obs}} \mid 1 + \Delta_{\mathrm{RG}} = 1 + \delta \sim \mathrm{Lognormal}(\mu, \sigma^2), \tag{7.6}$$

where the parameters $\mu = \mu(a, b)$ and $\sigma^2 = \sigma^2(a, b)$ are given by

$$\mu = \ln(1 + \delta) - \frac{1}{2}\ln\left(1 + a\,(1 + \delta)^{b-2}\right); \tag{7.7}$$

$$\sigma^2 = \ln\left(1 + a\,(1 + \delta)^{b-2}\right). \tag{7.8}$$

Here, a and b are parameters that characterise the heteroskedasticity. By considering the means and SDs of the relative densities of all selected LoTSS DR1 RGs, we found $a = 0.4$ and $b = 1.1$ for the fixed voxel method and $a = 0.4$ and $b = 0.9$ for the flexible voxel method.

In order to obtain a fit that is valid over many orders of magnitude in radio luminosity, rather than only over the order of magnitude in which most data are concentrated, we calculated the likelihood in such a way that all decades in radio luminosity that contain at least 20 RGs receive equal weight. We simply determined the likelihood function, and thus the posterior distribution for a flat prior, through a grid search over the four parameters $k(0)$, $\theta(0)$, $k'(0)$, and $\theta'(0)$. We define $l_{\nu,\mathrm{ref}} := 10^{23}\ \mathrm{W\ Hz^{-1}}$. We visualise this flat-prior posterior in Fig. 7.9. In addition, we show percentile scores

of 10%, 50%, 90%, and 95% corresponding to the MAP model in Fig. 7.8. Figures 7.19 and 7.20 demonstrate that, although the densities are lower, the results remain qualitatively the same upon switching to the fixed voxel method. We summarise the posterior for both methods in Table 7.2.

Table 7.2: Maximum a posteriori probability (MAP) and posterior mean and standard deviation (SD) estimates of the free parameters in RG Cosmic Web density–radio luminosity inference. As we assume a flat prior, the MAP parameters are also the MLE parameters.

fixed voxel method:

parameter	MAP	posterior mean and SD
$k(0)$	2.9	2.9 ± 0.1
$\theta(0)$	1.4	1.4 ± 0.1
$k'(0)$	0.04	0.04 ± 0.08
$\theta'(0)$	0.08	0.08 ± 0.04

flexible voxel method:

parameter	MAP	posterior mean and SD
$k(0)$	4.8	4.7 ± 0.2
$\theta(0)$	1.4	1.4 ± 0.1
$k'(0)$	0.2	0.2 ± 0.1
$\theta'(0)$	0.02	0.02 ± 0.04

As values of $k'(0)$ and especially $\theta'(0)$ tend to be positive, we find evidence for a positive radio luminosity–environmental richness relation, in line with the results of Croston et al. (2019). More quantitatively, we call the relation positive (at $l_\nu = l_{\nu,\mathrm{ref}}$) when

$$\left.\frac{\mathrm{d}\mathbb{E}[1 + \Delta_{\mathrm{RG}} \mid L_\nu = l_\nu]}{\mathrm{d}\lrcorner}\right|_{\lrcorner=0} = k(0)\theta'(0) + k'(0)\theta(0) > 0. \tag{7.9}$$

We combined the posterior with Eq. 7.9 to obtain the posterior probability that the radio luminosity–Cosmic Web density relation is positive — that is to say that more luminous RGs tend to live in denser regions of the Cosmic Web. For both the fixed and flexible voxel methods, this probability exceeds 99%.

Surface brightness selection

However, we must again be mindful of a possible surface brightness selection effect. RGs of a given radio luminosity may be able to grow larger in more tenuous Cosmic Web environments, as radio luminosity traces jet power (in close-to-linear fashion;

e.g. Willott et al., 1999; Shabala & Godfrey, 2013; Hardcastle, 2018). If this is indeed the case, then the lobe surface brightnesses of RGs in tenuous environments are likely lower than those of RGs with the same radio luminosity in dense environments. As a result, of all RGs with the same radio luminosity, those in tenuous environments are more likely to be missing from observational samples. This selection effect, that favours higher relative densities, holds for all radio luminosities, and could only explain the trend of Fig. 7.8 if the effect becomes more severe for higher radio luminosities. On the one hand, RGs with higher radio luminosities at a given relative density are expected to grow longer. On the other hand, as is clear from Fig. 7.7, the projected proper length range over which RGs retain detectable lobe surface brightnesses increases with radio luminosity. In other words, at higher radio luminosities, larger RGs remain detectable.

Whether the net result of these counteracting effects is less or more bias towards higher relative densities as radio luminosity increases, depends on the scaling between radio luminosity and growth. Self-similar growth predicts (Kaiser & Alexander, 1997) that an RG's projected proper length scales with jet power $Q_{\rm jet}$, Cosmic Web density at the host galaxy $1+\delta_{\rm g}$, and time since birth t as

$$l_{\rm p} \propto Q_{\rm jet}^{\frac{1}{5-\beta}} \cdot (1+\delta_{\rm g})^{\frac{-1}{5-\beta}} \cdot t^{\frac{3}{5-\beta}}, \tag{7.10}$$

where $-\beta$ is the exponent of the local Cosmic Web density profile:

$$(1+\delta)(r) \propto (1+\delta_{\rm g}) \cdot r^{-\beta}. \tag{7.11}$$

Under the additional assumption of equipartition between the relativistic electron–positron kinetic energy density and the magnetic field energy density, and upon neglecting electron–positron energy losses (such as adiabatic, synchrotron, and inverse Compton losses), the RG's radio luminosity obeys (Shabala & Godfrey, 2013)

$$l_\nu \propto Q_{\rm jet}^{\frac{5+p}{6}} l_{\rm p}^{3-\frac{(4+\beta)(5+p)}{12}}. \tag{7.12}$$

Here, $-p$ is the exponent of the initial post-acceleration electron–positron kinetic energy distribution $n(\gamma) \propto \gamma^{-p}$, valid for Lorentz factors between some minimum and maximum values $\gamma_{\rm min}$ and $\gamma_{\rm max}$. Solving for $Q_{\rm jet}$, substituting the result in Eq. 7.10, and solving for $l_{\rm p}$, yields — at fixed $1+\delta_{\rm g}$ and t —

$$l_{\rm p} \propto l_\nu^{\frac{12}{36+(5+p)(6-3\beta)}}. \tag{7.13}$$

In a constant-density environment, $\beta = 0$, and for the strong-shock spectral index $p = 2$, we find $l_\mathrm{p} \propto l_\nu^{2/13}$. In an environment where density falls off quadratically with distance, $\beta = 2$, and therefore $l_\mathrm{p} \propto l_\nu^{1/3}$.

Meanwhile, Eq. 7.1 predicts that the maximum detectable projected proper length for RGs with a given morphology, at a given redshift, and with given lobe spectral index, scales as $l_\mathrm{p,max} \propto \sqrt{l_\nu}$.

These arguments suggest that $l_\mathrm{p,max}$ increases more rapidly with l_ν than l_p does. Thus, amongst a family of RGs with varying radio luminosity but seen at the same age (i.e. time after birth) and with the same morphology and Cosmic Web density, the most luminous is most easily detectable. This provides some reassurance that the trend seen in Fig. 7.8 is real: surface brightness selection at a fixed relative density seems to become less severe as radio luminosity increases.

However, there are at least two caveats here. Firstly, more luminous RGs may live longer than less luminous RGs, as jointly suggested by the radio luminosity–stellar mass relation and the stellar mass–radio AGN occurrence relation of Sabater et al. (2019). As a result, in a snapshot of the RG population at a given instant, more luminous RGs may typically be older than less luminous RGs. We illustrate this point in Fig. 7.10.[9] If more luminous RGs indeed tend to be older than less luminous RGs, then the relation between projected length and radio luminosity at a fixed Cosmic Web density for a snapshot population will be steeper than the one of Eq. 7.13, which

[9] We define

$$\mathsf{F}_\mathrm{a} := \frac{\mathbb{E}[t_\mathrm{a}]}{\mathbb{E}[t_\mathrm{a} + t_\mathrm{q}]} = \frac{\mathbb{E}[t_\mathrm{a}]}{\mathbb{E}[t_\mathrm{a}] + \mathbb{E}[t_\mathrm{q}]} = \frac{\lambda_\mathrm{a}^{-1}}{\lambda_\mathrm{a}^{-1} + \lambda_\mathrm{q}^{-1}} = \left(1 + \frac{\lambda_\mathrm{a}}{\lambda_\mathrm{q}}\right)^{-1}. \tag{7.14}$$

We note that the more natural definition

$$\mathsf{F}_\mathrm{a} := \mathbb{E}\left[\frac{t_\mathrm{a}}{t_\mathrm{a} + t_\mathrm{q}}\right] = \mathbb{E}\left[\left(1 + \frac{t_\mathrm{q}}{t_\mathrm{a}}\right)^{-1}\right] \tag{7.15}$$

gives, for $r := \frac{\lambda_\mathrm{a}}{\lambda_\mathrm{q}}$,

$$\mathsf{F}_\mathrm{a} = \begin{cases} \frac{1}{2} & \text{if } r = 1; \\ \frac{r \ln r - r + 1}{(r-1)^2} & \text{if } r > 0, r \neq 1. \end{cases} \tag{7.16}$$

To arrive at Eq. 7.16, we use standard identities for the PDFs of the ratio distribution and the inverse distribution, and apply l'Hôpital's rule twice. When $r = 1$, that is to say when $\lambda_\mathrm{a} = \lambda_\mathrm{q}$, both definitions predict the same $\mathsf{F}_\mathrm{a} = \frac{1}{2}$. For $r < 1$, our definition predicts higher F_a than the natural definition; for $r > 1$, our definition predicts lower F_a. To find λ_a given λ_q and F_a under the natural definition, one must solve Eq. 7.16 numerically. As this procedure is less transparent, we generate Fig. 7.10 using the simpler definition for F_a.

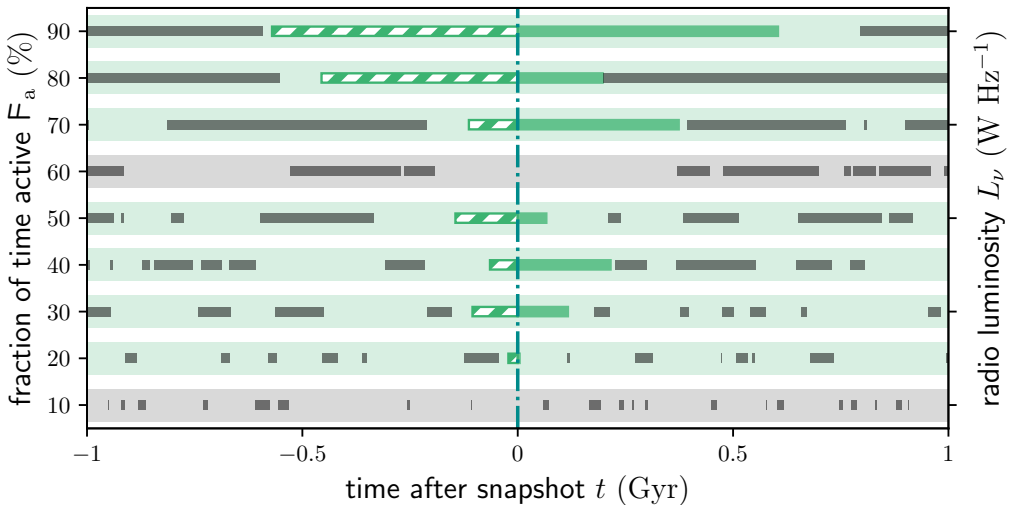


Figure 7.10: Monte Carlo simulation of a population of nine galaxies (thick horizontal bars) with successive active and quiescent phases. We mark the active phases (thin horizontal bars). Each galaxy spends a different fraction of the time F_a in the active phase. We model the time spent in both phases as draws from exponential distributions: $t_a \sim \mathrm{Exp}(\lambda_a)$ and $t_q \sim \mathrm{Exp}(\lambda_q)$. We assume that all galaxies share the same quiescent phase rate parameter $\lambda_q = 10\ \mathrm{Gyr}^{-1}$, although this may not be entirely correct (Turner & Shabala, 2015). This assumption then fixes the active phase rate parameter: $\lambda_a = (F_a^{-1} - 1)\lambda_q$. On human time scales, we can access only a snapshot view of this population. During the snapshot shown here (vertical dash-dotted line), seven galaxies generate RGs, and could be included in an RG sample (thick light-green horizontal bars). RG ages at the snapshot, represented by the hatched parts of the thin green horizontal bars, generally increase with F_a.

assumes equal ages. Secondly, we warn for the possibility — which remains subject of debate (Mingo et al., 2019) — that the distribution of morphologies at a fixed Cosmic Web density may change with radio luminosity.

Predicting the density distribution of observed giants by treating them as luminous general RGs

By applying standard PDF identities to the current context, we found

$$f_{1+\Delta_{\mathrm{RG}} \mid L_\nu \geq l_{\nu,\mathrm{min}}}(1+\delta) = \frac{\int_{l_{\nu,\mathrm{min}}}^{\infty} f_{1+\Delta_{\mathrm{RG}} \mid L_\nu = l_\nu}(1+\delta) \cdot f_{L_\nu}(l_\nu)\,\mathrm{d}l_\nu}{\int_{l_{\nu,\mathrm{min}}}^{\infty} f_{L_\nu}(l_\nu)\,\mathrm{d}l_\nu} \frac{\int_{l_{\nu,\mathrm{min}}}^{\infty} f_{1+\Delta_{\mathrm{RG}} \mid L_\nu = l_\nu}(1+\delta) \cdot f_{L_\nu \mid L_\nu \geq l_{\nu,\mathrm{min}}}(l_\nu)\,\mathrm{d}l_\nu}{\int_{l_{\nu,\mathrm{min}}}^{\infty} f_{L_\nu \mid L_\nu \geq l_{\nu,\mathrm{min}}}(l_\nu)\,\mathrm{d}l_\nu}, \tag{7.17}$$

which one can use to calculate the RG relative total matter density RV for radio luminosities of $l_{\nu,\mathrm{min}}$ and higher. This allowed us to test whether Fig. 7.6's higher densities for observed giants than for general RGs can be fully explained by the radio luminosity–Cosmic Web density relation. Clearly, to proceed, we had to find $f_{L_\nu}(l_\nu)$ for $l_\nu \geq l_{\nu,\mathrm{min}}$ — up to a constant, at least. We modelled $L_\nu \mid L_\nu \geq l_{\nu,\mathrm{min}} \sim \mathrm{Pareto}(l_{\nu,\mathrm{min}}, \lambda)$, where λ is the tail index or rate parameter of the associated exponential distribution.[10] For both $l_{\nu,\mathrm{min}} = 10^{23}\ \mathrm{W\,Hz^{-1}}$ and $l_{\nu,\mathrm{min}} = 10^{24}\ \mathrm{W\,Hz^{-1}}$, we found $\lambda_{\mathrm{MLE}} = 0.6$ (based on $1 \cdot 10^3$ and $3 \cdot 10^2$ RGs, respectively). For higher $l_{\nu,\mathrm{min}}$, λ_{MLE} increases somewhat, but becomes less reliable because of decreasing sample sizes. We therefore stuck with $\lambda_{\mathrm{MLE}} = 0.6$. We then evaluated Eq. 7.17 for $l_{\nu,\mathrm{min}} = 10^{25}\ \mathrm{W\,Hz^{-1}}$ and added heteroskedastistic noise in order to compare our prediction to observations. We propagated the correlated uncertainties on the parameters $k(0)$, $\theta(0)$, $k'(0)$, and $\theta'(0)$, which together determine $f_{1+\Delta_{\mathrm{RG}} \mid L_\nu = l_\nu}$, by rejection sampling from the joint posterior distribution. We show the resulting mean prediction with standard deviation half-width (pink) in Fig. 7.11. Although the prediction is based solely on observations of general RGs — and so is built without knowledge of the observed GRG distribution (solid green KDE) — it appears to fit it better than a direct MLE gamma fit to the GRG data (solid green curve) does. However, a *better* fit is not necessarily also a *good* fit. Is our sample of Cosmic Web densities for observed giants consistent with the prediction? A Kolmogorov–Smirnov tests suggests an answer in the affirmative, as it yielded a p-value of 0.7.

Section 7.4.1 presented a paradox: known giant radio galaxies reside in denser Cosmic Web environments than their smaller kin, in apparent violation of the long-standing hypothesis that giants should emerge predominantly in the dilute Cosmic Web. The modelling in this section suggests that the higher densities found for known giants than for general RGs can be explained by a combination of these giants' higher radio luminosities and the general relation between RG radio luminosity and Cosmic Web density. Fainter giants — which might inhabit the more dilute Cosmic Web — might exist, but are currently not observationally accessible.

[10] Equivalently, one could model

$$\ln\left(\frac{L_\nu \mid L_\nu \geq l_{\nu,\mathrm{min}}}{l_{\nu,\mathrm{min}}}\right) \sim \mathrm{Exp}(\lambda), \tag{7.18}$$

or

$$\log_{10}\left(\frac{L_\nu \mid L_\nu \geq l_{\nu,\mathrm{min}}}{l_{\nu,\mathrm{min}}}\right) \sim \mathrm{Exp}\left(\lambda \cdot \ln 10\right). \tag{7.19}$$

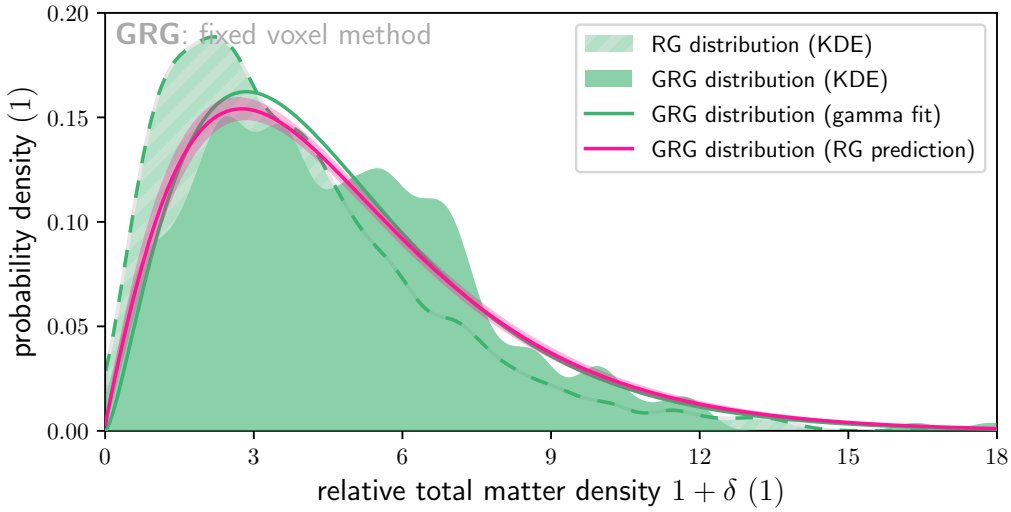


Figure 7.11: Density distribution predicted for observed giants combining Sect. 7.4.2's RG radio luminosity–Cosmic Web density relation and a radio luminosity cut-off $l_{\nu,\mathrm{min}} = 10^{25}$ W Hz^{-1} (pink); in addition, we show the KDE density distributions for giants with gamma distribution MLE fit (solid green) and for general RGs (hatched green) — all as in the left column of Fig. 7.6.

7.4.3 Radio galaxy number densities

After accreting baryons, supermassive black holes (SMBHs) can launch jets that give rise to radio galaxies. Although it is clear that SMBH jets and radio galaxies play an important role in galaxy evolution and cosmology, the physics of jet launching is not yet fully understood. Moreover, because SMBHs are astronomical unit–sized, galaxies are kiloparsec-sized (ratio 10^8–10^9), and the Cosmic Web is megaparsec-sized (ratio 10^{11}–10^{12}), it is not yet possible to build cosmological simulations in which a realistic interplay between SMBHs, their host galaxies, and the enveloping Cosmic Web naturally arises. Thus, for now, major advances in our understanding of radio galaxies on the cosmological scale must come from observations instead of from simulations. It is therefore of considerable interest to observationally constrain the occurrence of radio galaxies as a function of Cosmic Web density.

Enticingly, our relative total matter density measurements — in combination with the BORG SDSS HMC Markov chain samples in their entirety — allow us to determine the RG number density as a function of Cosmic Web density: $n_{\mathrm{RG}}(1+\delta)$. To see why, we first emphasise that the PDF $f_{1+\Delta_{\mathrm{RG}}}$, which we have approximated through observations and visualised in the bottom row of Fig. 7.6, provides a key ingredient for $n_{\mathrm{RG}}(1+\delta)$. It shows that if a volume of cosmological extent is surveyed, one will obtain more RGs with moderate (filament-like) densities than with low (void-like) or high (cluster-like) densities. We remark that in such a survey sample the number of

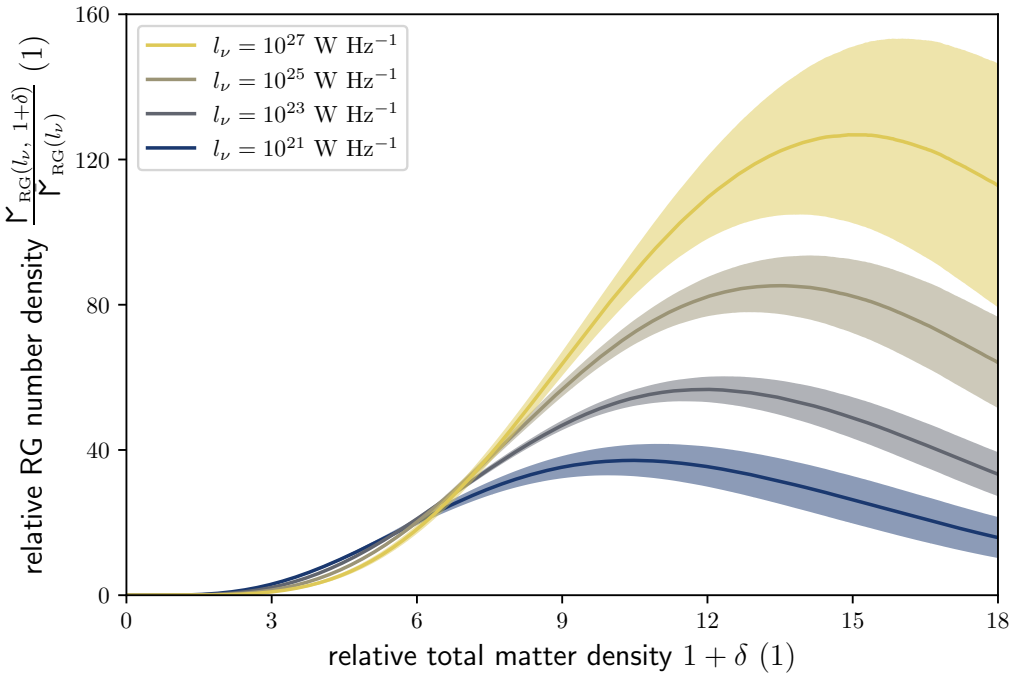


Figure 7.12: Number densities of RGs with a given radio luminosity l_ν, as a function of Cosmic Web density at the 2.9 Mpc h^{-1} scale. The number density functions reflect conditions in the Local Universe, and are given relative to the cosmic mean number density at l_ν. The solid curves reflect the posterior mean, whilst the shaded areas around the mean denote -1 to $+1$ posterior standard deviation ranges. We use flexible voxel method densities here.

RGs whose densities fall within a given interval depends not only on the RG number density at the interval's densities, but also on the prevalence of environments with such densities. As a result, if a sample contains more filament-inhabiting RGs than, say, cluster-inhabiting RGs, one need not necessarily conclude that the RG number density in filaments is higher than in clusters; in fact, the RG number density in the latter environment could well be higher, provided that clusters are sufficiently rare.

Conveniently, the density fields of the BORG SDSS HMC Markov chain samples enable us to quantify the rarity of each Cosmic Web environment with a PDF $f_{1+\Delta_{\mathrm{CW}}}$, which corresponds to the Cosmic Web relative density RV $1 + \Delta_{\mathrm{CW}}$. This RV represents the relative total matter density at a randomly chosen point in the Local Universe (and at the BORG SDSS resolution of 2.9 Mpc h^{-1}). We determine the distribution of $1 + \Delta_{\mathrm{CW}}$ simply by binning the density fields of a few hundred Markov chain samples. As a result, $f_{1+\Delta_{\mathrm{CW}}}$ is best determined at low densities, which correspond to the most common environments, and least so at high densities, which

are rare.[11] As we expect $f_{1+\Delta_{\rm CW}}$ to be smooth, we apply a Savitzky–Golay filter (of polynomial order 1) to the binned data; however, this step is not essential.

Appendix 7.A5 shows that $n_{\rm RG}(1+\delta)$, the RG number density as would arise in an environment of constant density $1+\delta$, obeys

$$\frac{n_{\rm RG}(1+\delta)}{\bar{n}_{\rm RG}} = \frac{f_{1+\Delta_{\rm RG}}(1+\delta)}{f_{1+\Delta_{\rm CW}}(1+\delta)}. \tag{7.20}$$

Here, $\bar{n}_{\rm RG}$ is the cosmic mean RG number density. Thus, a simple point-wise division of the PDFs of $1+\Delta_{\rm RG}$ and $1+\Delta_{\rm CW}$ yields the RG number density at a given Cosmic Web density relative to the cosmic mean value.

Unfortunately, as discussed in Sect. 7.4.1, our measurements of $f_{1+\Delta_{\rm RG}}$ suffer from surface brightness selection. To avoid selection effects, we must formulate a more specific version of Eq. 7.20. In particular, we consider the number density of RGs per unit of radio luminosity as a function of radio luminosity and relative density: $\Upsilon_{\rm RG}(l_\nu, 1+\delta)$ — again, relative to a cosmic mean value: $\bar{\Upsilon}_{\rm RG}(l_\nu)$. Appendix 7.A5 shows that this ratio equals

$$\frac{\Upsilon_{\rm RG}(l_\nu, 1+\delta)}{\bar{\Upsilon}_{\rm RG}(l_\nu)} = \frac{f_{1+\Delta_{\rm RG}|L_\nu=l_\nu}(1+\delta)}{f_{1+\Delta_{\rm CW}}(1+\delta)}. \tag{7.21}$$

This radio luminosity–dependent relative number density is, as before, equal to a point-wise division of PDFs. With respect to the RHS of Eq. 7.20, only the PDF in the numerator has changed. We have modelled this PDF, $f_{1+\Delta_{\rm RG}|L_\nu=l_\nu}$, in Sect. 7.4.2. It depends on the four parameters $k(0)$, $k'(0)$, $\theta(0)$, and $\theta'(0)$, for which the data shown in Fig. 7.8 induce the posterior distribution shown in Fig. 7.9. We evaluate the RHS of Eq. 7.21 by rejection sampling from the posterior, calculating $f_{1+\Delta_{\rm RG}|L_\nu=l_\nu}$ for each sampled quartet $(k(0), k'(0), \theta(0), \theta'(0))$, and dividing the resulting PDFs point-wise by $f_{1+\Delta_{\rm CW}}$. In Fig. 7.12, for four observationally accessible radio luminosities, we visualise the posterior mean alongside a range that is two standard deviations wide.

[11] We note that, at first sight, it might be tempting to approximate $f_{1+\Delta_{\rm CW}}$ by binning the BORG SDSS mean rather than individual samples. This, however, leads to incorrect results. The BORG SDSS mean tends to the cosmic mean wherever the structure of the Cosmic Web is not well constrained by the SDSS DR7 MGS. This tendency to revert to the cosmic mean in the absence of data constraints causes low-density environments to be overrepresented and high-density environments to be underrepresented. As a result, approximating $f_{1+\Delta_{\rm CW}}$ by binning the BORG SDSS mean would e.g. underestimate its high-density tail. More quantitatively, when using the BORG SDSS mean, the probability densities $f_{1+\Delta_{\rm CW}}(1+\delta)$ for $1+\delta \gtrsim 10$ are up to an order of magnitude lower than predicted by BORG SDSS samples.

As expected, RG number densities generally increase with Cosmic Web density. The more luminous RGs become, the more pronounced is the contrast between their number density in clusters and their cosmic mean number density. Our results indicate that high-luminosity RGs, with $l_\nu = 10^{25}$ W Hz^{-1} and $l_\nu = 10^{27}$ W Hz^{-1}, attain number densities in clusters that are a hundred times higher than average. Interestingly, for radio luminosities $l_\nu = 10^{21}$ W Hz^{-1} and $l_\nu = 10^{23}$ W Hz^{-1}, we obtain statistically significant, preliminary evidence that the RG number density peaks at a cluster-like density that depends positively on l_ν.[12] Beyond this density, these low-luminosity RGs become less prevalent per unit of volume, possibly because the underlying galaxy population starts generating more luminous RGs instead.

We stress that these results are tentative only, as both $f_{1+\Delta_{\rm RG}|L_\nu=l_\nu}$ and $f_{1+\Delta_{\rm CW}}$ suffer from systematic errors at higher densities. These systematic errors are not reflected in the uncertainties of Fig. 7.12. In the case of $f_{1+\Delta_{\rm RG}|L_\nu=l_\nu}$, the gamma distribution approximation loses validity in the tail: beyond some point, it will consistently under- or overestimate the probability density. Our LoTSS DR1 sample contains only a modest number ($\sim 10^2$) of RGs at relative densities $1+\delta \gtrsim 10$ — especially when selecting those around a given value of l_ν only (in which case just $\sim 10^1$ RGs remain). As a result, in its tail the gamma PDF takes on the character of an extrapolation function rather than of a fitting function, and should consequently be treated with caution. In the future, larger RG samples can resolve this situation. In the case of $f_{1+\Delta_{\rm CW}}$, the 2LPT model of structure formation used by the BORG SDSS loses validity at cluster densities. As a result, this work's estimate of $f_{1+\Delta_{\rm CW}}$ underpredicts the prevalence of cluster-like densities. Newer BORG runs, such as the BORG 2M++, employ an enhanced gravity solver (Jasche & Lavaux, 2019), relieving this limitation.

Although preliminary in nature, the inferences shown in Fig. 7.12 could in principle be used to calibrate cosmological simulations that feature RG feedback. The fact that our number densities are given with respect to a cosmic mean value is not problematic, because such a cosmic mean number density is straightforward to calculate from simulation snapshots. However, apart from making sure that the RG number densities in simulated environments of various densities are correctly proportioned, it is also important to calibrate the *absolute* level of RG feedback in simulations. Interestingly, there are variations on Eq. 7.20 for which an absolute number density can be explicitly calculated with current-day data. Consider, for example, non-giant radio galaxies (i.e. $l_{\rm p} < l_{\rm p,GRG} := 0.7$ Mpc) in the Local Universe with radio luminosities $l_\nu \geq 10^{24}$ W Hz^{-1}. Figure 7.7 suggests that this RG subpopulation is already fully accessible by the LoTSS. Calling these RGs luminous non-giant radio galaxies

[12] For radio luminosities $l_\nu = 10^{25}$ W Hz^{-1} and $l_\nu = 10^{27}$ W Hz^{-1}, the uncertainty margins are too large to claim evidence for a peak.

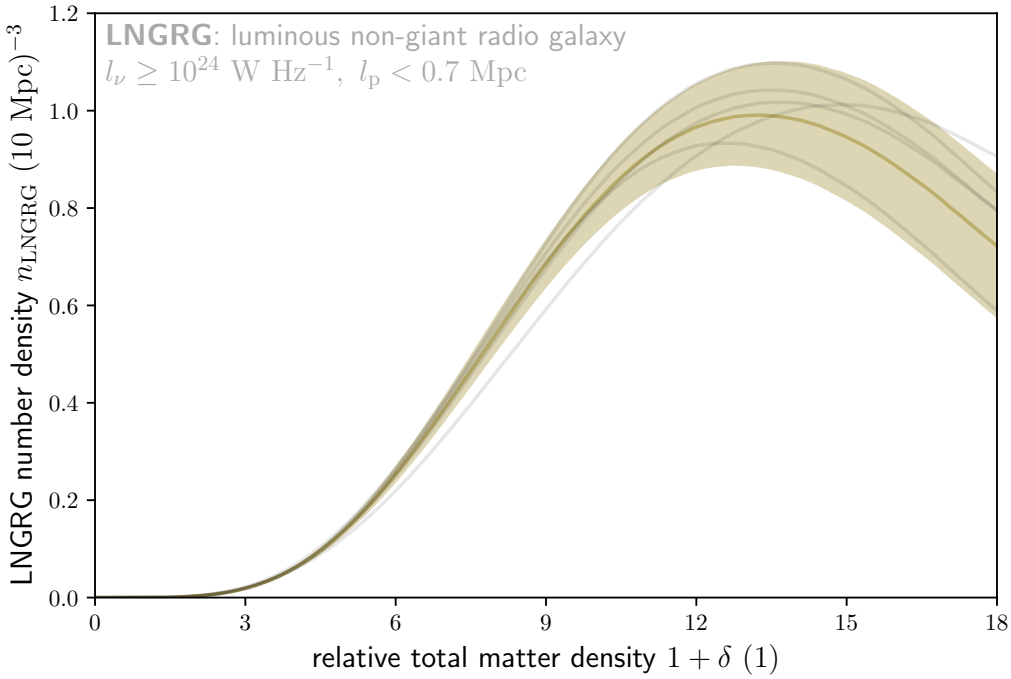


Figure 7.13: Number density of luminous non-giant radio galaxies in the Local Universe, as a function of Cosmic Web density. We define these as RGs with 150 MHz radio luminosities $l_\nu \geq 10^{24}$ W Hz^{-1}, projected proper lengths $l_\mathrm{p} < l_\mathrm{p,GRG} := 0.7$ Mpc, and redshifts $z < z_\mathrm{max} := 0.16$. The Cosmic Web densities encompass baryonic and dark matter, are defined on a 2.9 Mpc h^{-1} scale, and are given relative to the cosmic mean. We use flexible voxel method densities here. We show the posterior mean with a shaded -1 to +1 posterior standard deviation range. We also show five individual realisations (grey).

(LNGRGs), we have

$$n_\mathrm{LNGRG}(1+\delta) = \bar{n}_\mathrm{LNGRG} \cdot \frac{f_{1+\Delta_\mathrm{LNGRG}}(1+\delta)}{f_{1+\Delta_\mathrm{CW}}(1+\delta)}. \tag{7.22}$$

We calculate $\bar{n}_\mathrm{LNGRG}$ by counting all LNGRGs in the LoTSS DR1 footprint up to z_max, and find $N_\mathrm{LNGRG} = 172$. Because of calibration problems and higher noise levels towards the edges of the footprint, a fraction of the footprint is — artificially — devoid of LNGRGs. We thus apply a correction factor to N_LNGRG of 1.2, yielding $N_\mathrm{LNGRG} = 200$. Assuming that $N_\mathrm{LNGRG} \sim \mathrm{Poisson}(\lambda_\mathrm{LNGRG})$, maximum likelihood estimation simply suggests $\hat{\lambda}_\mathrm{LNGRG,MLE} = 200$. We combine this MLE Poisson distribution for N_LNGRG with the comoving volume $V = 16 \cdot 10^6$ Mpc^3 to obtain a probability distribution for $\bar{n}_\mathrm{LNGRG}$. We find $\bar{n}_\mathrm{LNGRG} = 12 \pm 1$ (100 Mpc)$^{-3}$. Next, we approximate $f_{1+\Delta_\mathrm{LNGRG}} \approx f_{1+\Delta_\mathrm{RG}\,|\,L_\nu \geq 10^{24}\,\mathrm{W\,Hz^{-1}}}$, and calculate the latter through

Eq. 7.17 and Sect. 7.4.2's subsequent procedure. Using the same $f_{1+\Delta_{\mathrm{CW}}}$ as before, we finally obtain a probability distribution over LNGRG number density functions, which we visualise in Fig. 7.13.

As it should, the LNGRG number density function follows the same trend as the functions of Fig. 7.12; the novelty here is the prediction of absolute number densities. We reiterate that systematic errors at cluster-like densities render only the left side of the plot reliable. It predicts that, in filaments, $n_{\mathrm{LNGRG}} \sim 10^{-1}\ (10\ \mathrm{Mpc})^{-3}$, or $n_{\mathrm{LNGRG}} \sim 10^{-4}\ \mathrm{Mpc}^{-3}$. As major filaments have volumes $V \sim 10^2\ \mathrm{Mpc}^3$, only one in ten major filaments harbours an LNGRG.

7.4.4 Cosmic Web dynamical state distributions

For each GRG and RG in the two samples, we obtain a probability distribution over T-web environment classes from the BORG SDSS extensions of Leclercq et al. (2015). The T-web scheme classifies the Cosmic Web at each point on the basis of its local gravitational dynamics. On an intuitive level, the scheme counts the number of dimensions along which the orbit of a test particle released at rest with respect to the environment is stable. More formally, at any location the Hessian of the gravitational potential, or tidal tensor T, has either three, two, one, or zero positive eigenvalues; the environment is interpreted correspondingly as a cluster, filament, sheet, or void.

Leclercq et al. (2015) have determined the T-web environment for each voxel of each BORG SDSS sample. By iterating over samples and counting the frequencies of the four environment classes on a per-voxel basis, one obtains a marginal environment class distribution $\mathbf{p} = (p_1, p_2, p_3, p_4)$ for each voxel. In Fig. 7.14, we show how each of $\mathbf{p}$'s components is distributed for Local Universe giants. We note the following.

- For 99% of giants, the most likely environment class is more than 50% probable. For 20% of giants, it is more likely than not that they inhabit a cluster (i.e. $p_1 > 50\%$); for 68% of giants, it is more likely than not that they inhabit a filament (i.e. $p_2 > 50\%$); and for 11% of giants, it is more likely than not that they inhabit a sheet (i.e. $p_3 > 50\%$). Finally, for 0% of giants, it is more likely than not that they inhabit a void (i.e. $p_4 > 50\%$).

- Some giants inhabit sheets (e.g. those in the bottom row of Fig. 7.3), but concluding this in individual cases rarely has a high degree of certainty: for just 0.5% of giants, it is very likely that they inhabit a sheet (i.e. $p_3 > 80\%$).

- A cluster environment is very unlikely (i.e. $p_1 < 20\%$) for a clear majority of 63% of giants, whilst a filament environment is very unlikely (i.e. $p_2 < 20\%$) for only 9% of giants; the filament environment hypothesis can thus seldom

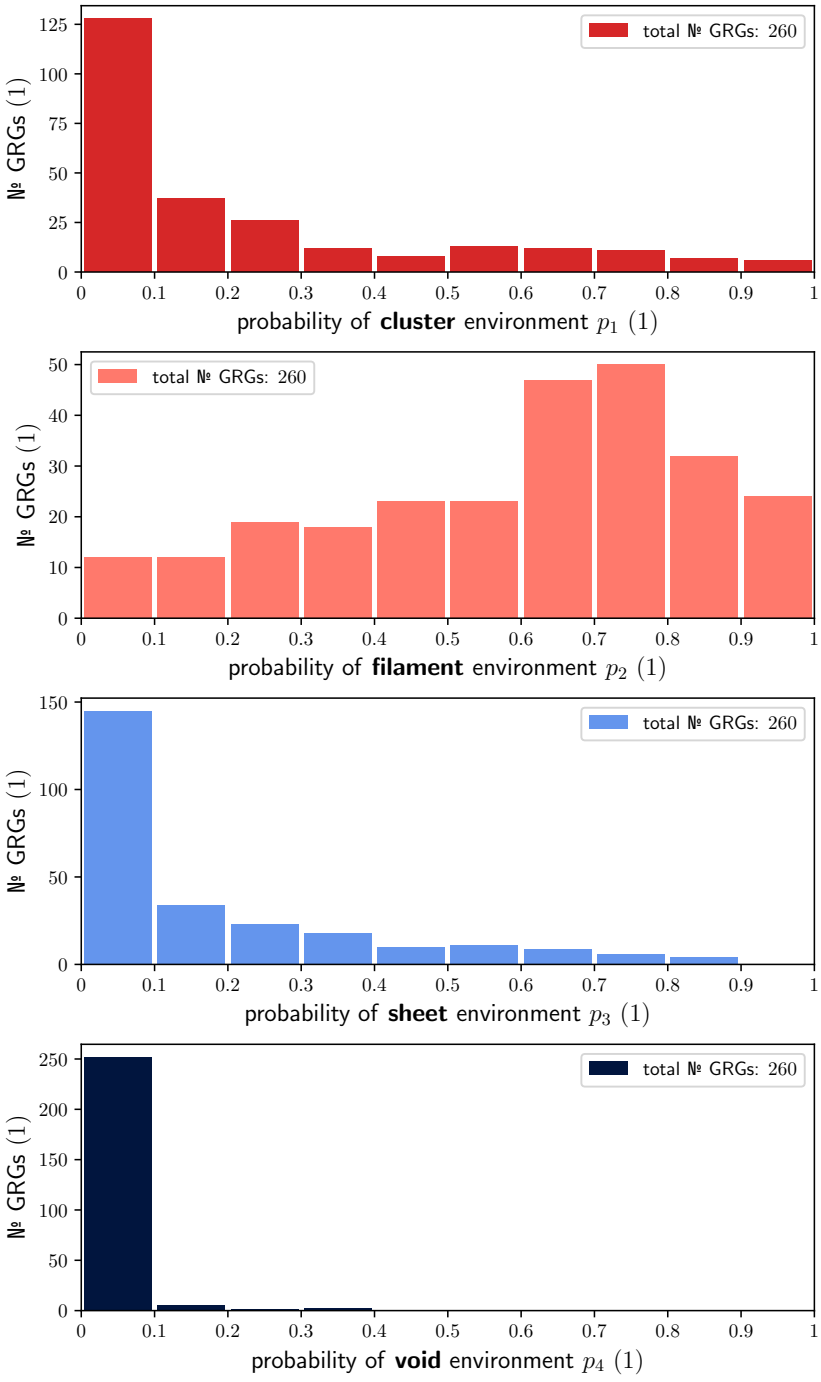


Figure 7.14: Bayesian Cosmic Web reconstructions allow one to determine a probability distribution over the four T-web environment classes — clusters, filaments, sheets, and voids — for each GRG. Here we show a statistical summary of these probability distributions. From top to bottom, we show distributions for the components of vector $\mathbf{p} = (p_1, p_2, p_3, p_4)$, which are probability parameters.

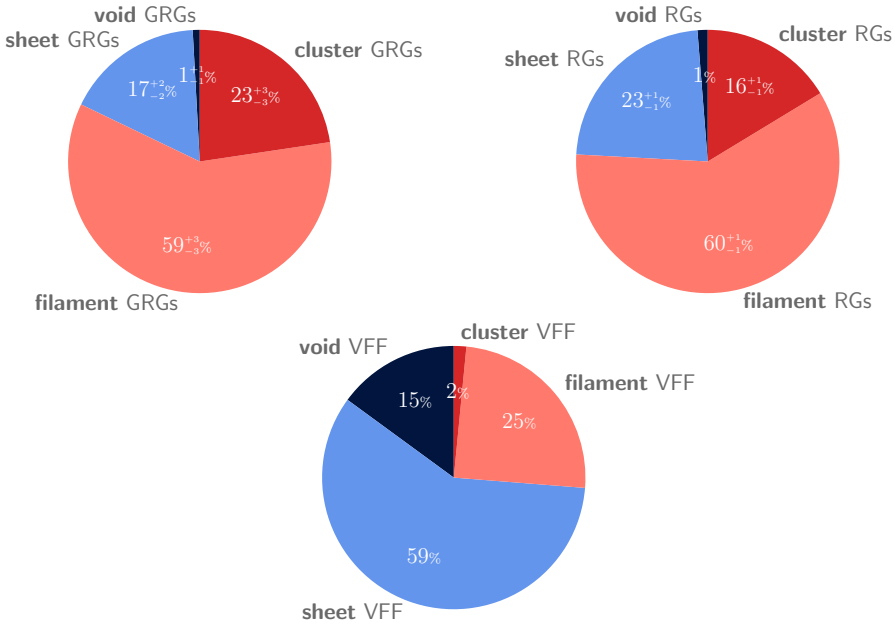


Figure 7.15: Dynamical environment classification of giant radio galaxies (left), radio galaxies in general (centre), and the Local Universe in its entirety (right; from Table 3 of Leclercq et al., 2015). We define clusters, filaments, sheets, and voids in the *T*-web sense. These distributions depend on the scale to which the Cosmic Web density field is smoothed; in this case, the smoothing scale is 2.9 Mpc h^{-1}. If giants and RGs were scattered uniformly throughout the Cosmic Web, their distributions would be similar to that of the volume-filling fractions (VFFs). Instead, observed RGs — and observed giants in particular — favour cluster and filament environments.

be excluded. A sheet environment is very unlikely (i.e. $p_3 < 20\%$) for 71% of giants; finally, a void environment is nearly always very unlikely (i.e. $p_4 < 20\%$) — this holds for 99% of giants.

The corresponding distributions for LoTSS DR1 RGs (not shown) are largely similar. This is also clear from Fig. 7.15, in which we compare the environment class distributions of giants, LoTSS DR1 RGs, and the Local Universe in its entirety. Evidently, giants and general RGs are not scattered uniformly throughout the Cosmic Web, but trace cluster and filament environments. Remarkably, nearly a quarter of all giants occurs in clusters, which have a volume-filling fraction (VFF) of just 1–2%. It is easy to derive that the ratio between a GRG environment class probability (left pie) and the corresponding VFF (right pie) yields the GRG number density of that class relative to the cosmic mean GRG number density. Of course, the analogous statement for RGs also holds. In Table 7.3, we present Local Universe relative number densities for giants and general RGs. Clusters are strongly (giant) radio galaxy overdense; voids are strongly (giant) radio galaxy underdense.

Table 7.3: Number densities of observed Local Universe giants and general RGs relative to their cosmic means. These numbers depend on environment class definitions and the scale to which the Cosmic Web density field is smoothed; in this case, we use the T-web classification scheme and a smoothing scale of 2.9 Mpc h^{-1}. Clusters and filaments are (giant) radio galaxy overdense; sheets and voids are underdense.

	cluster environment	**filament** environment	**sheet** environment	**void** environment
GRGs	15 ± 2	2.4 ± 0.1	0.29 ± 0.04	0.06 ± 0.04
RGs	10.6 ± 0.6	2.41 ± 0.05	0.39 ± 0.02	0.08 ± 0.02

The left and central pie charts of Fig. 7.15 can be regarded as prior distributions over Cosmic Web dynamical states for luminous giants and RGs in or near the Local Universe. This fact can be useful whenever direct BORG SDSS environment characterisation is not possible. For a randomly picked luminous giant, Fig. 7.15's left pie chart specifies the prior probability distribution over dynamical states to be $\mathbf{p} = (23\%, 59\%, 17\%, 1\%)$. Interestingly, there is often additional information available to constrain the dynamical state. For example, multiwavelength imagery demonstrates that Alcyoneus (Oei et al., 2022a), one of the longest known radio galaxies, does not inhabit a galaxy cluster; in addition, the number of luminous galaxies in its vicinity excludes the possibility of a void environment.[13] Combining this information with the prior, we obtain — after rounding — the posterior probability distribution $\mathbf{p} = (0\%, 80\%, 20\%, 0\%)$. We thus estimate the probability that Alcyoneus inhabits a filament to be 80%. Of course, this line of reasoning generalises to any luminous giant outside of galaxy clusters that has massive galactic neighbours in its Mpc-scale vicinity.

7.5 Discussion

7.5.1 Cosmic Web density accuracy

In this work, we have determined Mpc-scale total matter densities around radio galaxies in the Local Universe, and used these to compare the environments of luminous giants to those of other RGs. A natural question to ask is how accurate these densities are, and to which degree they can be used in other analyses, such as in the fitting of dynamical models to individual radio galaxies (e.g. Hardcastle, 2018). To quan-

[13] Characterising Alcyoneus's Cosmic Web environment is of scientific interest, as it could help understand how the largest radio galaxies form. However, at $z = 0.25$, it is too far away to probe directly with the BORG SDSS.

tify accuracy, one would ideally know the ground-truth densities that the inferred densities are meant to approach. Problematically, however, no ground-truth densities are known — especially for filaments, the environment type which most radio galaxies seem to inhabit. For Local Universe clusters though, approximate masses are known. We exploit this fact to test the accuracy of our cluster densities. We immediately remark that the accuracy of our densities are likely to vary with density, so that the accuracy of cluster densities might not be informative of the accuracy of filament densities. This is because the 2LPT gravity solver of the BORG SDSS has particularly limited validity in the strongly non-linear cluster regime.

To test the cluster density accuracy, we cross-match our 260 selected giants with the SDSS DR12–based galaxy cluster catalogue of Wen & Han (2015). First, we select all clusters in the Local Universe ($z < z_{\mathrm{max}}$) and discard the rest. To specify each cluster's location, this catalogue provides the right ascension, declination, and redshift of the brightest cluster galaxy (BCG). Next, we cross-match in an angular sense, and consider a radio galaxy matched with a BCG if their angular separation is less than 3″. We thus retain 60 BCG giants and 113 BCG LoTSS DR1 RGs. Figure 7.2 shows six members (10%) of this BCG giants sample. The perturbed jet and lobe shapes seen in the radio, and the massive ellipticals and crowded environments seen in the optical, together visually demonstrate the adequacy of our BCG giant identification. We can identify LoTSS DR1 BCG RGs with equal confidence.

As a quality check of the cluster masses, we test whether they are accurate enough to contain fingerprints of known physical effects. In particular, we test whether clusters with RG-generating BCGs are more massive than other Local Universe clusters. The AGN feedback paradigm suggests that this should be the case: in low-mass clusters, AGN feedback converts cool cores to non-cool cores within a gigayear, shutting down further AGN feedback for many gigayears to come; by contrast, high-mass clusters maintain stable AGN feedback cycles with gigayear-scale periods (Nobels et al., 2022), and thus their BCGs have a higher probability of being observed while generating RGs. The median mass of clusters with RG-generating BCGs is $M_{500} = 1.1 \cdot 10^{14}\, M_{\odot}$, while for other clusters it is $M_{500} = 1.0 \cdot 10^{14}\, M_{\odot}$. A Kolmogorov–Smirnov test yields a p-value $p = 7\%$, suggesting to reject the null hypothesis of equality in distribution only for high significance levels (e.g. $\alpha = 10\%$). We also test whether clusters with *giant*-generating BCGs are more massive than other clusters. The median mass of clusters with giant-generating BCGs is $M_{500} = 1.2 \cdot 10^{14}\, M_{\odot}$, while for other clusters it is $M_{500} = 1.0 \cdot 10^{14}\, M_{\odot}$. A Kolmogorov–Smirnov test yields $p = 5\%$, suggesting to reject equality in distribution only for rather high significance levels (e.g. $\alpha = 5\%$ or $\alpha = 10\%$).

As a first quality check of our densities, we test whether BCG giants have been

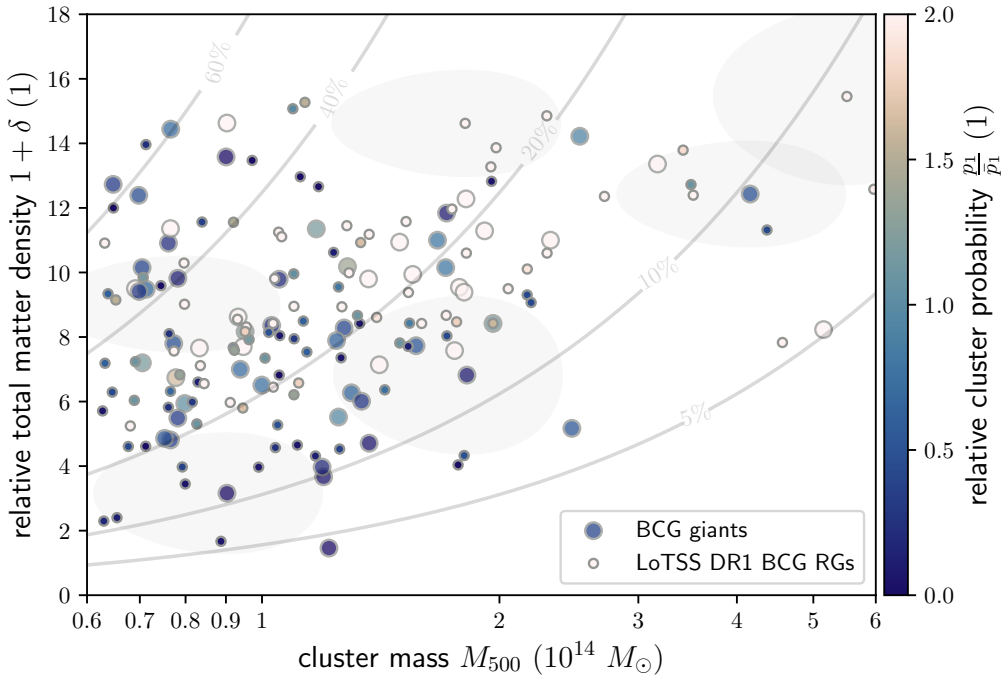


Figure 7.16: Comparison between environmental estimates for giants (large dots) and LoTSS DR1 RGs (small dots) generated by BCGs in the Local Universe. We compare cluster masses M_{500} from Wen & Han (2015) with relative densities $1+\delta$ from the BORG SDSS, and colour the dots based on their cluster probability p_1 (relative to the average cluster probability $\bar{p}_1$ for giants and LoTSS DR1 RGs). For three randomly chosen giants and three randomly chosen LoTSS DR1 RGs, we visualise uncertainties. Generally, as expected, radio galaxy–generating BCGs claimed to be in more massive clusters by Wen & Han (2015) have higher BORG SDSS densities and *T*-web cluster probabilities. The contours indicate fractions of the density expected through M_{500}, and show that our relative densities fall short significantly in the cluster regime — by a factor that ranges from less than two to more than an order of magnitude. Densities of low-mass clusters are less biased than densities of high-mass clusters. We use flexible voxel method densities here.

assigned higher densities than non-BCG giants. The median relative densities for BCG giants and non-BCG giants are $1+\delta = 8.5$ and $1+\delta = 7.2$, respectively. A Kolmogorov–Smirnov test yields $p = 0.2\%$, indeed suggesting to reject the null hypothesis of equality in distribution. Similarly, we test whether our LoTSS DR1 BCG RGs have been assigned higher densities than our LoTSS DR1 non-BCG RGs. The median relative densities for BCG RGs and non-BCG RGs are $1+\delta = 8.4$ and $1+\delta = 5.9$, respectively. A Kolmogorov–Smirnov test yields $p \ll 0.1\%$, again suggesting to reject equality in distribution.

Next, we show the relation between cluster mass and total matter density in Fig. 7.16. Although it is qualitatively clear that BCG RGs in more massive clusters have higher

inferred densities, the densities are all lower than one would expect from spreading out the cluster mass over a BORG SDSS voxel. (For reference, a cluster of mass $M \sim 10^{14}\, M_\odot$ should have a 2.9 Mpc h^{-1}–scale relative density $1 + \delta \sim 40$.) This systematic discrepancy cannot be attributed to errors in the cluster masses, as Wen & Han (2015) have made sure that their estimates are unbiased. The alternative is to conclude that our densities are biased low in the cluster regime. This is unsurprising, as the assumptions of the BORG SDSS's 2LPT gravity solver are violated in clusters. Our densities are least accurate for high-mass clusters, where they can be off by an order of magnitude, and most accurate for low-mass clusters, where they can be less than a factor two too low.

It is unclear whether, and if so how, these accuracy characterisations can be extrapolated to the filament density regime. The increased applicability of 2LPT to the filament density regime suggests that the accuracy will increase. We therefore expect our filament densities to be correct up to a small factor of order unity.

Until Bayesian Cosmic Web reconstructions are further developed, we urge caution in using our inferred densities as precision estimates: in filament environments, they should be trusted up to a factor of order unity only; in cluster environments, for the moment, other sources (such as cluster catalogues) appear to provide more accurate density estimates. If one would insist in using the cluster densities, then Fig. 7.16 suggests that one should correct them by multiplying by a factor of $\sim$5.

7.5.2 Cosmic Web dynamical state accuracy

Similar questions of accuracy can be raised about the Cosmic Web dynamical state distributions for luminous giants and RGs presented in the left and central pie charts of Fig. 7.15. To test the cluster GRG and cluster RG percentages, we can again use the cluster catalogue of Wen & Han (2015).

Cluster giants can be subdivided into BCG giants and non-BCG cluster giants. In Sect. 7.5.1, we determined that 60 out of 260 giants (23%) are BCG giants.[14] We define non-BCG cluster giants as giants generated by non-BCG galaxies for which the comoving distance to the BCG is less than some Mpc-scale threshold. We obtain approximate comoving coordinates from right ascensions, declinations, and spectroscopic redshifts, assuming vanishing peculiar motion. In order to allow for a fair comparison to the BORG SDSS, we set the Mpc-scale threshold to a BORG SDSS voxel side length: 2.9 Mpc h^{-1}. Although much larger than a galaxy cluster virial

[14] In comparison, Dabhade et al. (2020b) found that only 20 out of 128 LoTSS DR1 giants with $z \leq 0.55$ (16%) were BCG giants. We hypothesise that we find a higher occurrence of BCG giants because our study is restricted to the Local Universe ($z < z_{\mathrm{max}} := 0.16$), where cluster catalogues are generally more complete.

radius, this threshold does ensure that cluster galaxies with significant peculiar motion — a standard deviation $\sigma_{v_p} = 10^2$–10^3 km s^{-1} is typical — have a reasonable chance to be linked to their cluster. For this choice of threshold, we find 21 non-BCG cluster giants. We thus find a total of 81 cluster giants (31%), which is more than expected based on the BORG SDSS measurement (23^{+30}_{-3}%) shown in Fig. 7.15's GRG pie chart. This suggests that the BORG SDSS underpredicts the occurrence of cluster giants. Indeed, the strongly underpredicted densities of galaxy clusters discussed in Sect. 7.5.1 also leave their mark on the accuracy of our dynamical states. For example, the giant shown in the top-left panel of Fig. 7.2 and tabulated on the tenth row of Table 7.1, clearly occupies a cluster-like environment with an estimated $M_{500} = 1.3 \cdot 10^{14}\, M_\odot$, but is erroneously assigned an 80% chance to inhabit a filament.

Similarly, cluster RGs can be subdivided into BCG RGs and non-BCG cluster RGs. In Sect. 7.5.1, we determined that 113 out of 1443 selected LoTSS DR1 RGs (8%) are BCG RGs. By defining non-BCG cluster RGs in a way analogous to non-BCG cluster giants, we find 109 non-BCG cluster RGs.[15] We thus find a total of 222 cluster RGs (15%), which is consistent with the BORG SDSS measurement (16^{+10}_{-1}%) shown in Fig. 7.15's RG pie chart.

To assess the accuracy of the filament GRG and RG fractions shown in Fig. 7.15, we cross-match our giants and selected LoTSS DR1 RGs with the galaxy group and cluster catalogue of Tempel et al. (2017), which provides estimates of M_{200}. Wen & Han (2015) define clusters through $M_{500} \geq 0.6 \cdot 10^{14}\, M_\odot$, which translates to $M_{200} \geq 0.9 \cdot 10^{14}\, M_\odot$: a typical conversion is $M_{500} = 0.7\, M_{200}$ (e.g. Pierpaoli et al., 2003). To separate groups from clusters, we therefore define groups through $M_{200} < 0.9 \cdot 10^{14}\, M_\odot$. As with our cluster tests, we associate RGs to groups on the basis of estimated comoving coordinates, assuming vanishing peculiar motion. In the case of groups, this assumption is more reasonable. We again use a 2.9 Mpc h^{-1} association threshold. We associate 155 out of 260 giants (60%) and 903 out of 1443 RGs (63%) with galaxy groups. These percentages provide lower bounds to the occurrence of filament giants and RGs, as there are also filament giants and RGs that do not reside in galaxy groups. It is therefore likely that the filament GRG and RG fractions of Fig. 7.15 (59^{+30}_{-3}% and 60^{+10}_{-1}%, respectively) are underestimates.

Finally, as a probe of sheet and void giants and RGs, we explore the occurrence of giants and RGs for which the closest galaxy group or cluster is more than 10 Mpc away. Thus, we identify 18 out of 260 giants (7%) and 115 out of 1443 RGs (8%)

[15] Whereas three-quarters of all cluster giants are BCG giants, only half of all cluster RGs are BCG RGs. Thus, whenever a giant is found in a cluster, it is most likely generated by the BCG; by contrast, whenever a general RG is found in a cluster, a BCG origin is just as likely as a non-BCG origin.

as occupying sheets and voids. This suggests that the combined sheet and void GRG and RG fractions of Fig. 7.15 ($18^{+2}_{-2}\%$ and $24^{+1}_{-1}\%$, respectively) are significant overestimates.

Rounded to an accuracy of 10%, our ancillary environment analysis suggests an environment class probability distribution for luminous giants of $\mathbf{p} = (30\%, 60\%, 10\%)$, representing clusters, filaments, and a combination of sheets and voids, respectively. For RGs, it suggests $\mathbf{p} = (20\%, 70\%, 10\%)$, again representing clusters, filaments, and a combination of sheets and voids. Both distributions are in fair agreement with the BORG SDSS measurement $\mathbf{p} = (20\%, 60\%, 20\%)$. The ancillary analysis reinforces the idea that, in the Local Universe, luminous giants occupy denser environments than general RGs detectable at LoTSS-like image qualities.

7.5.3 Cosmic Web density resolution

All Cosmic Web density and dynamical state measurements presented in this work correspond to a density field with a resolution, or *smoothing scale*, of 2.9 Mpc h^{-1}. In fact, given the granular nature of matter, one must always adopt a smoothing scale in order to define an informative notion of density.[16] Ideally, this scale is chosen such that the resulting density field varies smoothly over the physical scales of interest — avoiding both unwanted stochastic variations over space and time (as would occur when the smoothing scale is too small), and excessive blurring of the physical phenomena under study (as would occur when the smoothing scale is too large). In the case of this work, the smoothing scale of the Cosmic Web density field could be considered too large. It is set by the limitations of the BORG SDSS 2LPT gravity solver, the ambition to cover the entire Local Universe in the direction of the SDSS MGS footprint, and — most practically — the relation between the number of voxels and the numerical cost of generating the HMC Markov chain. As galaxy clusters have typical radii of $\sim$1 Mpc h^{-1}, they are not resolved by the BORG SDSS; the same therefore certainly holds for galaxy groups. Thus, even if the gravitational dynamics of the reconstructions were exact, the utility of the BORG SDSS in improving dynamical model fits to individual RGs is limited: the smoothing scale of the density field is larger than the scale required to resolve (beta model–like) density profiles within filaments and clusters (0.1–1 Mpc h^{-1}). At best, our densities reflect the correct multi-Mpc-scale averages, but these are still much lower than the, say, central 1 Mpc–scale densities. This is because the massive galaxies that usually host RGs tend to occupy the bottom of their local gravitational potential wells, where the large-scale density is typically

[16] As there are many situations in which there exists a natural smoothing scale, the smoothing scale is often left unmentioned.

higher than everywhere in the vicinity. Relying on simulations, Oei et al. (2023a) have estimated the relation between the 2.9 Mpc h^{-1}–scale relative total matter density $1 + \delta$ (as used throughout this work), and the 1 Mpc–scale relative IGM density $1 + \delta_{\rm IGM}$. For galaxies with a stellar mass $M_\star = 10^{11}\,M_\odot$, their Fig. 11 suggests $1 + \delta_{\rm IGM} = 10 \cdot (1 + \delta)^{0.75}$. For luminous giants, the mode of the 2.9 Mpc h^{-1}–scale relative total matter density RV is $1 + \delta = 3.5$ for the fixed voxel method and $1+\delta = 6.3$ for the flexible voxel method (see Fig. 7.6); the corresponding host-centred 1 Mpc–scale relative IGM density is $1 + \delta_{\rm IGM} = 23$ for the fixed voxel method and $1 + \delta_{\rm IGM} = 40$ for the flexible voxel method. This relation in principle allows one to perform Bayesian inference of the distributions of the 1 Mpc–scale Cosmic Web density around RG hosts — possibly even conditioned on RG radio luminosity, as in Sect. 7.4.2 — by converting parametrised high-resolution distributions to the lower BORG SDSS resolution, applying heteroskedastic measurement errors, and finally comparing the resulting distribution to data. We leave this to future work.

Concerning resolution, we finally warn that care must be taken when calculating (AGN feedback–related) quantities that relate quadratically to density, such as the local bremsstrahlung emissivity: these tend to be underestimated if computed from low-resolution densities. To see why, we consider ρ, the density at some natural smoothing scale, alongside $\langle\rho\rangle$, the density averaged over some larger-than-natural smoothing scale. We assume that only the latter is known. Next, we pick a quantity $y = a\rho^2$, which we aim to compute at the larger smoothing scale: $\langle y\rangle = \langle a\rho^2\rangle = a\langle\rho^2\rangle$. Naively, one would be tempted to proceed in calculating $\langle y\rangle$ by assuming $\langle\rho^2\rangle \approx \langle\rho\rangle^2$, as to use $\langle y\rangle \approx a\langle\rho\rangle^2$. However, by doing so, one risks underestimating $\langle y\rangle$, because $\langle\rho^2\rangle \geq \langle\rho\rangle^2$ by Jensen's inequality. In Appendix 7.A6, we explicitly calculate the ratio between $\langle\rho^2\rangle$ and $\langle\rho\rangle^2$ for sheets and voids, filaments, and clusters. For sheets and voids, this ratio is at most $\frac{4}{3}$; for filaments, it is at most a number of order unity; for clusters, it can exceed a thousand.

7.5.4 Spectroscopic redshift selection

To localise RGs in the Cosmic Web and subsequently measure their relative total matter densities, we require them to have a spectroscopic redshift. By selecting on spectroscopic redshift availability, we implicitly select on both Cosmic Web density and radio luminosity. The key reason is that spectroscopic redshifts are preferentially available for galaxies with high optical flux densities — so galaxies that are nearby,[17]

[17] There is an exception to this rule: physically large galaxies at low redshifts, such as ellipticals in nearby galaxy clusters, can be angularly too extended for SDSS fibers and are therefore sometimes excluded from SDSS's spectroscopic survey. Excluding the RGs associated to these galaxies from our

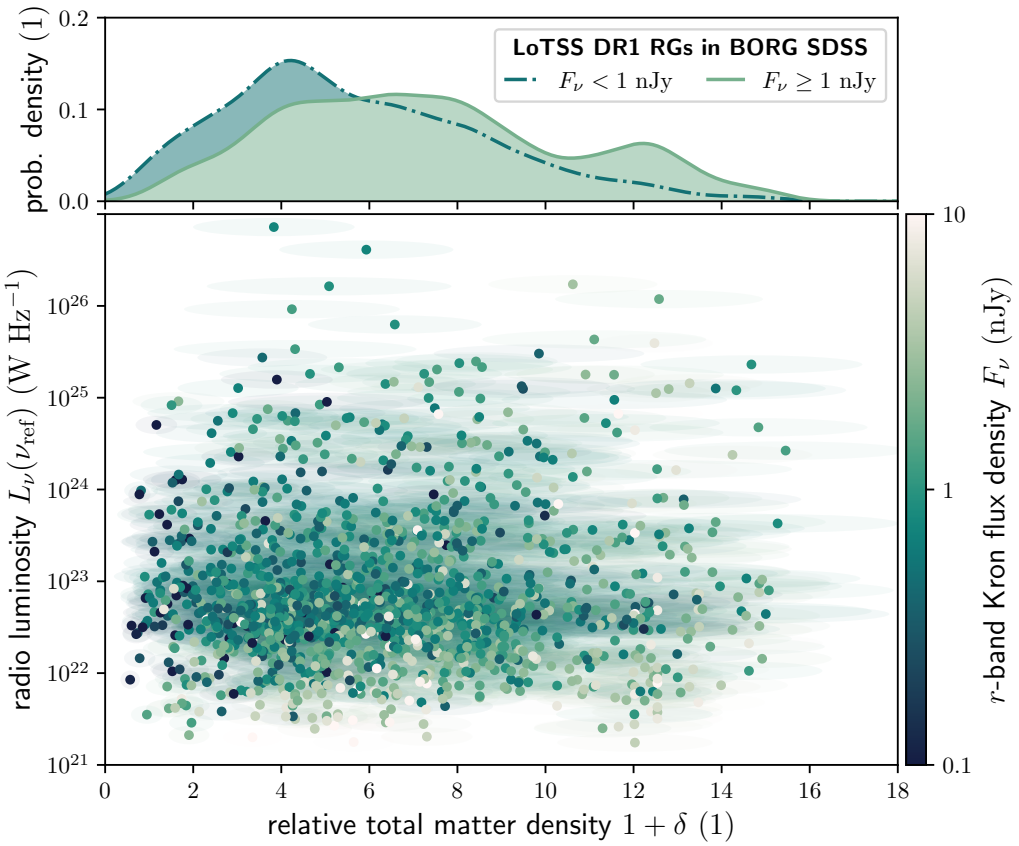


Figure 7.17: Observations underpinning the RG radio luminosity–Cosmic Web density relation of Fig. 7.8, but with Pan-STARRS r-band Kron flux densities F_ν of the host galaxies indicated. We show uncertainty ellipses whose semi-major and semi-minor axes represent one standard deviation along each dimension. In the top panel, we show KDE Cosmic Web density distributions for $F_\nu < 1$ nJy, and for $F_\nu \geq 1$ nJy. These suggest that spectroscopic redshift selection biases the observed Cosmic Web density distributions high.

and galaxies that have high optical luminosities. Galaxies with high optical luminosities tend to inhabit dense environments and, when active, tend to generate RGs with high radio luminosities.

To demonstrate that spectroscopic redshift selection indeed affects our RG radio luminosity–Cosmic Web density relation, Fig. 7.17 revisits the same Local Universe LoTSS DR1 RG data as shown in Fig. 7.8, but now with the Pan-STARRS r-band Kron flux densities F_ν of the host galaxies indicated. The top panel of Fig. 7.17 shows

samples introduces a bias against high density environments. This artificially vacates the right sides of Figs. 7.5 and 7.8, and leads towards an underestimation of the occurrence of cluster RGs in our environment analysis.

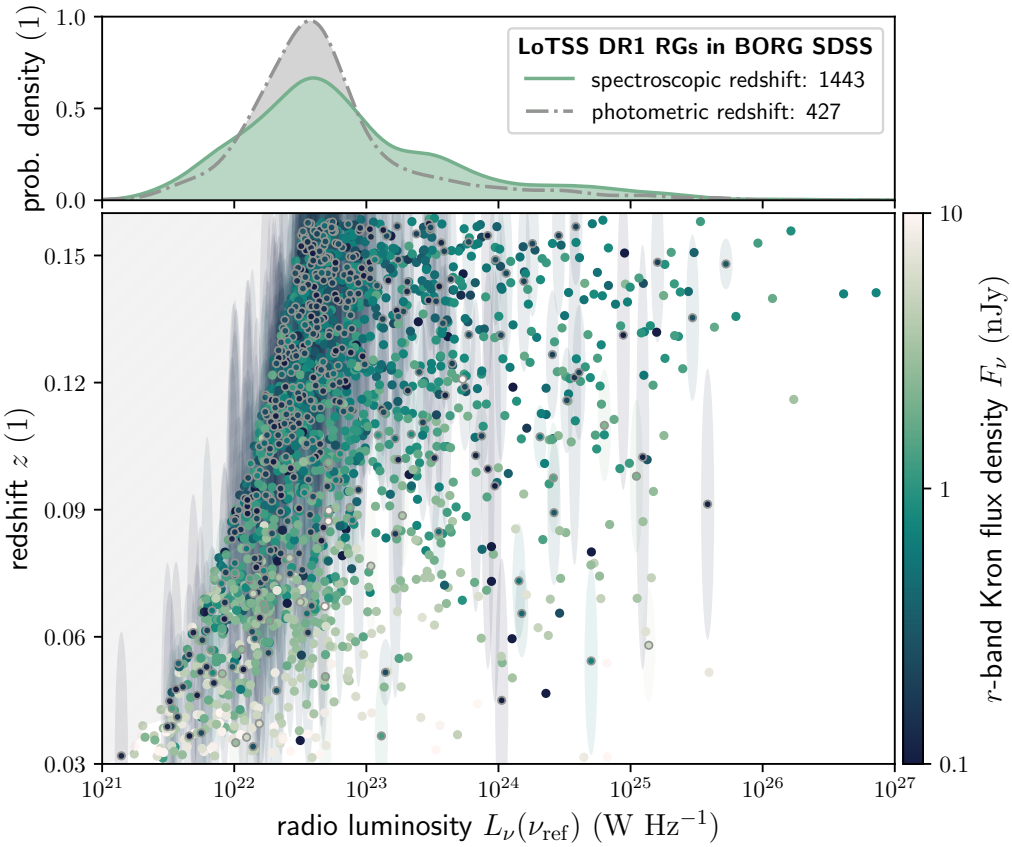


Figure 7.18: Overview of radio luminosity selection resulting from spectroscopic redshift selection, radio surface brightness limitations, and the correlation between radio and optical luminosity. The top panel compares radio luminosity distributions at $\nu_{\mathrm{ref}} = 150$ MHz of LoTSS DR1 RGs in the constrained part of the BORG SDSS, distinguishing between those with spectroscopic redshifts (green) and those without (grey). (We perform KDE on samples of $\log_{10}(L_\nu(\nu_{\mathrm{ref}}) \cdot (\mathrm{W\,Hz^{-1}})^{-1})$ rather than of $L_\nu(\nu_{\mathrm{ref}})$, and thus the probability densities on the vertical axis are dimensionless.) Radio galaxies with spectroscopic redshifts dominate (77%) the population. The bottom panel shows the two subpopulations in radio luminosity–redshift parameter space, with RGs without spectroscopic redshifts marked (grey circles). We show uncertainty ellipses whose semi-major and semi-minor axes represent half a standard deviation along each dimension. One retrieves the top panel by collapsing the data on the horizontal axis.

KDEs of $1 + \Delta_{\mathrm{RG,obs}}$ for hosts with $F_\nu < 1$ nJy, and for hosts with $F_\nu \geq 1$ nJy. Hosts with low optical flux densities tend to inhabit more tenuous environments than hosts with high optical flux densities. By requiring spectroscopic redshifts, we have discarded RGs with only photometric redshifts, whose hosts almost exclusively have $F_\nu < 1$ nJy. As a result, our estimates of $1+\Delta_{\mathrm{RG,obs}}$, $1+\Delta_{\mathrm{RG}}$, and $1+\Delta_{\mathrm{RG}} \mid L_\nu = l_\nu$ are biased high. The severity of the bias depends on the percentage of RGs that have been

selected out. As discussed in Sect. 7.3.2, 23% of all LoTSS DR1 RGs within the constrained BORG SDSS volume lack a spectroscopic redshift. Our sample of luminous giants is much less affected: only 7% of the population within the constrained BORG SDSS volume lack spectroscopy. Thus, correcting for spectroscopic redshift selection would shift the Cosmic Web density distribution for general RGs to somewhat lower densities, while it would not appreciably shift the corresponding distribution for luminous giants. The result would be an even bigger discrepancy between the Cosmic Web density distributions of luminous giants and general RGs than established in this work, strengthening our results.

Spectroscopic redshift selection also leaves an imprint on the sample's radio luminosities. The top panel of Fig. 7.18 shows that the radio luminosities of the 1870 LoTSS DR1 RGs in the constrained BORG SDSS volume are distributed differently between the subpopulations with and without spectroscopic redshifts. In particular, the subpopulation with spectroscopic redshifts is more spread out over radio luminosity compared to the subpopulation without: both very low and very high values are more probable.[18] To understand why, we turn to the bottom panel of Fig. 7.18. The LoTSS 6″ noise level, the inverse square law, and cosmological surface brightness dimming together determine an RG radio luminosity detection threshold that increases with redshift. This threshold induces the hatched region, which is inaccessible to LoTSS-like radio observations. As a straightforward consequence, RGs with low radio luminosities can be detected only at low redshifts. It is not immediately clear whether the host galaxies of low–radio luminosity RGs at low redshifts have low or high optical flux densities. On the one hand, low radio luminosities suggest low optical luminosities: both forms of spectral luminosity correlate positively with stellar mass (as evinced by e.g. Figs. 2 of Mahajan et al., 2018; Sabater et al., 2019) and therefore[19] positively with eachother. However, one must again be cautious of selection effects, which can inject strong but spurious luminosity–luminosity correlations in observations (Singal et al., 2019). On the other hand, observing low-redshift galaxies is easier, again because of the inverse square law and cosmological surface brightness dimming. The bottom panel of Fig. 7.18, and in particular its lower left corner, shows that the correlation between radio and optical luminosity is not strong enough to deny the observational advantages of low redshifts: low–radio luminosity RGs at low redshifts have high optical flux densities. As a result, spectroscopic redshifts are more common for such RGs. Similarly, the correlation between radio and optical luminosity explains that spectroscopic redshifts are more common for high–radio lu-

[18] The distribution is not only more dispersed, but is also shifted rightwards: the median and mean radio luminosities of the spectroscopic subpopulation are 1.3 and 3.2 times larger, respectively.

[19] There exist cases in which this argument does not hold: correlation is not necessarily transitive.

minosity RGs.

This indirect radio luminosity selection effect is not expected to change the radio luminosity–Cosmic Web density relation inferred in this work and shown in e.g. Figs. 7.8 and 7.9. This is because the inferred relation concerns the Cosmic Web density distribution at given radio luminosity. Changes in the number of RGs available in each decade of radio luminosity — and in particular an increase in the number of RGs with 10^{22} W Hz^{-1} $< L_\nu(\nu = 150\ \mathrm{MHz}) < 10^{23}$ W Hz^{-1} — do not materially affect our analysis, as we already perform data rebalancing that downweights our sample's most populated decades.[20]

7.5.5 Future outlook

Excitingly, there appear to be many opportuntities to improve our understanding of the large-scale environments of radio galaxies through Cosmic Web reconstructions, both with current and near-future data.

In this work, we have used LoTSS DR1 Local Universe RLAGN with spectroscopic redshifts from Hardcastle et al. (2019). With the release of the LoTSS DR2 radio data (Shimwell et al., 2022) and the associated optical cross-matching results (Hardcastle et al., 2023), it will become possible to redo our analysis on a footprint that is more than 13 times larger. Doing so would increase the size of the general RG sample by an order of magnitude, from $\sim 10^3$ to $\sim 10^4$. In turn, this would allow us to revisit Sect. 7.4.2's RG radio luminosity–Cosmic Web density relation, but with all of Fig. 7.8's six radio luminosity decades fully evenly weighted. If the positive correlation between radio luminosity and Cosmic Web density is genuine, it will emerge more clearly upon weighting evenly. Such a LoTSS DR2 sample might even be large enough to meaningfully investigate possible differences in the radio luminosity–Cosmic Web density relation for different types of RGs, such as those generated by quasars, and those generated by other AGN. In the near future, the WEAVE–LOFAR Survey (Smith et al., 2016) will generate $\sim 10^6$ spectra of LOFAR-selected galaxies, which will help alleviate the effects of spectroscopic redshift selection described in Sect. 7.5.4. Finally, the availability of optical spectra for all AGN in the current analysis and for those in similar follow-ups means that a comprehensive classification into HERGs and LERGs, and into HEGRGs and LEGRGs, appears possible. In turn, this classification would allow for studies into the relationships

[20] The hatched region of Fig. 7.18 shows that radio surface brightness selection causes our Local Universe RG sample to be strongly incomplete in the radio luminosity decades 10^{21}–10^{22} and 10^{22}–10^{23} W Hz^{-1}. Extending our sample in the lowest of these decades upweights this decade's importance, likely revealing that the positive scaling between RG radio luminosity and Cosmic Web density is stronger than we suggest in Sect. 7.4.2.

between AGN accretion mode, RG growth, and the surrounding Cosmic Web. As these AGN reside in the Local Universe, many could be further characterised by cross-matching with X-ray catalogues.

Another promising extension of our work would be to use Cosmic Web reconstructions that are complementary to or improve upon the BORG SDSS. Notably, the BORG 2M++ (Jasche & Lavaux, 2019) offers one such possibility: it provides density reconstructions for the entire sky and at a higher spatial resolution, and features more accurate selection functions, bias modelling, and gravitational dynamics than the BORG SDSS. However, its reconstructions are limited to $z_{\mathrm{max}} = 0.1$, whereas those of the BORG SDSS extend to $z_{\mathrm{max}} = 0.16$. Using the BORG SDSS and the BORG 2M++ in conjunction would allow for an extension of the number of giants for which a density can be determined: the catalogue of giants by Oei et al. (2023a), introduced in Sect. 7.2, contains 83 giants with spectroscopic redshifts $z_{\mathrm{s}} < 0.1$ that lie outside of the constrained BORG SDSS volume. Adding these to the 260 giants for which we have already determined densities would increase the number of BORG giants by 32%. Perhaps more importantly, there are 89 giants for which both BORG SDSS and BORG 2M++ density estimates are possible. Comparing these densities allows for a quality assessment of the BORG SDSS density measurements at large. In particular, given its better gravity solver, we expect BORG 2M++ cluster densities to perform much better in Sect. 7.5.1's Cosmic Web density–cluster mass test. In addition to the BORG 2M++, additional BORG runs are on their way. Besides, there exist comparable Cosmic Web reconstruction frameworks, such as COSMIC BIRTH (e.g. Kitaura et al., 2021). In any case, thanks to the combination of new spectroscopic galaxy surveys and advances in computing power, high-fidelity Bayesian Cosmic Web reconstructions will become available for an increasingly large fraction of the Local Universe and beyond. We expect that Cosmic Web density $1 + \delta$ will therefore become a standard RG observable, alongside properties such as projected proper length l_{p} and radio luminosity l_{ν}.

By building on the work of Leclercq et al. (2015), we have presented probabilistic Cosmic Web environment classifications, based on the T-web definition, for 260 giants and 1443 LoTSS DR1 RGs. However, Leclercq et al. (2015) provides additional environmental classifications for the BORG SDSS volume that allow us to measure and understand the Cosmic Web environments of giants and other radio galaxies in complementary ways. Future work could explore and compare RG environments in the T-web, DIVA, ORIGAMI, and LICH sense (Leclercq et al., 2016, 2017).

7.6 Conclusion

Using Bayesian Cosmic Web reconstructions of the Local Universe, we compare the large-scale environments of giant radio galaxies with those of the radio galaxy population in general. In particular, we measure multi-Mpc-scale Cosmic Web densities around the hosts of $\sim 10^2$ luminous giants and $\sim 10^3$ general RGs. This reveals that the currently observable population of giants inhabits denser regions of the Cosmic Web than general RGs — contradicting the popular hypothesis that giants should primarily form in the dilute Cosmic Web. Our interpretation is that high jet powers, as implied by the high radio luminosities of the known population of giants, enable these RGs to overcome the IGM's resistance to their Mpc-scale growth.

Next, we quantify the relation between radio luminosity, a proxy for jet power, and Cosmic Web density; to our knowledge, this is the first time this has been done. Our radio luminosity–conditioned Cosmic Web density distributions reinforce the idea that AGN that generate weak jets primarily reside in the dilute Cosmic Web, while AGN that generate powerful jets primarily reside in the dense Cosmic Web. We subsequently show that the radio luminosity–Cosmic Web density relation, which we have inferred using our sample of general RGs only, can accurately explain the density distribution observed for luminous giants.

Additional evidence corroborates that luminous giants form more often in the dense Cosmic Web than radio galaxies in general. Classifying Cosmic Web environments as clusters, filaments, sheets, or voids on the basis of the local gravitational field, we find that luminous giants occur more often in clusters than radio galaxies in general. We also find evidence that clusters with giant-generating BCGs are more massive than other clusters, while the evidence for the analogous claim for clusters with RG-generating BCGs is weaker.

We now summarise our main findings in more detail.

1. Using the BORG SDSS, we have performed the most physically principled measurements yet of the Cosmic Web environments of giant radio galaxies and radio galaxies in general. Our characterisations require spectroscopic redshifts, and are currently confined to the Local Universe ($z < z_{\rm max} := 0.16$). In particular, we have determined 2.9 Mpc h^{-1}–scale total matter densities and gravitational environment classifications for 260 giants, of which 208 (80%) have recently been discovered through the LoTSS DR2, and for 1443 LoTSS DR1 RGs.

2. While the marginal probability distributions of the Cosmic Web density around individual RGs are approximately lognormal, the probability distribution of

the Cosmic Web density for the RG population as a whole is approximately gamma. We provide a physical argument that motivates why the gamma distribution could have arisen. The probability distribution of the Cosmic Web density for luminous giants also appears roughly gamma, but favours higher densities. We show that heteroskedastic measurement errors, which are a general feature of our density determination method, spuriously shift the observed density distributions towards lower densities. We demonstrate a forward modelling method to correct for heteroskedasticity.

3. We quantify the radio luminosity–Cosmic Web density relation for general radio galaxies. We find that the 2.9 Mpc h^{-1}–scale Cosmic Web density distribution at a given 150 MHz radio luminosity is well described by $1+\Delta_{\mathrm{RG}} \mid L_\nu = l_\nu \sim \Gamma(k,\theta)$, where $k = 4.8 + 0.2 \cdot \lrcorner$, $\theta = 1.4 + 0.02 \cdot \lrcorner$, and $\lrcorner := \log_{10}(l_\nu \cdot (10^{23}\ \mathrm{W\,Hz^{-1}})^{-1})$. This result shows that more luminous RGs tend to live in denser regions of the Cosmic Web. Treating giants as ordinary radio galaxies with $l_\nu \geq l_{\nu,\mathrm{min}} := 10^{25}\ \mathrm{W\,Hz^{-1}}$, we use this relation to predict the Cosmic Web density distribution for luminous giants. The prediction is consistent with the *observed* distribution for giants: a Kolmogorov–Smirnov test yields a p-value of 0.7. Whether less luminous giants — assuming they exist — also form more often in the dense Cosmic Web, remains to be seen when surveys with increased surface brightness sensitivity commence. If such giants obey the same radio luminosity–Cosmic Web density relation as other RGs, this will not be the case.

4. We show that our methodology enables the inference of radio galaxy number densities as a function of Cosmic Web density. Whether or not the RG number density is a strong function of Cosmic Web density, depends on the radio luminosity considered. In clusters, the number densities of high-luminosity RGs (e.g. 10^{25} or $10^{27}\ \mathrm{W\,Hz^{-1}}$) are $\sim 10^2$ times higher than on average; in the same environment type, the number densities of low-luminosity RGs (e.g. 10^{21} or $10^{23}\ \mathrm{W\,Hz^{-1}}$) are just $\sim 10^1$ times higher than on average. Furthermore, we obtain tentative evidence that the number densities of low-luminosity RGs peak in low-mass clusters, before decreasing again. A possible explanation could be that, as cluster mass increases, the underlying population of active galaxies starts favouring the generation of high-luminosity RGs over low-luminosity RGs.

5. We use the BORG SDSS T-web classification to generate probability distributions over four Cosmic Web structure types — clusters, filaments, sheets, and

voids — for both luminous giants and general RGs. Luminous giants inhabit clusters more often ($23^{+3}_{-3}\%$ versus $16^{+1}_{-1}\%$). Independently, by cross-matching our radio galaxies with a cluster catalogue, we find that the former percentage presumably is an underestimate, while confirming the latter percentage. The same cross-matching procedure reveals tentative evidence ($p = 5\%$) that clusters with giant-generating BCGs are more massive than other clusters. At the same time, evidence that clusters with RG-generating BCGs are more massive than other clusters is weaker ($p = 7\%$), even though the sample involved is almost twice as big.

Although RG environment characterisations with Bayesian Cosmic Web reconstructions are still far from perfect, we have demonstrated that the current generation of reconstructions already allows for addressing open questions in radio galaxy research. Given the pitfalls of the BORG SDSS, we advise to treat every measured RG density with skepsis, and in particular warn that the densities of cluster RGs are strongly underestimated. Nevertheless, given the ever-increasing number of galaxies with spectroscopic redshifts, a growing list of technical improvements, and the widespread availability of computing power, exciting contemporary and near-future opportunities exist to expand and improve upon the results of this work. It is well possible that Cosmic Web density will soon become a staple quantity that helps to unravel the physics of radio galaxies.

M.S.S.L. Oei, R.J. van Weeren, and A. Botteon acknowledge support from the VIDI research programme with project number 639.042.729, which is financed by the Dutch Research Council (NWO). M.S.S.L. Oei warmly thanks Rafaël Mostert, Erik Osinga, and Wendy Williams for helpful discussions and the sharing of data. LOFAR data products were provided by the LOFAR Surveys Key Science project (LSKSP; https://lofar-surveys.org/) and were derived from observations with the International LOFAR Telescope (ILT). LOFAR (van Haarlem et al., 2013) is the Low Frequency Array designed and constructed by ASTRON. It has observing, data processing, and data storage facilities in several countries, which are owned by various parties (each with their own funding sources), and which are collectively operated by the ILT foundation under a joint scientific policy. The efforts of the LSKSP have benefited from funding from the European Research Council, NOVA, NWO, CNRS-INSU, the SURF Co-operative, the UK Science and Technology Funding Council and the Jülich Supercomputing Centre.

Data availability

Oei et al. (2023a)'s LoTSS DR2 GRG data are available at the Centre de Données astronomiques de Strasbourg (CDS) via https://cdsarc.cds.unistra.fr/cgi-bin/qcat?J/A+A/. The LoTSS DR1 RG data with optical identifications from the value-added catalogue of Williams et al. (2019) are available at https://lofar-surveys.org/dr1_release.html; those interested in Hardcastle et al. (2019)'s RLAGN subsample should contact Martin J. Hardcastle. The BORG SDSS data release is publicly available at https://github.com/florent-leclercq/borg_sdss_data_release. In-

dividual BORG SDSS HMC Markov chain samples are not available in the public domain; those interested should contact Florent Leclercq. The cluster catalogue of Wen & Han (2015) is available at the CDS.

7.A1 Cosmic Web localisation accuracy with spectroscopic and photometric redshifts

In this appendix, we show that RGs with spectroscopic redshifts can be reliably localised within Cosmic Web reconstructions, while RGs with only photometric redshifts generally cannot.

To determine the accuracy of our radio galaxy localisations in the Cosmic Web, we Monte Carlo simulate a radial comoving distance distribution for each RG and compare its dispersion to the BORG SDSS voxel length. Each RG will have a peculiar velocity v_p with respect to us. We treat v_p as a zero-mean Gaussian random variable (RV): $v_\mathrm{p} \sim \mathcal{N}(0, \sigma_{v_\mathrm{p}})$, and choose a standard deviation σ_{v_p} representative of conditions in low-mass galaxy clusters: $\sigma_{v_\mathrm{p}} = 100\ \mathrm{km\ s^{-1}}$. Similarly, we treat the measured redshift z_m as an RV, and again assume Gaussianity: $z_\mathrm{m} \sim \mathcal{N}\left(\mu_{z_\mathrm{m}}, \sigma_{z_\mathrm{m}}\right)$. Our catalogue provides the parameters μ_{z_m} and σ_{z_m} for each RG. Using Eqs. 7.23 from left to right, we calculate the relative peculiar velocity RV β_p, the peculiar velocity redshift RV z_p, and the cosmological redshift RV z_c:

$$\beta_\mathrm{p} := \frac{v_\mathrm{p}}{c}; \quad z_\mathrm{p} = \sqrt{\frac{1+\beta_\mathrm{p}}{1-\beta_\mathrm{p}}} - 1; \quad z_\mathrm{c} = \frac{1+z_\mathrm{m}}{1+z_\mathrm{p}} - 1. \tag{7.23}$$

Finally, we calculate the radial comoving distance RV $r = r(z_\mathrm{c}) = ||\mathbf{r}||_2$. The distributions of r are approximately Gaussian: for each of the 248 giants in the BORG SDSS volume with a spectroscopic redshift and an associated error, we perform the Shapiro–Wilk test on 1000 Monte Carlo samples and reject the null hypothesis that r is Gaussian for one at significance level $\alpha = 0.01$.[21] For these giants, 95% of the standard deviations σ_r are between 1.00 Mpc h^{-1} and 1.12 Mpc h^{-1}; the sample median is 1.06 Mpc h^{-1}. By contrast, the 21 giants in the BORG SDSS volume with only a photometric redshift and an associated error have a σ_r that ranges from 20 Mpc h^{-1} to 110 Mpc h^{-1}; for details, see Table 7.4. To localise a giant in the BORG SDSS, precision up to the scale of a voxel, which are approximately 2.9 Mpc h^{-1} long, is

[21] If one uses sufficiently many Monte Carlo samples and a fixed significance level, any deviation from Gaussianity will become significant. Indirectly, the sample size and significance level together specify a degree of non-Gaussianity identifiable with such Shapiro–Wilk tests.

necessary — but sub-voxel localisation is redundant. We conclude that giants with a spectroscopic redshift can be tied to an individual voxel, even when taking peculiar velocity into account; on the other hand, giants with only a photometric redshift can be mislocalised more than a typical filament length, and are therefore not subjectable to a precise environment analysis. Of course, these conclusions extend to any radio galaxy with an associated host, or indeed to any galaxy.[22]

As only few (i.e. $\sim 10^2$) giants are currently known in the volume reconstructed by the BORG SDSS, it is worthwhile to perform host galaxy spectroscopic follow-up for those with photometric redshifts only. We provide a list with targets in Table 7.4.

7.A2 Cosmic Web density distribution: the gamma Ansatz

In this work, we have modelled several RG Cosmic Web density RVs, such as $1 + \Delta_{\mathrm{RG,obs}}$, $1 + \Delta_{\mathrm{RG}}$, and $1 + \Delta_{\mathrm{RG}} \mid L_\nu = l_\nu$, as gamma variates. This choice is largely driven by data (and a preference for simplicity) rather than by theory, although the requirements of continuity and a strictly positive support have a physical basis. Gamma distributions adequately fit $1 + \Delta_{\mathrm{RG,obs}}$ (see the bottom row of Fig. 7.6), and their use in modelling $1 + \Delta_{\mathrm{RG}} \mid L_\nu = l_\nu$ is justified by the RG prediction of $1 + \Delta_{\mathrm{GRG,obs}}$, which matches observations (see Fig. 7.11). Here, we post hoc theorise why the gamma distribution arises in the current context.

The late-time density in a voxel is, of course, equal to the late-time mass in the voxel divided by its (fixed) volume of $(2.9\ \mathrm{Mpc}\ h^{-1})^3$. Therefore, up to a dimensionful constant, the late-time density and the late-time mass have the same probability distribution. The late-time mass M in the voxel equals the sum of the mass aggregated over cosmic time. If we are to consider a simplistic treatment of mass aggregation, or structure formation, which environment is most relevant? The central pie chart of Fig. 7.15 suggests that most observed RGs have filament environments. In addition, Pasini et al. (2021) suggest that within filaments, most RGs inhabit galaxy groups, which grow by merging with other groups. Therefore, we consider a proto-filament — i.e. a massive structure that extends essentially along a single dimension — on which we consider a Poisson point process with constant spatial rate. The Poisson points represent peaks in the density field where galaxy groups arise. As with any homogeneous Poisson process, the distances between the points are exponentially distributed. It is now helpful to partition the filament in cells, each associated to a single

[22] By jointly inferring the large-scale density field and the radial distances to the galaxies used to constrain the former, it is possible to reduce individual photometric redshift uncertainties (Jasche & Wandelt, 2012; Tsaprazi et al., 2023), but these reductions are not extensive enough to alter our conclusions.

Table 7.4: Overview of all giants in the constrained BORG SDSS volume without a spectroscopic redshift. Roughly speaking, this volume corresponds to the SDSS DR7 MGS footprint up to $z_{\max} = 0.16$. The total matter density of the ambient Cosmic Web can only be reliably determined after spectroscopic follow-up. The giants are sorted (in descending order) by their radial comoving distance standard deviation σ_r obtained through Monte Carlo simulation, taking into account photometric redshift errors and a random, zero-centred peculiar velocity component with standard deviation $\sigma_{v_p} = 100$ km s^{-1}. After spectroscopic follow-up, the projected proper length l_p must be revised.

rank ↓	SDSS DR12 name	photometric redshift (1)	σ_r (Mpc h^{-1})	l_p (Mpc)
1	J122009.84+194833.0	0.15 ± 0.04	110	0.8
2	J155503.00+280430.9	0.10 ± 0.03	100	1.3
3	J134211.92+565839.3	0.11 ± 0.03	70	0.7
4	J130444.37+511119.0	0.09 ± 0.02	70	0.8
5	J122129.95+662644.0	0.10 ± 0.02	70	1.1
6	J152151.93+570635.0	0.07 ± 0.02	50	0.8
7	J112033.54+233559.0	0.12 ± 0.02	50	0.9
8	J140046.38+301900.0	0.06 ± 0.01	40	0.8
9	J090640.80+142522.9	0.13 ± 0.01	40	1.2
10	J112248.98+565243.5	0.08 ± 0.01	30	0.8
11	J085022.99+383547.2	0.14 ± 0.01	30	0.7
12	J145102.13+301227.0	0.16 ± 0.01	30	0.8
13	J152229.23+281911.9	0.13 ± 0.01	30	0.8
14	J140044.26+125219.8	0.11 ± 0.01	30	0.7
15	J090128.15+145158.1	0.14 ± 0.01	30	1.2
16	J145827.21+331312.0	0.11 ± 0.01	20	1.0
17	J152024.13+310557.6	0.06 ± 0.01	20	0.9
18	J102135.20+420022.3	0.13 ± 0.01	20	1.5
19	J172715.85+585220.5	0.15 ± 0.01	20	2.1
20	J141033.40+405932.4	0.14 ± 0.01	20	1.0
21	J134339.20+195301.7	0.15 ± 0.01	20	0.9

point. Assuming that the filament has a constant cross-sectional area and an approximately constant density, the volumes and masses of the cells are simply proportional to the distances between the points, and are thus also exponentially distributed. Finally, if a fixed fraction of the mass in each cell collapses into the cell's galaxy group, then the group masses within the filament — at early times — are, again, exponentially distributed. A group that exists at late times has built up its mass M by aggregating the masses of early-time groups. Assuming a constant temporal rate of early-time group aggregation, there exists a typical number N of early-time groups that con-

tribute to the mass of a given late-time group: $M = \sum_{i=1}^{N} M_i$. Now we invoke the fact that the sum of a fixed number of independent and identically distributed RVs with an exponential distribution is an RV with a gamma distribution. Thus, as the M_i are exponentially distributed, M is gamma distributed. This, in turn, implies that the late-time densities of the voxels in which late-time groups fall, are gamma distributed also. As late-time RGs trace late-time groups (and especially the more massive ones), the late-time RG density distribution should be approximately gamma.

In the future, by cross-matching RG catalogues with galaxy group catalogues with accurate masses, the mass distribution of RG-hosting groups in the Local Universe can be sampled and tested against the gamma distribution hypothesis.

7.A3 Modelling relative baryon density measurement heteroskedasticity

We must describe the measured RG relative baryon density given a true RG relative baryon density with a continuous probability distribution that has the positive half-line as its support; preferably, we use a simple analytical prescription.

Motivated by the quality of the fits shown in Fig. 7.5, we assume $1 + \Delta_{\mathrm{RG,obs}} \,|\, 1 + \Delta_{\mathrm{RG}} = 1 + \delta_i \sim \mathrm{Lognormal}(\mu, \sigma^2)$, and likewise for giants. The parameters μ and σ^2 are determined by assumptions on the RV's mean and variance. In particular, we assume that measured RG relative baryon densities are unbiased estimators of the underlying true RG relative baryon densities and that their variances are the same when the underlying true RG relative baryon densities are the same:

$$\mathbb{E}\left[1 + \Delta_{\mathrm{RG,obs}} \,|\, 1 + \Delta_{\mathrm{RG}} = 1 + \delta_i\right] = 1 + \delta_i; \tag{7.24}$$

$$\mathbb{V}\left[1 + \Delta_{\mathrm{RG,obs}} \,|\, 1 + \Delta_{\mathrm{RG}} = 1 + \delta_i\right] = g\left(1 + \delta_i\right), \tag{7.25}$$

where $g\left(1 + \delta\right)$ is some non-negative function characterising the heteroskedasticity. For the current work, we adopt a two-parameter power-law variance model $g\left(1 + \delta\right) := a\left(1 + \delta\right)^b$. From basic identities of the lognormal distribution, we find for this choice of g

$$\mu = \ln\left(1 + \delta\right) - \frac{1}{2}\ln\left(1 + a\left(1 + \delta\right)^{b-2}\right); \tag{7.26}$$

$$\sigma^2 = \ln\left(1 + a\left(1 + \delta\right)^{b-2}\right). \tag{7.27}$$

Using the fact that

$$f_{1+\Delta_{\mathrm{RG,obs}}}(1+\delta) = \int_0^\infty f_{1+\Delta_{\mathrm{RG,obs}}\,|\,1+\Delta_{\mathrm{RG}}=x}(1+\delta)\cdot f_{1+\Delta_{\mathrm{RG}}}(x)\ \mathrm{d}x, \tag{7.28}$$

we can deduce the distribution of $1+\Delta_{\mathrm{RG}}$ given some parametrisation of it and a procedure like MLE. We choose $1+\Delta_{\mathrm{RG}} \sim \Gamma(k,\theta)$ on the basis that $1+\Delta_{\mathrm{RG,obs}}$ appears to be much better described by a gamma distribution than by a lognormal distribution and that the effect of heteroskedasticity appears minor. We determine a and b from data.[23]

7.A4 Radio luminosity–Cosmic Web density relation: fixed voxel method

In Sect. 7.4.2, we have presented a quantification of the relation between RG radio luminosity and Cosmic Web density. In particular, we infer a probability distribution for the 2.9 Mpc h^{-1}–scale total matter density around the hosts of Local Universe radio galaxies with a given 150 MHz radio luminosity — which can range from 10^{21} to 10^{27} W Hz^{-1}. The MLE model shown in Fig. 7.8 and the one- and two-dimensional posterior marginals shown in Fig. 7.9, are based on densities measured through Sect. 7.3.1's flexible voxel method. In this appendix, through Figs. 7.19 and 7.20, we present the analogous results based on densities measured through Sect. 7.3.1's fixed voxel method. Both methods reveal a positive scaling between RG radio luminosity and Cosmic Web density.

7.A5 Relative number density derivation

We let the RV $1+\Delta_{\mathrm{CW}}$ represent the 2.9 Mpc h^{-1}–scale relative density at a randomly chosen point in the contemporary Cosmic Web, and let $f_{1+\Delta_{\mathrm{CW}}}$ be its PDF. Similarly, we let the RV $1+\Delta_{\mathrm{RG}}$ represent the 2.9 Mpc h^{-1}–scale relative density at a randomly chosen RG in the contemporary Cosmic Web, and let $f_{1+\Delta_{\mathrm{RG}}}$ be its PDF. We consider a cosmologically sized, comoving volume of extent V, in which a total of N_{RG} radio galaxies exist. The subvolume in which the relative density is between $1+\delta$ and $1+$

[23] For the fixed voxel method, we find $a = 0.4$ and $b = 1.1$; for the flexible voxel method, we find $a = 0.4$ and $b = 0.9$. We stick to a simple prescription here, but there appears to be enough BORG SDSS data to describe the heteroskedasticity with a more accurate (though more complex) model while still avoiding overfitting.

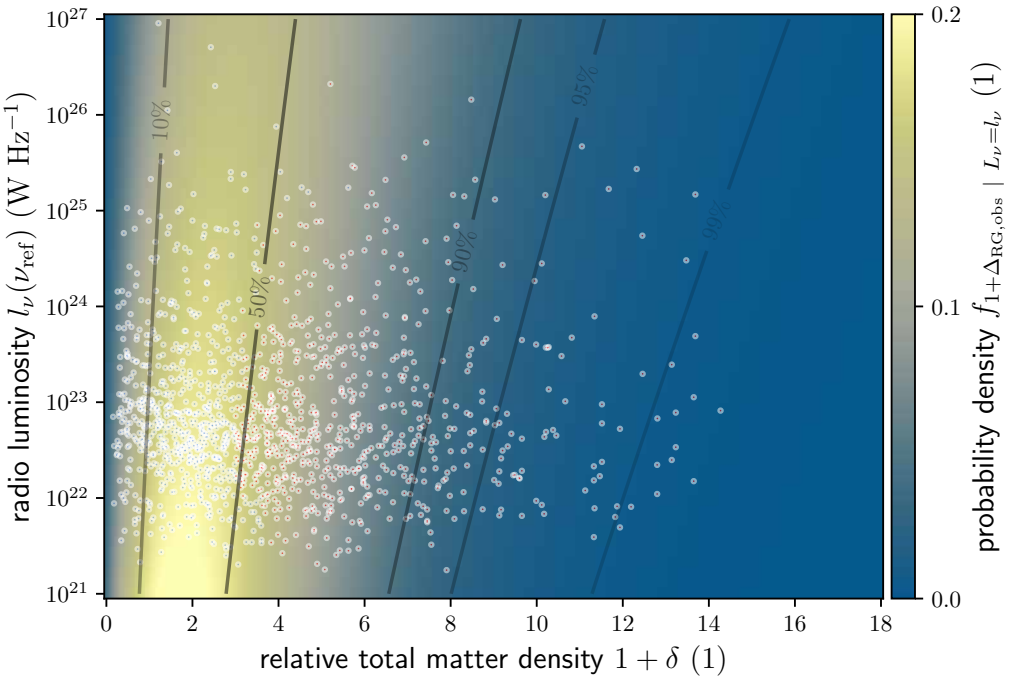


Figure 7.19: PDFs of the observed RG relative total matter density RV given a radio luminosity at $\nu_{\rm ref} = 150$ MHz, using MLE parameter values for the model described in Sect. 7.4.2. The black contours denote CDF values. We overplot all 1443 selected LoTSS DR1 RGs (dots), with those above the empirical median density coloured red, and those below coloured blue. We use fixed voxel method densities here. For flexible voxel method densities, see Fig. 7.8.

$\delta + \mathrm{d}\delta$ has extent

$$\mathrm{d}V = V f_{1+\Delta_{\rm CW}}(1+\delta)\,\mathrm{d}\delta. \tag{7.29}$$

The number of RGs in this subvolume is

$$\mathrm{d}N_{\rm RG} = n_{\rm RG}(1+\delta)\,\mathrm{d}V, \tag{7.30}$$

where $n_{\rm RG}(1+\delta)$ is the number density of RGs that would arise in an environment of constant density $1+\delta$. In practice, one never encounters environments of constant density; the average RG number density in the part of the Cosmic Web where the relative density is between $1+\delta_1$ and $1+\delta_2$ equals

$$\bar{n}_{\rm RG}(1+\delta_1, 1+\delta_2) = \frac{\int_{\delta_1}^{\delta_2} n_{\rm RG}(1+\delta) f_{1+\Delta_{\rm CW}}(1+\delta)\,\mathrm{d}\delta}{\int_{\delta_1}^{\delta_2} f_{1+\Delta_{\rm CW}}(1+\delta)\,\mathrm{d}\delta}. \tag{7.31}$$

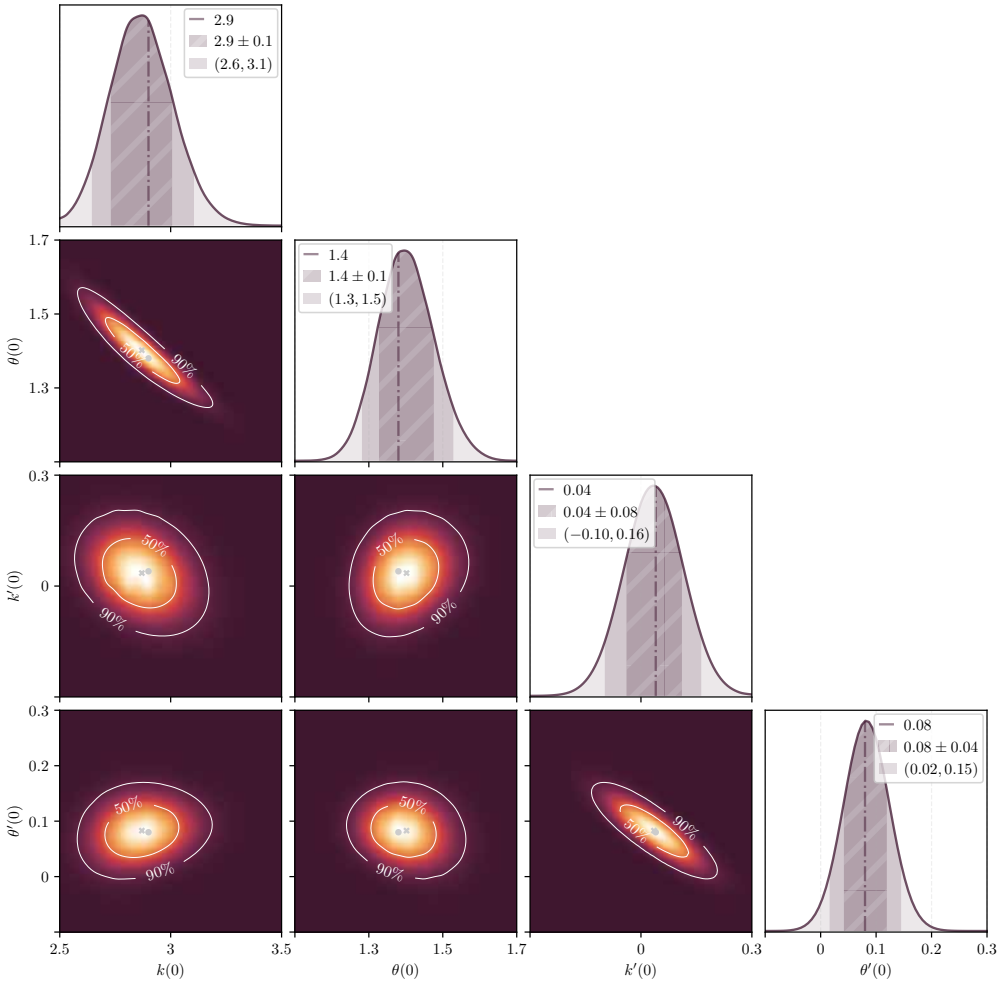


Figure 7.20: Posterior distribution over $k(0)$, $\theta(0)$, $k'(0)$, and $\theta'(0)$, based on selected Local Universe LoTSS DR1 RGs. We show all two-parameter marginals of the likelihood function, with contours enclosing 50% and 90% of total probability. We mark the maximum likelihood estimate (MLE) values (grey dot) and the maximum a posteriori (MAP) values (grey cross). The one-parameter marginals again show the MLE (dash-dotted line), a mean-centred interval of standard deviation–sized half-width (hashed region), and a median-centred 90% credible interval (shaded region). We use fixed voxel method densities here. For flexible voxel method densities, see Fig. 7.9.

As a result, the cosmic mean RG number density $\bar{n}_{\mathrm{RG}}$ is given by

$$\bar{n}_{\mathrm{RG}} := \int_{-1}^{\infty} n_{\mathrm{RG}}(1+\delta) f_{1+\Delta_{\mathrm{CW}}}(1+\delta)\, \mathrm{d}\delta. \tag{7.32}$$

The probability that an RG has a relative density between $1+\delta$ and $1+\delta+\mathrm{d}\delta$, $f_{1+\Delta_{\mathrm{RG}}}(1+\delta)\,\mathrm{d}\delta$, is

$$f_{1+\Delta_{\mathrm{RG}}}(1+\delta)\,\mathrm{d}\delta = \frac{\mathrm{d}N_{\mathrm{RG}}}{N_{\mathrm{RG}}} = \frac{n_{\mathrm{RG}}(1+\delta)\,V f_{1+\Delta_{\mathrm{CW}}}(1+\delta)\,\mathrm{d}\delta}{N_{\mathrm{RG}}}. \tag{7.33}$$

We note that

$$N_{\mathrm{RG}} = \bar{n}_{\mathrm{RG}} V, \tag{7.34}$$

and combine the last two equations to find

$$\frac{n_{\mathrm{RG}}(1+\delta)}{\bar{n}_{\mathrm{RG}}} = \frac{f_{1+\Delta_{\mathrm{RG}}}(1+\delta)}{f_{1+\Delta_{\mathrm{CW}}}(1+\delta)}. \tag{7.35}$$

In other words, through point-wise division of the PDFs of $1+\Delta_{\mathrm{RG}}$ and $1+\Delta_{\mathrm{CW}}$, both of which we determine in this work, we obtain the radio galaxy number density that would arise at a given density, up to a constant. This constant is (the reciprocal of) the cosmic mean RG number density; the LHS of Eq. 7.35 can thus be interpreted as the relative RG number density at $1+\delta$ — that is, the number density relative to the cosmic mean value. Fully analogously, the relative GRG number density at $1+\delta$ is given by

$$\frac{n_{\mathrm{GRG}}(1+\delta)}{\bar{n}_{\mathrm{GRG}}} = \frac{f_{1+\Delta_{\mathrm{GRG}}}(1+\delta)}{f_{1+\Delta_{\mathrm{CW}}}(1+\delta)}. \tag{7.36}$$

By multiplying both sides of Eq. 7.35 by $f_{L_\nu|1+\Delta_{\mathrm{RG}}=1+\delta}(l_\nu)$, using the identity $f_{X|Y=y}(x)f_Y(y) = f_{Y|X=x}(y)f_X(x)$, and dividing both sides by $f_{L_\nu}(l_\nu)$, we obtain

$$\frac{n_{\mathrm{RG}}(1+\delta)\cdot f_{L_\nu|1+\Delta_{\mathrm{RG}}=1+\delta}(l_\nu)}{\bar{n}_{\mathrm{RG}}\cdot f_{L_\nu}(l_\nu)} = \frac{f_{1+\Delta_{\mathrm{RG}}|L_\nu=l_\nu}(1+\delta)}{f_{1+\Delta_{\mathrm{CW}}}(1+\delta)}. \tag{7.37}$$

The numerator and denominator at the left-hand side are both number density *densities*: they denote radio galaxy numbers per unit of comoving volume and unit of radio luminosity. As their physical dimensions are distinct from those of ordinary number densities, we will denote them differently. Calling the numerator $\curlyvee_{\mathrm{RG}}(l_\nu, 1+\delta)$ and

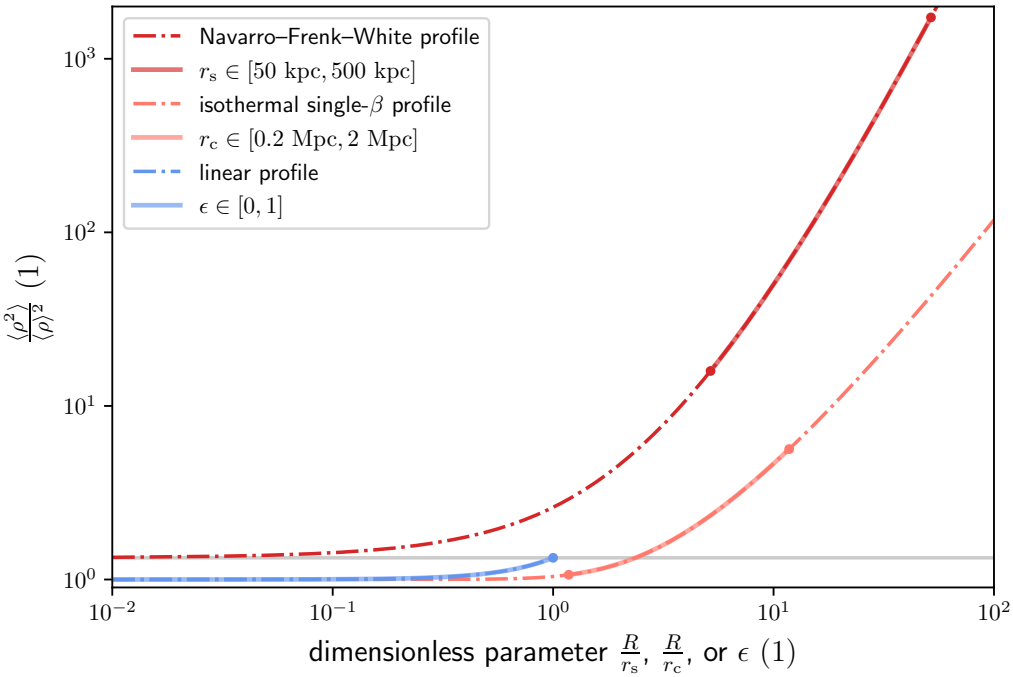


Figure 7.21: Multiplicative discrepancy between the mean squared density $\langle\rho^2\rangle$ and the squared mean density $\langle\rho\rangle^2$ in voxelised Cosmic Web reconstructions, for density profiles of clusters (red), filaments (orange), and sheets or voids (blue). The length scale $R \sim L$, where L is the voxel side length; here, $L = 2.9\ \mathrm{Mpc}\ h^{-1}$. The asymptote (grey) corresponds to value $\frac{4}{3}$.

the denominator $\bar{\daleth}_{\mathrm{RG}}(l_\nu)$, we write[24]

$$\frac{\daleth_{\mathrm{RG}}(l_\nu, 1+\delta)}{\bar{\daleth}_{\mathrm{RG}}(l_\nu)} = \frac{f_{1+\Delta_{\mathrm{RG}}|L_\nu=l_\nu}(1+\delta)}{f_{1+\Delta_{\mathrm{CW}}}(1+\delta)}. \tag{7.38}$$

7.A6 Disparity between mean squared density and squared mean density

Several physical quantities that relate to AGN and RGs, such as the bremsstrahlung emissivity around the host galaxy, scale with the square of IGM density ρ. In Sect. 7.5.3, we have argued qualitatively that such quantities are likely to be underestimated if calculated using low-resolution densities $\langle\rho\rangle$, such as the ones offered by the BORG SDSS. In this appendix, we explicitly calculate how much bigger $\langle\rho^2\rangle$ is than $\langle\rho\rangle^2$ — in the sense of a multiplicative Jensen's gap. Our calculations demonstrate a strong

[24] As all n-like symbols in the Roman and Greek alphabets are already in use, we use the Phoenician root $\daleth$.

dependence on Cosmic Web environment.

When the true density field varies over scales much larger than a voxel side length L, such as in sheets or voids, the density field variation within a voxel resembles a gradient from $(1-\varepsilon)\langle\rho\rangle$ to $(1+\varepsilon)\langle\rho\rangle$, where $\varepsilon \in [0,1]$. If the planes of constant density are parallel to two of the voxel's faces, we can derive

$$\frac{\langle\rho^2\rangle}{\langle\rho\rangle^2} = \frac{\varepsilon^2}{3} + 1, \text{ so that } 1 \leq \frac{\langle\rho^2\rangle}{\langle\rho\rangle^2} \leq \frac{4}{3}. \tag{7.39}$$

We show the multiplicative Jensen's gap of Eq. 7.39 in Fig. 7.21 (blue curve).

For a filament whose baryon density is modelled with an isothermal single-β model (Cavaliere & Fusco-Femiano, 1976, 1978) with $\beta = \frac{2}{3}$ and core radius r_c, we have

$$\frac{\langle\rho^2\rangle}{\langle\rho\rangle^2} = \frac{\left(\frac{R}{r_c}\right)^4}{\left(1+\left(\frac{R}{r_c}\right)^2\right)\ln^2\left(1+\left(\frac{R}{r_c}\right)^2\right)}, \tag{7.40}$$

where we consider a cylinder with length L and radius R, where R is such that the area of a cylindrical section perpendicular to the axis equals L^2, the area of a voxel face. Thus, $R = \frac{1}{\sqrt{\pi}}L$. We show the multiplicative Jensen's gap of Eq. 7.40 in Fig. 7.21 (orange curve).

When the true density field varies over scales much smaller than a voxel side length L, such as in clusters, the ratio between the mean of the squared density and the square of the mean density can be much larger than 1. In particular, for a Navarro–Frenk–White (NFW) profile (Navarro et al., 1996) with scale radius r_s,

$$\frac{\langle\rho^2\rangle}{\langle\rho\rangle^2} = \frac{1-\left(\frac{r_s}{r_s+R}\right)^3}{9\left(\frac{r_s}{R}\right)^3\left(\ln\frac{r_s}{r_s+R}+1-\frac{r_s}{r_s+R}\right)^2}. \tag{7.41}$$

where we consider a sphere with radius R, where R is such that the volume of the sphere is equal to the volume of the voxel: $R = \left(\frac{3}{4\pi}\right)^{\frac{1}{3}} L$. We show the multiplicative Jensen's gap of Eq. 7.41 in Fig. 7.21 (red curve). In the limit $\frac{r_s}{R} \to \infty$, $\frac{\langle\rho^2\rangle}{\langle\rho\rangle^2} \to \frac{4}{3}$.

The history of astronomy is a history of receding horizons.
Edwin P. Hubble, American astronomer, *The Realm of the Nebulae (1936)*

8

Constraining the giant radio galaxy population with machine learning–accelerated detection and Bayesian inference

R. I. J. Mostert, **M. S. S. L. Oei**, B. Barkus, L. Alegre, M. J. Hardcastle, K. J. Duncan, H. J. A. Röttgering, R. J. van Weeren, M. Horton — *Astronomy & Astrophysics*, submitted

Abstract

Context Large-scale sky surveys at low frequencies, like the LOFAR Two-metre Sky Survey (LoTSS), allow for the detection and characterisation of unprecedented numbers of giant radio galaxies (GRGs, or 'giants'). This, in turn, enables us to study giants in a cosmological context. A tantalising prospect of such studies is a measurement of the contribution of giants to cosmic magnetogenesis. However, finding large GRG samples requires the creation of radio–optical catalogues for well-resolved radio sources and a suitable statistical framework to infer intrinsic GRG population properties.

Aims By automating the creation of radio–optical catalogues, we aim to expand significantly the census of known giants. With the resulting sample and a forward model mindful of selection effects, we aim to constrain their intrinsic length distri-

bution, number density, and lobe volume-filling fraction (VFF) in the Cosmic Web.

Methods We combine five existing codes into a single machine learning–driven (ML) pipeline that automates radio source component association and optical host identification for well-resolved radio sources. We create a radio–optical catalogue for the entire LoTSS Data Release 2 (DR2) footprint and subsequently select all sources that qualify as possible giants. We combine the list of ML pipeline GRG candidates with an existing list of LoTSS DR2 crowd-sourced GRG candidates and visually confirm or reject all members of the merged sample. To infer intrinsic GRG properties from GRG observations, we develop further a population-based forward model and constrain its parameters using Bayesian inference.

Results Roughly half of all radio sources that our ML pipeline identifies as giants (of at least $l_{\mathrm{p,GRG}} := 0.7$ Mpc long) indeed turn out to be upon visual inspection, whereas the success rate is one in eleven for the previous best giant-finding ML technique in the literature. We confirm $5,596$ previously unknown giants from the crowd-sourced LoTSS DR2 catalogue and $2,592$ previously unknown giants from the ML pipeline. Our confirmations and discoveries bring the total number of known giants to at least $11,524$. Our forward model for the intrinsic GRG population is able to provide a good fit to the data. Our posterior indicates that the projected lengths of giants are consistent with a curved power law probability density function whose initial tail index $\xi(l_{\mathrm{p,GRG}}) = -2.8 \pm 0.2$ changes by $\Delta\xi = -2.4 \pm 0.3$ over the interval up to $l_{\mathrm{p}} = 5$ Mpc. We predict a comoving GRG number density $n_{\mathrm{GRG}} = 13 \pm 10\ (100\ \mathrm{Mpc})^{-3}$, close to a current estimate of the number density of luminous non-giant radio galaxies. With the projected length distribution, number density, and additional assumptions, we derive a current-day GRG lobe VFF $\mathcal{V}_{\mathrm{GRG-CW}}(z=0) = 1.1 \pm 0.9 \cdot 10^{-5}$ in clusters and filaments of the Cosmic Web.

Conclusions We have created a state-of-the-art ML-accelerated pipeline for finding giants, whose complex morphologies, arcminute extents, and radio-emitting surroundings pose challenges. Our data analysis suggests that giants are more common than previously thought. More work is needed to make estimates of the GRG lobe VFF reliable, but the first results indicate that it is possible that magnetic fields originating from giants permeate significant ($\sim$10%) fractions of today's Cosmic Web.

Key words: Surveys – Methods: data analysis – Catalogues – Galaxies: active – Radio continuum: galaxies

8.1 Introduction

Recent radio Stokes-I imaging and rotation measure observations show that filaments of the Cosmic Web are magnetised (e.g. Govoni et al., 2019; de Jong et al., 2022; Carretti et al., 2023) with $B \sim 10^0$–10^2 nG (e.g. Vazza et al., 2021a). However, the origin of these magnetic fields remains highly uncertain. In a primordial magnetogenesis scenario (e.g. Subramanian, 2016), the seeds of intergalactic magnetic fields can be traced to the Early Universe. This scenario is not problem-free: primordial magnetic fields that arise before the end of inflation are typically too weak to match observations, while fields that arise after inflation (but before recombination) typically have coherence lengths that are too small. Alternatively, in an astrophysical magnetogenesis scenario, the seeds of intergalactic magnetic fields are predominantly spread by energetic astrophysical phenomena in the more recent Universe, such as radio galaxies (RGs) and supernova explosion driven winds (e.g. Vazza et al., 2017). In this latter scenario, giant radio galaxies (GRGs, or 'giants') may play a significant role in the magnetisation of the intergalactic medium (IGM), as their associated jets can carry magnetic fields of strength $B \sim 10^2$ nG from host galaxies to cosmological, megaparsec-scale distances (e.g. Oei et al., 2022a).

Efforts to measure the contribution of giants to astrophysical magnetogenesis in filaments of the Cosmic Web have only recently begun, with the advent of systematically processed, sensitive, low-frequency sky surveys such as the Low Frequency Array (LOFAR; van Haarlem et al., 2013) Two-metre Sky Survey (LoTSS; Shimwell et al., 2017). By carrying out a manual search for giants in LoTSS DR2 (Shimwell et al., 2022) pipeline products and a subsequent statistical analysis, Oei et al. (2023a) inferred a key statistic: the volume-filling fraction (VFF) of GRG lobes within clusters and filaments of the Local Universe, $\mathcal{V}_{\mathrm{GRG-CW}}(z = 0)$. However, considerable uncertainty remains as to its precise value, which requires inference of both the intrinsic GRG length distribution and the intrinsic GRG number density, as well as information about the typical shape of GRG lobes.

As the number of observed radio galaxies rapidly increases with decreasing angular length, the time taken, as part of the manual process of associating radio source components and identifying optical host galaxies, logically increases. Machine learning (ML)–based techniques have the potential to massively accelerate the detection of specific radio sources, to complement or eventually replace manual searches (e.g. Proctor, 2016; Gheller et al., 2018; Lochner & Bassett, 2021; Mostert et al., 2023). The potential for detecting GRGs was demonstrated by Dabhade et al. (2020a), who visually inspected the 1,600 ML-predicted GRG candidates of Proctor (2016) and thereby discovered 151 giants. By combining multiple ML-based and rule-based algo-

rithms that automate both the radio component association process and the optical host identification process into a single pipeline, we aim to improve upon the 9% precision achieved by the ML-predictions of Proctor (2016).

In the current work, we construct a LoTSS DR2 GRG sample of unparalleled size, by combining results from a manual visual search (Oei et al., 2023a), a citizen science–based visual search (Hardcastle et al., 2023), and a machine learning–accelerated search (this article; Sect. 8.4). With a definitive LoTSS DR2 GRG sample in hand, we refine the Bayesian forward model presented in Oei et al. (2023a), and finally constrain several key geometric quantities pertaining to giants.

In Sect. 8.2, we briefly recap, generalise, and enrich the statistical GRG geometry theory of Oei et al. (2023a). In Sect. 8.3, we introduce the LoTSS DR2 data in which we search for giants. In Sect. 8.4, we describe the methods that we use to build our definite LoTSS DR2 GRG sample, and explain how we use the theory of Sect. 8.2 in practice to infer GRG quantities of interest. In Sect. 8.5, we present our findings regarding the projected proper length distribution for giants, their comoving number density, and their instantaneous lobe volume-filling fraction (VFF) in clusters and filaments of the Cosmic Web. In Sect. 8.6, we discuss caveats of the current work, compare our results with previous results, and propose promising directions for future work, before we conclude in Sect. 8.7.

We assume a flat, inflationary ΛCDM model with parameters adopted from Planck Collaboration et al. (2020); i.e. $h = 0.6766$, $\Omega_{\mathrm{BM},0} = 0.0490$, $\Omega_{\mathrm{M},0} = 0.3111$, $\Omega_{\Lambda,0} = 0.6889$, where $H_0 := h \cdot 100\ \mathrm{km\ s^{-1}\ Mpc^{-1}}$. We define giants as radio galaxies with a projected proper[1] length $l_{\mathrm{p}} \geq l_{\mathrm{p,GRG}} := 0.7$ Mpc. We define the spectral index α such that it relates to flux density F_ν at frequency ν as $F_\nu \propto \nu^\alpha$; under this convention, most radio spectral indices are negative.

8.2 Theory

To infer the intrinsic length distribution, number density, and lobe volume-filling fraction of giants, we use a Bayesian forward modelling approach that incorporates selection effects. We adopt the framework described in Oei et al. (2023a), but generalise a few key formulae. Furthermore, in a change that allows for the extraction of tighter parameter constraints from the data, we now predict joint projected proper length–redshift histograms rather than projected proper length distributions.

[1] In Cosmic Web filament environments, where giants appear most common (Oei et al., in prep.), lobes may expand along the Hubble flow, rendering their proper and comoving extents different. To avoid ambiguity, we stress that our projected lengths are proper instead of comoving. A less precise synonym for 'projected proper length' often found in the literature is 'largest linear size' (LLS).

8.2.1 RG total and projected proper lengths

The central geometric quantity predicted by models of radio galaxy evolution (e.g. Turner & Shabala, 2015; Hardcastle, 2018) is, simply, the RG's intrinsic proper length l. Once the probability distribution of the intrinsic proper length random variable (RV) L is known, one can estimate other geometric quantities of interest, such as the VFF of RG lobes in the Cosmic Web. However, for the vast majority of observed RGs only a projected proper length l_p is available, as accurate measurements of jet inclination angles θ are currently challenging. In order to fit statistical models to data from surveys such as LoTSS DR2, models should therefore predict the distribution of the projected proper length RV L_p.

8.2.2 GRG projected proper length: general

We now show, first without adopting a specific parametric form for the distribution of L, how the cumulative density function (CDF) and probability density function (PDF) of the GRG projected proper length RV $L_\mathrm{p} \mid L_\mathrm{p} \geq l_\mathrm{p,GRG}$ can be calculated. In particular, suppose that L has support from some length $l_\mathrm{min} \geq 0$ onwards. It holds that $L_\mathrm{p} = L \sin \Theta$, where Θ is the inclination angle RV. Assuming that — at least on cosmological scales — all RG orientations in three dimensions are equally likely, the CDF of L_p relates to the PDF of L via

$$F_{L_\mathrm{p}}(l_\mathrm{p}) = \begin{cases} 0 & \text{if } l_\mathrm{p} \leq 0; \\ 1 - \int_{l_\mathrm{min}}^{\infty} \sqrt{1 - \left(\frac{l_\mathrm{p}}{l}\right)^2} f_L(l)\,\mathrm{d}l & \text{if } 0 < l_\mathrm{p} \leq l_\mathrm{min}; \\ 1 - \int_{l_\mathrm{p}}^{\infty} \sqrt{1 - \left(\frac{l_\mathrm{p}}{l}\right)^2} f_L(l)\,\mathrm{d}l & \text{if } l_\mathrm{p} > l_\mathrm{min}. \end{cases} \tag{8.1}$$

We note that, in the usual scenario $l_\mathrm{min} = 0$, the second case disappears. Equation 8.1 generalises Eq. A.8 from Oei et al. (2023a); its derivation closely follows the one presented there.

The CDF of the GRG projected proper length RV $L_\mathrm{p} \mid L_\mathrm{p} \geq l_\mathrm{p,GRG}$ is

$$F_{L_\mathrm{p} \mid L_\mathrm{p} \geq l_\mathrm{p,GRG}}(l_\mathrm{p}) = \begin{cases} 0 & \text{if } l_\mathrm{p} < l_\mathrm{p,GRG}; \\ 1 - \dfrac{\int_{l_\mathrm{p}}^{\infty} \sqrt{1 - \left(\frac{l_\mathrm{p}}{l}\right)^2} f_L(l)\,\mathrm{d}l}{\int_{l_\mathrm{p,GRG}}^{\infty} \sqrt{1 - \left(\frac{l_\mathrm{p,GRG}}{l}\right)^2} f_L(l)\,\mathrm{d}l} & \text{if } l_\mathrm{p} \geq l_\mathrm{p,GRG}. \end{cases} \tag{8.2}$$

This result follows from combining Eq. 8.1 and Eq. A.12 from Oei et al. (2023a).[2] As PDFs follow from CDFs by differentiation, we find that the PDFs of L_p and $L_\mathrm{p} \mid L_\mathrm{p} \geq l_\mathrm{p,GRG}$ relate as

$$f_{L_\mathrm{p} \mid L_\mathrm{p} \geq l_\mathrm{p,GRG}}(l_\mathrm{p}) = \begin{cases} 0 & \text{if } l_\mathrm{p} < l_\mathrm{p,GRG}; \\ \dfrac{f_{L_\mathrm{p}}(l_\mathrm{p})}{\int_{l_\mathrm{p,GRG}}^{\infty} \sqrt{1-\left(\frac{l_\mathrm{p,GRG}}{l}\right)^2} f_L(l)\,\mathrm{d}l} & \text{if } l_\mathrm{p} \geq l_\mathrm{p,GRG}. \end{cases} \tag{8.3}$$

We note that, throughout the support of $L_\mathrm{p} \mid L_\mathrm{p} \geq l_\mathrm{p,GRG}$, $f_{L_\mathrm{p} \mid L_\mathrm{p} \geq l_\mathrm{p,GRG}}(l_\mathrm{p})$ and $f_{L_\mathrm{p}}(l_\mathrm{p})$ are directly proportional — the quantity in the denominator of Eq. 8.3 is merely a normalisation constant.[3]

To find $f_{L_\mathrm{p}}(l_\mathrm{p})$ if $l_\mathrm{p} > l_\mathrm{min}$, it can be helpful to perform a change of variables. By defining $\eta := \frac{l}{l_\mathrm{p}}$, we rewrite

$$F_{L_\mathrm{p}}(l_\mathrm{p}) = 1 - l_\mathrm{p} \int_1^{\infty} \sqrt{1 - \frac{1}{\eta^2}} f_L(l_\mathrm{p}\eta)\,\mathrm{d}\eta \quad \text{if } l_\mathrm{p} > l_\mathrm{min}. \tag{8.5}$$

This form has the advantage that — within the integral — l_p occurs only in the integrand, whereas the form of Eq. 8.1 features l_p in both the integrand and in the lower integration limit. By differentation,

$$\begin{aligned} f_{L_\mathrm{p}}(l_\mathrm{p}) = &- \int_1^{\infty} \sqrt{1 - \frac{1}{\eta^2}} f_L(l_\mathrm{p}\eta)\,\mathrm{d}\eta \\ &- l_\mathrm{p} \int_1^{\infty} \sqrt{1 - \frac{1}{\eta^2}} \frac{\mathrm{d}f_L(l_\mathrm{p}\eta)}{\mathrm{d}l_\mathrm{p}}\,\mathrm{d}\eta \quad \text{if } l_\mathrm{p} > l_\mathrm{min}. \end{aligned} \tag{8.6}$$

To arrive at concrete expressions for the GRG projected proper length PDF of Eq. 8.3, we must choose a specific parametric form for the distribution of L or L_p.

8.2.3 GRG projected proper length: curved power law

Oei et al. (2023a) have shown that models that assume a Paretian tail for the RG in-

[2] We have also assumed that $l_\mathrm{p,GRG} > l_\mathrm{min}$, which is the obvious case to consider.

[3] This is an example of a more general rule: for any RV X,

$$f_{X \mid X \geq y}(x) = \begin{cases} 0 & \text{if } x < y; \\ \frac{f_X(x)}{1-F_X(y)} & \text{if } x \geq y. \end{cases} \tag{8.4}$$

trinsic proper length distribution, and that include angular and surface brightness selection effects, can tightly reproduce the observed GRG projected proper length distribution. The PDF of a Pareto-distributed RV is a simple power law, which is fully specified by a lower cut-off l_{min} and a tail index ξ. However, there is a good reason to believe that the true GRG projected proper length PDF deviates from simple power law behaviour. The true RG projected proper length PDF $f_{L_{\mathrm{p}}}$ will peak around a value set by the typical jet power, environment, lifetime, and inclination angle (amongst other properties). Below this value, $f_{L_{\mathrm{p}}}$ will necessarily be an increasing function of l_{p}; above this value, $f_{L_{\mathrm{p}}}$ will be a decreasing function.[4] As giants embody the large-length tail of the distribution of L_{p}, it is likely that the slope of $f_{L_{\mathrm{p}} \mid L_{\mathrm{p}} \geq l_{\mathrm{p,GRG}}}(l_{\mathrm{p}})$ at least somewhat decreases (i.e. steepens) as l_{p} increases — even in log–log space.

To remain close to the seemingly effective Pareto assumption of Oei et al. (2023a), we assume in this work that, at least for $l_{\mathrm{p}} \geq l_{\mathrm{p,GRG}}$, the RG projected proper length PDF is a curved power law:

$$f_{L_{\mathrm{p}}}(l_{\mathrm{p}}) \propto \left(\frac{l_{\mathrm{p}}}{l_{\mathrm{p,GRG}}}\right)^{\xi(l_{\mathrm{p}})} \quad \text{if } l_{\mathrm{p}} \geq l_{\mathrm{p,GRG}}, \tag{8.7}$$

where the exponent

$$\xi(l_{\mathrm{p}}) := \xi(l_{\mathrm{p,1}}) + \frac{l_{\mathrm{p}} - l_{\mathrm{p,1}}}{l_{\mathrm{p,2}} - l_{\mathrm{p,1}}} \left(\xi(l_{\mathrm{p,2}}) - \xi(l_{\mathrm{p,1}})\right) \tag{8.8}$$

is a linear function of l_{p}. As long as $l_{\mathrm{p,1}} \neq l_{\mathrm{p,2}}$, both projected proper length constants can be chosen arbitrarily; however, $l_{\mathrm{p,1}} := l_{\mathrm{p,GRG}}$ seems to be a natural choice. Adopting this choice, and defining $\Delta\xi := \xi(l_{\mathrm{p,2}}) - \xi(l_{\mathrm{p,1}})$, leads to the final exponent formula

$$\xi(l_{\mathrm{p}}) = \xi(l_{\mathrm{p,GRG}}) + \frac{l_{\mathrm{p}} - l_{\mathrm{p,GRG}}}{l_{\mathrm{p,2}} - l_{\mathrm{p,GRG}}} \Delta\xi. \tag{8.9}$$

We adopt $\xi(l_{\mathrm{p,GRG}})$ and $\Delta\xi$ as two parameters of our model. We furthermore choose $l_{\mathrm{p,2}} := 5$ Mpc, which is close to the largest currently known radio galaxy projected proper length (Oei et al., 2022a, 2023a). Being the first-order Taylor polynomial of an arbitrary function $\xi(l_{\mathrm{p}})$ at $l_{\mathrm{p,GRG}}$, Eq. 8.8 represents a natural generalisation of the constant tail index assumption of Oei et al. (2023a). In particular, if model parameter $\Delta\xi = 0$, we recover the earlier Paretian model.

By the same reasoning as before, we find that if the RG projected proper length

[4] This line of reasoning implicitly assumes that the distribution of L_{p} is unimodal.

PDF is a curved power law for $l_\mathrm{p} \geq l_\mathrm{p,GRG}$, then the GRG projected proper length PDF is also a curved power law over this range:

$$f_{L_\mathrm{p} \mid L_\mathrm{p} \geq l_\mathrm{p,GRG}}(l_\mathrm{p}) \propto \left(\frac{l_\mathrm{p}}{l_\mathrm{p,GRG}}\right)^{\xi(l_\mathrm{p})} \quad \text{if } l_\mathrm{p} \geq l_\mathrm{p,GRG}. \tag{8.10}$$

The factors required to normalise $f_{L_\mathrm{p}}(l_\mathrm{p})$ and $f_{L_\mathrm{p} \mid L_\mathrm{p} \geq l_\mathrm{p,GRG}}(l_\mathrm{p})$ can be obtained numerically.

Whereas Oei et al. (2023a) parametrised $f_L(l)$ and derived $f_{L_\mathrm{p}}(l_\mathrm{p})$ and $f_{L_\mathrm{p} \mid L_\mathrm{p} \geq l_\mathrm{p,GRG}}(l_\mathrm{p})$, we now parametrise $f_{L_\mathrm{p}}(l_\mathrm{p})$ and derive only $f_{L_\mathrm{p} \mid L_\mathrm{p} \geq l_\mathrm{p,GRG}}(l_\mathrm{p})$. It is possible to start modelling at the level of $f_L(l)$, also in the context of curved power law PDFs, but the resulting expressions for $f_{L_\mathrm{p}}(l_\mathrm{p})$ and $f_{L_\mathrm{p} \mid L_\mathrm{p} \geq l_\mathrm{p,GRG}}(l_\mathrm{p})$ become tedious and rather uninsightful. For simplicity, we therefore choose to parametrise $f_{L_\mathrm{p}}(l_\mathrm{p})$; we explore the alternative set-up in Appendix 8.A1.

8.2.4 GRG observed projected proper length

Equation 8.10 describes a distribution of GRG projected proper lengths in the absence of observational selection effects. Unfortunately, this distribution cannot be directly tested against GRG samples obtained from surveys, which are always affected by selection. For a thorough description and derivation of selection effect modelling in the context of our framework, we refer the reader to Sect. 2.8 and Appendix A.8 of Oei et al. (2023a); here, we shall only briefly introduce the expressions that we require.

A key result, adopted from Eq. 21 of Oei et al. (2023a), is that the GRG observed projected proper length RV $L_\mathrm{p,obs} \mid L_\mathrm{p,obs} \geq l_\mathrm{p,GRG}$ can be expressed as

$$f_{L_\mathrm{p,obs} \mid L_\mathrm{p,obs} \geq l_\mathrm{p,GRG}}(l_\mathrm{p}) = \begin{cases} 0 & \text{if } l_\mathrm{p} < l_\mathrm{p,GRG}; \\ \dfrac{C(l_\mathrm{p}) f_{L_\mathrm{p}}(l_\mathrm{p})}{\int_{l_\mathrm{p,GRG}}^{\infty} C(l''_\mathrm{p}) f_{L_\mathrm{p}}(l''_\mathrm{p})\, \mathrm{d}l''_\mathrm{p}} & \text{if } l_\mathrm{p} \geq l_\mathrm{p,GRG}, \end{cases} \tag{8.11}$$

where $C(l_\mathrm{p}) = C(l_\mathrm{p}, z_\mathrm{max})$ is the completeness function. More precisely, $C(l_\mathrm{p}, z_\mathrm{max})$ denotes the fraction of all RGs with projected proper length l_p in the volume up to cosmological redshift z_max that is detected and identified through the survey considered — in this work, this will be LoTSS DR2. The repeated factors in numerator and denominator reveal that, in order to compute $f_{L_\mathrm{p,obs} \mid L_\mathrm{p,obs} \geq l_\mathrm{p,GRG}}(l_\mathrm{p})$, we need to know $f_{L_\mathrm{p}}(l_\mathrm{p})$ for $l_\mathrm{p} \geq l_\mathrm{p,GRG}$ only — and within this range up to a constant only. More concerningly, we also see that selection effects that reduce the completeness by the same factor for all $l_\mathrm{p} \geq l_\mathrm{p,GRG}$ leave no imprint on $f_{L_\mathrm{p,obs} \mid L_\mathrm{p,obs} \geq l_\mathrm{p,GRG}}(l_\mathrm{p})$. Therefore, such selection effects cannot be constrained by a GRG observed projected proper length

analysis alone.

Under the assumption that the RG projected proper length PDF $f_{L_\mathrm{p}}(l_\mathrm{p})$ does not evolve between redshifts $z = z_\mathrm{max}$ and $z = 0$, the completeness function becomes

$$C\left(l_\mathrm{p}, z_\mathrm{max}\right) = \frac{\int_0^{z_\mathrm{max}} p_\mathrm{obs}\left(l_\mathrm{p}, z\right) r^2\left(z\right) E^{-1}\left(z\right)\ \mathrm{d}z}{\int_0^{z_\mathrm{max}} r^2\left(z\right) E^{-1}\left(z\right)\ \mathrm{d}z}, \tag{8.12}$$

where the observing probability $p_\mathrm{obs}(l_\mathrm{p}, z)$ is the probability that an RG of projected proper length l_p at redshift z is detected by a survey and its subsequent analysis steps (such as the machine learning pipeline considered in this work), r denotes comoving radial distance, and $E(z)$ is the dimensionless Hubble parameter[5]. The appropriate form of $p_\mathrm{obs}(l_\mathrm{p}, z)$ is determined by the selection effects relevant to the survey of interest and its analysis.

In this work, we will consider GRG lobe surface brightness selection, which at present renders part of the GRG population inherently undetectable (on an individual level, at least), and selection by limitations of our analysis steps, which causes in principle detectable giants to evade sample inclusion. We describe the former effect parametrically, and determine the latter effect empirically. The effects yield functions $p_\mathrm{obs,SB}(l_\mathrm{p}, z)$ and $p_\mathrm{obs,ID}(l_\mathrm{p}, z)$, respectively, which then combine to form a single observing probability function through

$$p_\mathrm{obs}(l_\mathrm{p}, z) = p_\mathrm{obs,SB}(l_\mathrm{p}, z) \cdot p_\mathrm{obs,ID}(l_\mathrm{p}, z). \tag{8.14}$$

Selection effects: surface brightness limit

RG lobes whose surface brightnesses are lower than some threshold value $b_{\nu,\mathrm{th}}$, which typically equals the survey noise level σ times a low factor of order unity, cannot be detected. Following Sect. 2.8.3 of Oei et al. (2023a), we model surface brightness (SB) selection by assuming that the lobe surface brightnesses $B_\nu(\nu, l, z)$ at $\nu = \nu_\mathrm{obs}$ of radio galaxies of intrinsic proper length $l = l_\mathrm{ref}$ that reside at redshift $z = 0$ are lognormally distributed. More precisely, we parametrise $B_\nu(\nu_\mathrm{obs}, l_\mathrm{ref}, 0) = b_{\nu,\mathrm{ref}}\, S$, where $b_{\nu,\mathrm{ref}}$ is the median SB, and S is a lognormally distributed RV with median 1 and dispersion

[5] In a *flat* FLRW universe, the dimensionless Hubble parameter E is

$$E\left(z\right) := \frac{H\left(z\right)}{H_0} = \sqrt{\Omega_{\mathrm{R},0}\left(1+z\right)^4 + \Omega_{\mathrm{M},0}\left(1+z\right)^3 + \Omega_{\Lambda,0}}. \tag{8.13}$$

parameter σ_{ref}. The observing probability due to SB selection then is

$$p_{\mathrm{obs,SB}}\left(l_{\mathrm{p}}, z\right) = \int_{s_{\mathrm{min}}}^{\infty} \sqrt{1 - \left(\frac{s_{\mathrm{min}}}{s}\right)^{-\frac{2}{\zeta}}} f_S\left(s\right) \mathrm{d}s; \tag{8.15}$$

$$s_{\mathrm{min}} = \frac{b_{\nu,\mathrm{th}}}{b_{\nu,\mathrm{ref}}} \left(\frac{l_{\mathrm{p}}}{l_{\mathrm{ref}}}\right)^{-\zeta} (1+z)^{3-\alpha}; \tag{8.16}$$

$$f_S\left(s\right) = \frac{1}{\sqrt{2\pi}\sigma_{\mathrm{ref}} s} \exp\left(-\frac{\ln^2 s}{2\sigma_{\mathrm{ref}}^2}\right). \tag{8.17}$$

Here, α is the typical RG lobe spectral index, which we will assume fixed at $\alpha = -1$. The exponent ζ determines how the SB distribution scales with projected proper length l_{p}.

In contrast to the choice made in Oei et al. (2023a), we do not fix $\zeta = -2$, but rather leave ζ a free parameter which we fit to the data. Deviations from $\zeta = -2$ occur in at least two cases: when giant growth is not shape-preserving, and if the radio luminosity distributions of giants of different l_{p} are distinct. Dynamical models of radio galaxies in general predict that both cocoons (e.g. Fig. 4 of Turner & Shabala, 2015) and lobes (e.g. Fig. 9 of Hardcastle, 2018) change shape over time, with a dependence on jet power. There remains considerable uncertainty as to how shapes change throughout the giant phase: axial ratio–like measures generally show that RG lobes become more elongated during growth, but this trend could possibly reverse for giants, whose lobes might protrude from the clusters and filaments in which they are born. Simulations suggest that, for such protrusions, the usual constant power law profile assumptions for the ambient baryon density and temperature break down (e.g. Fig. 8 of Gheller & Vazza, 2019). If lobes of giants widen over time, then ζ would decrease. The second case occurs if the end-of-life lengths of radio galaxies increase with jet power, so that the subpopulation that survives up to some l_{p} has its jet power distribution — and thus its radio luminosity distribution — shifted upwards with respect to subpopulations at smaller l_{p}. This effect, which appears plausible given models (e.g. Fig. 8 of Hardcastle, 2018), would increase ζ. At present, it seems hard to predict the net result on ζ of these counteracting effects.

Selection effects: nonidentification

Every current-day survey search method (such as visual inspection by scientists, visual inspection by citizen scientists, and machine learning–based approaches) will fail to identify some giants that are in principle identifiable (in the sense that they lie above the detection threshold set by the noise). For automated approaches, such

as the machine learning–based approach presented in this work, identification can become more challenging for larger angular lengths φ: one reason being the increased number of unrelated, interloping radio sources that cover the solid angle occupied by the RG. We call the probability that an identifiable RG is indeed identified — and thus makes it into the final sample — $p_{\mathrm{obs,ID}}(l_{\mathrm{p}}, z)$.

Say we have M methods to search for giants in the same survey. Let $\mathcal{G} = \{g_1, g_2, \ldots, g_N\}$ be the set of all identified giants (so that $|\mathcal{G}| = N$), and let $\mathcal{G}_i \subseteq \mathcal{G}$ be the subset identified by method i. Figure 8.1 provides an overview of the set-up. We take $l_{\mathrm{p}}(g)$ and $z(g)$ to mean the projected proper length and cosmological redshift of giant g. To determine the identification probability $p_{\mathrm{obs,ID},i}(l_{\mathrm{p}}, z)$ for method i, we first assume it to be of logistic form

$$p_{\mathrm{obs,ID},i}(l_{\mathrm{p}}, z) = \frac{1}{1 + \exp\left(-(\beta_{0,i} + \beta_{l_{\mathrm{p}},i} \cdot l_{\mathrm{p}} + \beta_{z,i} \cdot z)\right)}. \tag{8.18}$$

We obtain best-fit parameters $\hat{\beta}_{0,i}$, $\hat{\beta}_{l_{\mathrm{p}},i}$, and $\hat{\beta}_{z,i}$ by performing binary logistic regression with two explanatory variables on the set of pairs $\mathcal{D}_i$, where

$$\mathcal{D}_i := \left\{ \left(\left[l_{\mathrm{p}}(g), z(g) \right], \mathbb{I}(g \in \mathcal{G}_i) \right) \mid g \in \bigcup_{j=1, j \neq i}^{M} \mathcal{G}_j \right\}. \tag{8.19}$$

The first element of each pair is a point in projected length–redshift space, whilst the second element is 0 or 1: $\mathbb{I}$ denotes the indicator function. Qualitatively, $\mathcal{D}_i$ stores for each giant in the union of all GRG subsets except $\mathcal{G}_i$ its projected length–redshift coordinates, together with the success or failure of its identification by method i.

The implicit assumption here is that all $g \in \bigcup_{j=1, j \neq i}^{M} \mathcal{G}_j$ are typical examples of identifiable giants at the relevant projected proper length and redshift. We caution that this might not be true: giants with a peculiar morphology, or those lying in parts of the sky where optical identification is hard (e.g. towards the Galactic Plane or crowded regions of large-scale structure), may be identifiable in a radio surface brightness sense, but will nonetheless evade sample inclusion more often than other giants. As a result, giants that *do* end up in a sample — such as $\bigcup_{j=1, j \neq i}^{M} \mathcal{G}_j$ — will have more regular morphologies than giants in general and will lie in regions of the sky where optical identification is easier than for giants in general. Typically, such giants are also more likely to be found by method i, and as a result our approach will probably render $p_{\mathrm{obs,ID},i}$ biased high.

Given a set of M functions $\{p_{\mathrm{obs,ID},i}(l_{\mathrm{p}}, z) \mid i \in \{1, 2, \ldots, M\}\}$, several possibilities exist to combine them into a single $p_{\mathrm{obs,ID}}(l_{\mathrm{p}}, z)$. At the minimum, $p_{\mathrm{obs,ID}}(l_{\mathrm{p}}, z)$ is

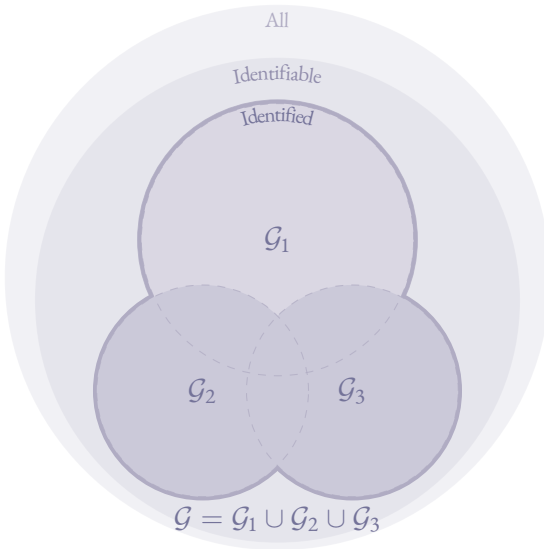


Figure 8.1: Schematic of a three-method search for giants. Of all giants in the survey footprint up to $z = z_{\max}$, only those for which the lobe surface brightness at the observing frequency $\nu_{\rm obs}$ is above detection threshold $b_{\nu,\rm th}$ are identifiable. $\mathcal{G}$ denotes the actually identified set of giants. $\mathcal{G}_1$, $\mathcal{G}_2$, and $\mathcal{G}_3$ are the subsets identified by each method individually. As an example, we shade $\mathcal{G}_2 \cup \mathcal{G}_3$, which has overlap with $\mathcal{G}_1$, and which can be used to measure $p_{\rm obs,ID,1}(l_{\rm p}, z)$.

given by a point-wise maximum:

$$p_{\rm obs,ID}(l_{\rm p}, z) = \max_{i \in \{1,2,\ldots,M\}} p_{\rm obs,ID,i}(l_{\rm p}, z), \tag{8.20}$$

which is appropriate if methods tend to find the same identifiable giants — as in our case.[6]

8.2.5 GRG number density

The preceding theory allows us to find the intrinsic, comoving number density of giants, $n_{\rm GRG}$, if we know the observed number of giants within a solid angle of extent Ω and in the volume up to $z_{\max}$, $N_{\rm GRG,obs}(\Omega, z_{\max})$. We assume that, up to this redshift, $n_{\rm GRG}$ remains constant. We note that we cannot calculate $n_{\rm GRG}$ using Eq. 30 from

[6] In case methods tend to find independent subsets of identifiable giants,

$$p_{\rm obs,ID}(l_{\rm p}, z) = 1 - \prod_{i=1}^{M} (1 - p_{\rm obs,ID,i}(l_{\rm p}, z)). \tag{8.21}$$

We note that it is possible to design methods that find subsets of identifiable giants that have even less overlap than independent subsets have.

Oei et al. (2023a): this equation assumes $\xi(l_{\mathrm{p},1}) = \xi(l_{\mathrm{p},2})$. We derive a more general expression by first noting that the number of giants observed within a solid angle of extent Ω in the volume up to $z_{\max}$ and with projected proper lengths between l_{p} and $l_{\mathrm{p}} + \mathrm{d}l_{\mathrm{p}}$ is

$$\mathrm{d}N_{\mathrm{GRG,obs}}(l_{\mathrm{p}}, \Omega, z_{\max}) = \frac{\Omega}{4\pi}\, n_{\mathrm{GRG}} f_{L_{\mathrm{p}} \mid L_{\mathrm{p}} \geq l_{\mathrm{p,GRG}}}(l_{\mathrm{p}})\, \mathrm{d}l_{\mathrm{p}} \cdot \int_0^{z_{\max}} p_{\mathrm{obs}}(l_{\mathrm{p}}, z)\, 4\pi r^2(z)\, \frac{\mathrm{d}r}{\mathrm{d}z}\, \mathrm{d}z. \tag{8.22}$$

Because

$$N_{\mathrm{GRG,obs}}(\Omega, z_{\max}) = \int_{l_{\mathrm{p,GRG}}}^{\infty} \mathrm{d}N_{\mathrm{GRG,obs}}(l_{\mathrm{p}}, \Omega, z_{\max}), \tag{8.23}$$

we find — by isolating n_{GRG} — that

$$n_{\mathrm{GRG}}(l_{\mathrm{p,GRG}}, z_{\max}) = \frac{H_0}{c}\, \frac{4\pi}{\Omega}\, N_{\mathrm{GRG,obs}}(\Omega, z_{\max}) \cdot \left(\int_{l_{\mathrm{p,GRG}}}^{\infty} f_{L_{\mathrm{p}} \mid L_{\mathrm{p}} \geq l_{\mathrm{p,GRG}}}(l_{\mathrm{p}}) \int_0^{z_{\max}} p_{\mathrm{obs}}(l_{\mathrm{p}}, z)\, 4\pi r^2(z)\, E^{-1}(z)\, \mathrm{d}z\, \mathrm{d}l_{\mathrm{p}} \right)^{-1}. \tag{8.24}$$

This expression is valid also beyond the context of power law or curved power law PDFs $f_{L_{\mathrm{p}} \mid L_{\mathrm{p}} \geq l_{\mathrm{p,GRG}}}(l_{\mathrm{p}})$. We remark that n_{GRG} can depend sensitively on the projected proper length used to define giants, $l_{\mathrm{p,GRG}}$.

In contrast to the approach of Oei et al. (2023a), in this work we do not calculate n_{GRG} in a step *following* inference of the framework's parameters, but rather include it as a parameter to be constrained *during* inference.

8.2.6 GRG lobe volume-filling fraction

To constrain the contribution of giants to astrophysical magnetogenesis, we wish to know the volume-filling fraction of their lobes in clusters and filaments of the Cosmic Web. Under the approximation that GRG shapes are independent of their volumes, we have

$$\mathcal{V}_{\mathrm{GRG-CW}}(z = 0) = \mathbb{E}[\Upsilon_{\mathrm{p}} \mid L_{\mathrm{p}} \geq l_{\mathrm{p,GRG}}] \cdot \mathbb{E}[L_{\mathrm{p}}^3 \mid L_{\mathrm{p}} \geq l_{\mathrm{p,GRG}}] \cdot n_{\mathrm{GRG}} \cdot (\mathcal{V}_{\mathrm{CW}}(z = 0))^{-1}, \tag{8.25}$$

where $\Upsilon_\mathrm{p} \mid L_\mathrm{p} \geq l_\mathrm{p,GRG}$ is a random variable denoting the ratio between a giant's combined lobe volume and its cubed projected proper length: $\Upsilon_\mathrm{p} \mid L_\mathrm{p} \geq l_\mathrm{p,GRG} := \frac{V}{L_\mathrm{p}^3} \mid L_\mathrm{p} \geq l_\mathrm{p,GRG}$. The distribution of $\Upsilon_\mathrm{p} \mid L_\mathrm{p} \geq l_\mathrm{p,GRG}$ can be inferred by fitting geometric models to GRG images, as has been explored in Oei et al. (2022a, 2023b). $\mathcal{V}_\mathrm{CW}(z = 0)$ denotes the volume-filling fraction of clusters and filaments in the Local Universe.

8.2.7 GRG angular lengths

An object's angular length φ, projected proper length l_p, and cosmological redshift z are related through

$$\varphi(l_\mathrm{p}, z) = \frac{l_\mathrm{p}(1+z)}{r(z)}. \tag{8.26}$$

Due to the expansion of the Universe, there exists a minimum angular length for objects of a given projected proper length. If one defines giants as radio galaxies with projected proper lengths $l_\mathrm{p} \geq l_\mathrm{p,GRG} := 0.7$ Mpc, as in this work, then all giants have an angular length $\varphi \geq 1.3'$ (Oei et al., 2023a). This fact has important consequences for GRG search campaigns. At the LoTSS resolution of $\theta_\mathrm{FWHM} = 6''$, it implies that giants are always resolved sources, spanning at least 13 resolution elements. Thus, to model the detectability of giants at this resolution, one must consider their surface brightness (profiles), rather than their flux densities.

8.2.8 Inference

Finally, we detail how the framework's six free parameters $\boldsymbol{\theta} := [\xi(l_\mathrm{p,GRG}), \Delta\xi, b_{\nu,\mathrm{ref}}, \sigma_\mathrm{ref}, \zeta, n_\mathrm{GRG}]$ can be inferred from a dataset containing a projected length and redshift for each observed giant. In particular, we consider a rectangle in projected proper length–cosmological redshift parameter space, within which our model assumptions are expected to hold. We partition this rectangle into N_b equiareal bins of width Δl_p and height Δz. We denote the coordinates of bin i's centre as $(l_{\mathrm{p},i}, z_i)$.

On the data side, we construct a two-dimensional histogram using these bins. The number of giants found in bin i, N_i, is a random variable with a Poisson distribution: $N_i \sim \mathrm{Poisson}(\lambda_i)$. Its expectation λ_i depends on the model parameters $\boldsymbol{\theta}$. Assuming that the $\{N_i\}$ are independent, the log-likelihood becomes

$$\ln \mathcal{L}(\{N_i\} \mid \boldsymbol{\theta}) = \sum_{i=1}^{N_\mathrm{b}} N_i \ln \lambda_i(\boldsymbol{\theta}) - \lambda_i(\boldsymbol{\theta}) - \ln(N_i!). \tag{8.27}$$

The last term on the right-hand side of Eq. 8.27 is the same for all θ, and need not be calculated if one is interested in $\mathcal{L}$ up to a global constant only.[7] Following Eq. 8.22, but avoiding integration over z and assuming narrow bins in both dimensions, we approximate

$$\lambda_i \approx n_{\mathrm{GRG}} V_i \cdot f_{L_{\mathrm{p}} \mid L_{\mathrm{p}} \geq l_{\mathrm{p,GRG}}}(l_{\mathrm{p},i}) \Delta l_{\mathrm{p}} \cdot p_{\mathrm{obs}}(l_{\mathrm{p},i}, z_i), \tag{8.29}$$

where the volume in which the giants of bin i fall, V_i, is

$$V_i = \Omega r^2(z_i) \Delta r_i, \text{ with } \Delta r_i = \frac{c}{H_0} \frac{\Delta z}{E(z_i)}. \tag{8.30}$$

Appendix 8.A2 details a particularly efficient trick to compute the likelihood for a range of n_{GRG}, whilst leaving the other parameters fixed. By multiplying the likelihood function with a prior distribution, for which we shall choose a uniform distribution, we obtain a posterior distribution over θ — up to a constant.

8.3 Data

We applied our automated radio–optical catalogue creation methods to all total intensity (Stokes-I) maps from LoTSS DR2 (Shimwell et al., 2022).[8] The observations in LoTSS DR2 cover the 120–168 MHz frequency range, have a 6″ resolution, a median RMS sensitivity of 83 μJy beam^{-1}, and a flux density scale uncertainty of approximately 10%. The observations are split into a region centred at 12h45m +44°30′ and a region centred at 1h00m +28°00′; both avoid the Galactic Plane. These regions span 4,178 and 1,457 square degrees respectively, and together cover 27% of the Northern Sky. The observations consist of 841 partly overlapping pointings with diameters of 4.0°. The vast majority of the pointings were observed for 8h, all within the 2014-05-23 to 2020-02-05 time frame.

Apart from the LoTSS DR2 Stokes-I maps, the ML radio catalogue pipeline that we describe in this manuscript (Sect. 8.4), relies on an infrared-optical source catalogue. This catalogue combines the positions, magnitudes and colour information of

[7] If one includes the term, it only needs to be calculated once. For numerical stability, it is helpful to note that

$$-\sum_{i=1}^{N_{\mathrm{b}}} \ln(N_i!) = -\sum_{i=1}^{N_{\mathrm{b}}} \sum_{j=2}^{N_i} \ln j. \tag{8.28}$$

[8] LoTSS DR2 is publicly available at `https://lofar-surveys.org/dr2_release.html`.

the infrared sources of the unWISE data release (Schlafly et al., 2019) from the Wide-field Infrared Survey Explorer (WISE; Wright et al., 2010), and the optical sources in the DESI Legacy Imaging Surveys DR9 (Dey et al., 2019).

We tried to maximise the identification of giants $p_{\mathrm{obs,ID}}$ within the LoTSS DR2 data. To do so, we complemented the sample of giants detected by our ML pipeline with all giants in the value-added LoTSS DR2 radio catalogue (Hardcastle et al., 2023). For large ($\varphi > 15''$) radio components (and thus most giants), the radio source component association and most of the optical host identifications of the value-added LoTSS DR2 radio catalogue, described by Hardcastle et al. (2023), were performed via a public project named 'LOFAR Galaxy Zoo' on Zooniverse. We will thus refer to the value-added LoTSS DR2 radio catalogue as the 'LGZ catalogue' and to the giants in that catalogue as the 'LGZ giants'. Zooniverse is an online citizen science platform for crowd-sourced visual inspection.[9] The detailed source component information provided by the LGZ catalogue allowed us to homogenise the angular length estimates of our ML pipeline giants and the LGZ giants (see Sect. 8.4.6). We further complemented our GRG sample with other GRG samples in the literature, see Sect. 8.4.8.

8.4 Methods

To derive the projected length distribution, number density, and lobe VFF for the intrinsic population of giants, we followed a two-stage approach. In the first stage, we gathered all giants that we detected in the LoTSS DR2 Stokes-I images using our automatic ML pipeline and added all other giants that we found in the LGZ catalogue. We re-evaluated and homogenised the source size estimates over the combined GRG sample, and manually inspected the plausibility of the associated radio source components and optical/infrared host galaxy. Finally, we merged this GRG sample with the other GRG samples from the literature. In the second stage, we search for the most likely parameters for the forward model presented in Sect. 8.2 that describe the GRG observed projected proper length distribution and the selection effects of the merged GRG sample. Figure 8.2 shows an overview of our approach.

8.4.1 Radio emission detection

We started out with the publicly available calibrated LoTSS DR2 Stokes-I images (Shimwell et al., 2022). For each of the 841 pointings, we ran the PyBDSF radio blob detection software (Mohan & Rafferty, 2015) using the same parameters as used in

[9] The Zooniverse website is `https://zooniverse.org`.

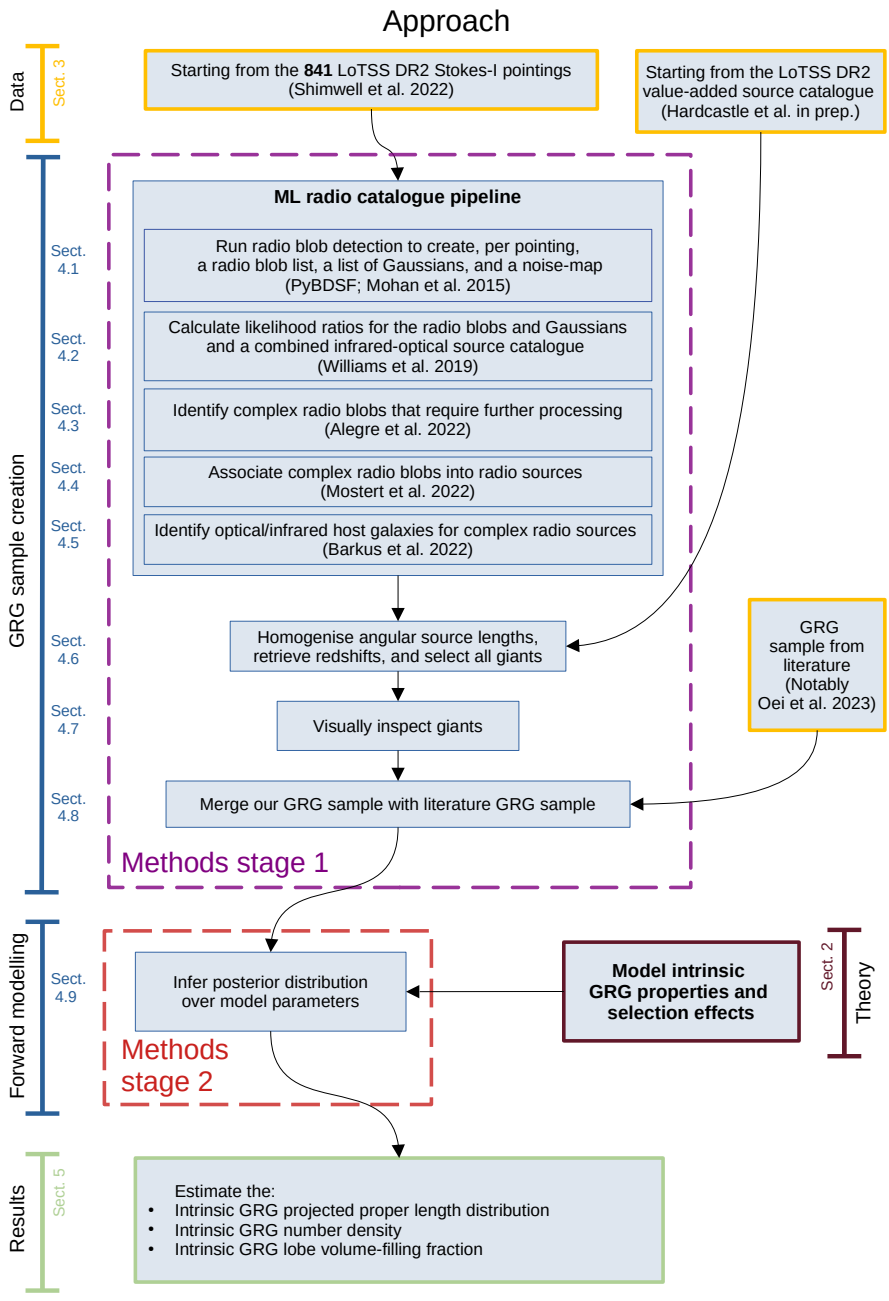


Figure 8.2: Overview of our approach, which consists of two stages. In the first stage we built a GRG sample, and in the second stage we inferred the properties of the intrinsic GRG population using a forward model. The brackets indicate the different parts of our approach and mention the sections containing the corresponding details.

LoTSS DR2 (Shimwell et al., 2022) — notably, this means that we used a 5σ detection threshold. Appendix 8.A3 provides the full list of PyBDSF parameters and their values.

The output we generated consists of a list of radio blobs with their location and properties. PyBDSF can decompose each radio blob it detects, into one or more 2D Gaussians. For each radio blob, we also saved the corresponding list of Gaussians. These Gaussians function as a source model for each radio blob and will be used in later steps in the ML pipeline.[10]

8.4.2 Calculating radio to optical/infrared likelihood ratios

For radio sources, the location of the host galaxy on the sky is close to the flux-weighted centre of the radio source.[11] The likelihood ratio method, which exploits this idea, quantifies the likelihood that a source in one spectral window is the correct counterpart to a source in another spectral window (e.g. Richter, 1975; de Ruiter et al., 1977; Sutherland & Saunders, 1992). Williams et al. (2019) used this method to cross-match the unresolved — and some resolved — radio sources of LoTSS DR1 to a combined catalogue of infrared and optical sources. More specifically, the infrared sources came from AllWISE (Cutri et al., 2021), whilst the optical sources came from the Panoramic Survey Telescope and Rapid Response System 1 (Pan-STARRS1; Chambers et al., 2016) DR1 3π steradian survey. The likelihood ratio function that Williams et al. (2019) used is a function of the angular distance between the flux-weighted centre of the radio source and the flux-weighted centre of the optical or infrared source, the magnitude of the optical or infrared source, and the colour of the optical or infrared source. The likelihood ratio function also takes into account uncertainties in each of these three dependencies.

We adopted the same procedure as detailed by Williams et al. (2019) to cross-match our simple radio sources (where 'simple' is to be understood as in Sect. 8.4.3) to a combined catalogue of infrared and optical sources. The infrared sources came from unWISE (Schlafly et al., 2019), and the optical sources were now taken from the DESI Legacy Imaging Surveys DR9 (Dey et al., 2019), which boasts deeper imagery than Pan-STARRS1 DR1 used for LoTSS DR1. The unWISE (Schlafly et al., 2019) and DESI Legacy Imaging Surveys DR9 source catalogues will be used for LoTSS DR2 cross-matching more generally (Hardcastle et al., 2023). Per pointing, we applied the likelihood ratio method to the full list of radio blobs and to the full list of Gaussians.

[10] However, as we discuss in Sect. 8.4.6 these source models are not always adequate for extended, well-resolved radio sources.

[11] In this context, 'close' refers to angular distances comparable to the survey's resolution.

For both the blobs and the Gaussians, we stored the identifier of the optical or infrared source that produced the highest likelihood ratio, alongside this highest likelihood ratio itself.

8.4.3 Sorting radio emission with a gradient boosting classifier

Most radio sources that consist of a single radio blob (mostly unresolved or barely resolved radio sources) can be cross-matched using the likelihood ratio method. However, some resolved radio sources, and certainly most resolved giant radio galaxies, consist of multiple radio blobs, and thus require radio blob association and cannot be cross-matched using the likelihood ratio alone. To separate the simple from the complex radio blobs in LoTSS DR1, a considerable amount of visual inspection was applied (Williams et al., 2019). For LoTSS DR2, Alegre et al. (2022) trained a gradient boosting classifier (GBC; Breiman, 1997; Friedman, 2001) to classify radio blobs as either 'simple' or 'complex' based on the properties of the radio blobs, the properties of the Gaussians fitted to these blobs, the likelihood ratios for each, and the distance to and properties of the nearest neighbours.

We adopted the procedure of Alegre et al. (2022) and use their trained GBC to separate the simple radio blobs from those that require radio component association beyond PyBDSF's capabilities and/or optical host identification beyond the scope of the likelihood ratio method as described by Sutherland & Saunders (1992). We expect most giant radio galaxies to fall in the latter case.

8.4.4 Associating radio emission into radio sources

We proceeded with automatic radio source component association for the complex radio blobs. Following the procedure laid out by Mostert et al. (2022), for each of these radio blobs, we created a $300'' \times 300''$ LoTSS DR2 image cutout centred on the radio blob. Next, a Fast region-based convolutional neural network (Fast R-CNN; Girshick, 2015), adapted and trained for this purpose by Mostert et al. (2022), was applied to these cutouts to predict which (if any) other radio blobs — whether they be complex or simple — should be associated to the centred radio blob for it to form a single physical radio source. For example, the two lobes of an RG, each represented by a radio blob, might be associated together to form a single physical radio source. Due to the fixed $300'' \times 300''$ image size for which the Fast R-CNN was trained, we expect most radio sources that are associated in our pipeline to have an angular length $\varphi < 424''$.[12]

[12] If predicted associations from neighbouring cutouts have an overlapping radio blob, the associations will be merged. For example: in cutout 1 lobe A and core B are associated and in cutout 2

The result is a radio source catalogue in which some of the radio blobs have been merged, and a component catalogue that lists for each radio blob to which radio source it belongs. The radio and the component catalogue were completed by appending to them the remaining list of simple radio blobs.

8.4.5 Optical or infrared host galaxy identification

Barkus et al. (2022) created a method to identify an extended radio source's optical or infrared host. The method described by Barkus et al. (2022) takes the radio morphology into account by drawing a ridgeline along the regions of high flux density. The method continues with the application of the likelihood ratio method to quantify which pairs of host galaxy candidates and radio sources are a plausible match. The likelihood ratio LR used in this context follows Eq. 1 of Sutherland & Saunders (1992), with the slight simplification of having the latter's dependence on two angular offsets replaced by a dependence on a *radial* angular offset only:

$$LR = \frac{q(m,c)f(r)}{n(m,c)}, \tag{8.31}$$

where $q(m,c)$ is a prior on the magnitude m and colour c of the optical host, $f(r)$ is a function of the angular offset between the optical centroid and the radio centroid, and $n(m,c)$ normalises for the number density of optical sources with a certain magnitude m and colour c in the catalogue used for the cross-matching.

To adapt the likelihood ratio for use in the case of extended radio sources, Barkus et al. (2022) implemented the different components of the ratio as follows. For $n(m,c)$, Barkus et al. (2022) estimated the probability density over m and c for a distribution of 50, 000 randomly sampled sources from a combined Pan-STARRS–AllWISE catalogue in the region of the sky that overlapped with LoTSS DR1. For $q(m,c)$, they estimated the probability over m and c for sources from the combined Pan-STARRS–AllWISE catalogue that were manually selected to be the most likely optical/near-infrared host for a sample of 950 radio sources with angular length $\varphi > 15''$. For both $n(m,c)$ and $q(m,c)$, the AllWISE W1 magnitudes were used for m, the Pan-STARRS i-band magnitudes minus the AllWISE W1 magnitudes were used for colour c, and the PDF was formed using a 2D kernel density estimator (KDE; e.g. Pedregosa et al., 2011) with a Gaussian kernel and a bandwidth of 0.2. For extended asymmetric or bent radio galaxies, the optical host is not likely to be found at the radio centroid.

core B and lobe C are associated, then the set (lobe A, core B, and lobe C) will enter the catalogue as a single radio source. Thereby creating the possibility of detecting radio sources with angular length $\varphi > 424''$.

Thus, Barkus et al. (2022) proposed $f(r)$ to be a function of both the distance between the radio centroid and the optical source $r_{\text{opt,centroid}}$ and the smallest distance between the optical source and a ridgeline fitted to the radio source $r_{\text{opt,ridge}}$. Specifically,

$$f(r) = f_{\text{ridge}}(r_{\text{opt,ridge}}) \cdot f_{\text{centroid}}(r_{\text{opt,centroid}}), \tag{8.32}$$

with

$$f_{\text{ridge}}(r_{\text{opt,ridge}}) = \frac{1}{2\pi\sigma_r^2} e^{\frac{-r_{\text{opt,ridge}}^2}{2\sigma_r^2}}, \tag{8.33}$$

and

$$f_{\text{centroid}}(r_{\text{opt,centroid}}) = \frac{1}{2\pi\sigma_c^2} e^{\frac{-r_{\text{opt,centroid}}^2}{2\sigma_c^2}}, \tag{8.34}$$

where $\sigma_r^2 = \sigma_{\text{opt}}^2 + \sigma_{\text{radio}}^2 + \sigma_{\text{astr}}^2$ and the chosen value for the astrometric uncertainty σ_{astr} is $0.2''$, the optical position uncertainties σ_{opt} are taken from the optical catalogue (generally $\sim 0.1''$), the radio position uncertainty σ_{radio} is fixed to $3''$, and the uncertainty in the centroid position σ_c is empirically estimated at 0.2 times the length of the considered radio source. The parameters in f_{ridge} are expressed in arcsec and those in f_{centroid} as a fraction of the radio source length. For the 30 optical sources closest to the radio ridgeline Barkus et al. (2022) calculate the value of LR, whereby the optical source with the highest LR value is considered to be the most likely host galaxy.

We use the method by Barkus et al. (2022) but propose three minor adaptations. First, we introduce explicit regularisation for $q(m, c)$ and $n(m, c)$. As the PDF estimates for $q(m, c)$ and $n(m, c)$ are 2D KDEs over sampled (m, c)-distributions, the parts of the (m, c)–parameter space that are sparsely sampled can lead to probabilities that are effectively zero when the realistic theoretical probability should be small but non-zero. Through the q/n-fraction in Eq. 8.31, the resulting values of LR in the sparsely sampled parts of the (m, c)–parameter space blow up to unrealistic large values or collapse to almost 0 (see Fig. 8.12). In practice, these unsampled parts of parameter space are almost never visited by new sources for which we calculate LR. Even so, we add a constant factor to the KDE estimate of q and n to get more robust LR values (see Fig. 8.13) and to express the model uncertainties in our functions of q and n. Using 10-fold cross-validation, we empirically select the bandwidths for the KDEs leading to q and n to be 0.4. Second, we propose an alternate form of $f(r)$. For giants, $f(r)$ is rarely dominated by errors in the position of the optical source or that of the radio source. As $r_{\text{opt,centroid}}$ and $r_{\text{opt,ridge}}$ are slightly correlated, multiplication of $f_{\text{ridge}}(r_{\text{opt,ridge}})$ and $f_{\text{centroid}}(r_{\text{opt,centroid}})$ underestimates the chance of low values of $r_{\text{opt,centroid}}$ or $r_{\text{opt,ridge}}$. Therefore, we combine $r_{\text{opt,centroid}}$ and $r_{\text{opt,ridge}}$ into a single parameter r_{mean} that is the mean of the two distance parameters. Furthermore, we ob-

serve that the empirical distributions of $r_{\text{opt,centroid}}$, $r_{\text{opt,ridge}}$ and r_{mean} for a sample of radio sources with angular length $> 1'$ $A_{\text{radio,opt}}$ for which optical counterparts were determined via visual inspection do not follow a normal distribution as assumed by Barkus et al. (2022) but rather a lognormal distribution (see Fig. 8.14). Instead of estimating the values of the different error components (astrometric error, error in optical position, error in radio position) we use the empirical values of the distribution of $f(r)$ for the sources in $A_{\text{radio,opt}}$; see Appendix 8.A4 for details. Third, we replaced the Pan-STARRS1 DR1 catalogue (from which colour c was derived) with the DESI Legacy Imaging Surveys DR9 catalogue, as the latter goes up to an i-band magnitude of 24.

We applied the modified ridgeline method to all radio sources in our pipeline catalogue with angular lengths larger than $1'$ and brighter than 10 mJy. We limit the ridgeline procedure to these sources to save time, as the procedure takes multiple seconds per radio source.

After detecting the optical host galaxies for our radio sources, we checked for corresponding spectroscopic redshift estimates from SDSS (VizieR catalogue `V/147/sdss12`, Ahn et al., 2012), or if not available, for photometric redshift estimates from DESI (VizieR catalogue `VII/292`, Duncan, 2022). The SDSS catalogue also provides us with velocity dispersions and a quasar flag. The DESI VizieR catalogue includes a flag (FCLEAN $= 1$) that indicates that the optical source used for the photometric redshift prediction is free from blending or image artefacts. The catalogue also includes a column (PSTAR) that estimates how likely it is that the optical source is a star based on its colours. We discard all sources in both the pipeline catalogue and the LGZ catalogue for which either FCLEAN $\neq 1$ or PSTAR > 0.2.

8.4.6 Re-assessing angular source lengths

Next, we proceeded to re-assess the angular source lengths, for both the radio catalogue created using the ML pipeline and those reported by the LGZ catalogue described by Hardcastle et al. (2023). Up to this point, the angular source lengths in both catalogues are the full width at half maximum (FWHM) of the combined Gaussian components that make up a source, if the source is only composed of a single radio blob. If the radio source is composed of multiple radio blobs, the reported size is the distance between the two furthest removed points on a convex hull that encloses the FWHMs of the blobs that make up the radio source. However, in the literature, the length of a giant is often reported to be the maximum distance between the signal of a radio source that exceeds three times the image noise σ.

To get the 3σ angular lengths, we applied five steps to all sources in both catalogues

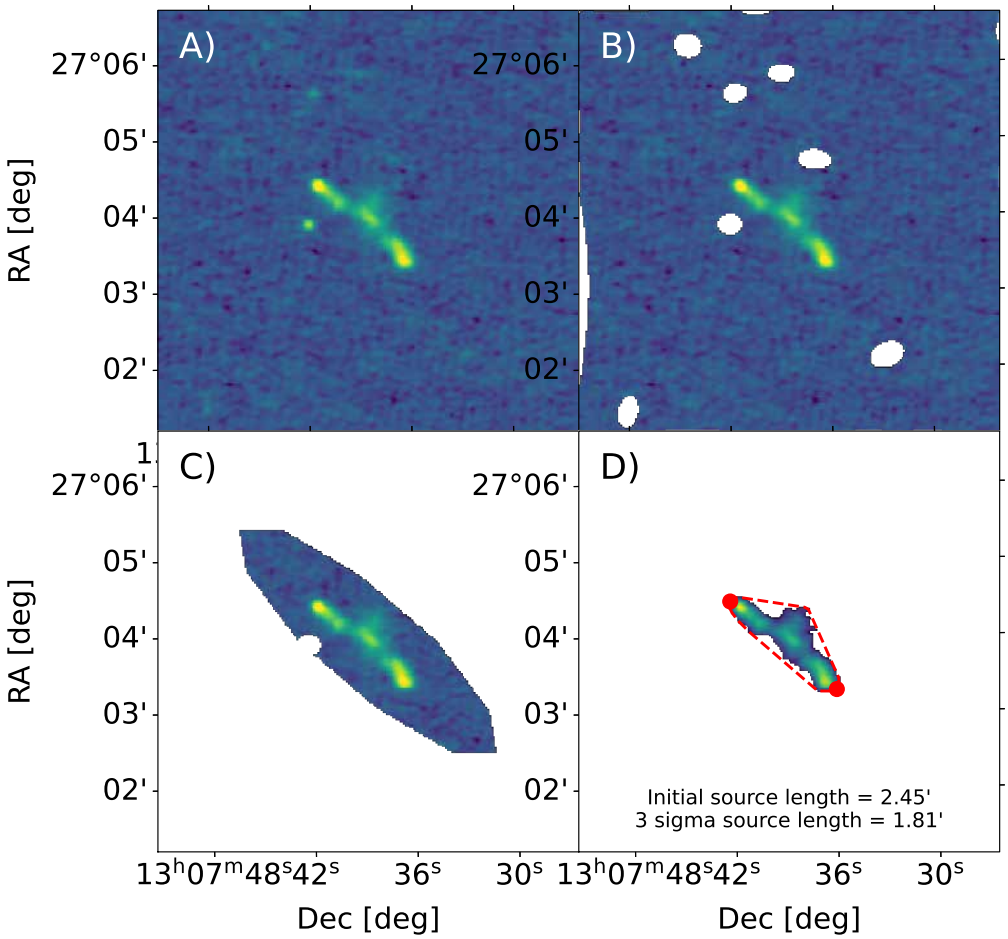


Figure 8.3: Four panels, depicting the re-evaluation of the angular length of radio source ILTJ130738.79+270355.1. Panel A) shows the initial cutout, B) shows the removal of neighbouring sources, C) the masking of emission outside a convex hull based on the old source length, D) shows the emission that is left after masking all emission below 3 times the local noise σ. The red dashed line shows the convex hull around the left-over emission, and the red points indicate the furthest removed points on this convex hull. The distance between these points yields the 3σ angular length.

with a reported angular length $\varphi > 1'$. First, we created a square image cutout with a width and height equal to 1.5 times the old angular source length. Second, we mask all neighbouring radio emission. Third, we mask all emission outside an ellipse with a major axis that is the old source length, and a minor axis that is 1.1 times the old source width or a quarter of the old source length if that value is bigger. These numbers are a result of the observation that, with respect to the 3σ angular lengths, the old lengths were almost always significantly overestimated, while the source width tended to be underestimated. Fourth, we mask all remaining emission that is below 3 times

the local noise. Fifth, we fitted a convex hull around the remaining emission and determined the distance between the points on this convex hull that were farthest apart. See Fig. 8.3 for an illustrative example.

The entire process from source detection (Sect. 8.4.1) to source list with optical identifications and updated angular lengths (this subsection) took roughly half an hour to one hour per LoTSS DR2 pointing, depending on the detected number of sources. Each pointing can be processed independently, which allowed us to spread the processing of all 841 LoTSS DR2 pointings over 5 nodes of a heterogeneous computer cluster with 80 physical CPU cores in total for three to four days.

Finally, for both the ML pipeline and LGZ catalogues, we calculate the projected proper lengths using the 3σ angular lengths and the redshift estimates corresponding to each source, and discard all sources that do not meet the $l_{\rm p} \geq l_{\rm p,GRG} = 0.7$ Mpc GRG criterion. For the ML pipeline catalogue, we discarded all internally duplicate GRG candidates using a $1'$ cone search. The LGZ catalogue did not contain any internal duplicates. That left us with $7,001$ GRG candidates in the pipeline catalogue and $7,044$ GRG candidates in the LGZ catalogue.

8.4.7 Manual verification of obtained GRG sample

The following step we took in the creation of our GRG sample, was a manual visual inspection of all giant candidates. For the LGZ giant candidates, as described by Hardcastle et al. (2023), at least five different volunteers already inspected the radio and corresponding optical emission. The purpose of our manual visual inspection was therefore to exclude only those sources where either the radio component association or the host identification was obviously incorrect. For each giant, a single expert looked at a panel showing the giant with its neighbouring sources masked and most neighbouring emission masked (akin to panel C in Fig. 8.3) and a panel showing the giant in its wider context (akin to panel A in Fig. 8.3); additionally, the location of the optical host was indicated. We sorted the GRG candidates into three categories: candidates that looked reasonable, candidates that clearly missed (or included too many) significant radio components, and candidates that showed a very unlikely host galaxy location. For the ML pipeline giant candidates, we initially followed the same procedure as for the LGZ giant candidates. To make visual inspection feasible, we skipped the $4,272$ ML pipeline giant candidates that overlapped with the verified LGZ giants. After inspecting the ML pipeline giant candidates once, we subjected all giants that were not rejected to a second round of visual inspection. The second round was aided by inspecting LoTSS DR2 radio contours over a Legacy Survey DR9 (g, r, z) image cube, where sources from the combined optical–infrared catalogue within the

field of view were highlighted.

For the LGZ catalogue, we judged $6,550$ (93%) GRG candidates to be without issues, 389 (6%) to have radio component issues, and 105 (1%) to have been assigned an unlikely host galaxy. For the $5,864$ (unique) ML pipeline giant candidates, we judged $2,722$ (47%) candidates to be without issues, $1,963$ (33%) to have radio component issues, and $1,179$ (20%) to have been assigned an unlikely host galaxy. From the $6,550$ LGZ giants, we find $5,596$ of those to be newly discovered (not appearing in previous literature), and for the $2,722$ ML pipeline giants, we find $2,592$ to be newly discovered.

Qualitatively, from the visual inspection, we noticed that the verified ML pipeline GRG sample contained more symmetric giants with colinear jets, while the verified LGZ GRG sample contained more giants with complex, bent structures indicative of interaction with the IGM. The ML pipeline did also detect giants with complex structures, but was often unable to fully separate them from all neighbouring unrelated emission. An in-depth comparison between the ML pipeline and LGZ GRG samples is beyond the scope of this work. Figures 8.4 and 8.5 each show six examples of previously unknown giants found through our ML-based approach. Through cutouts covering $3' \times 3'$, Fig. 8.4 shows angularly compact giants; through cutouts covering $6' \times 6'$, Fig. 8.5 shows more angularly extended specimen.

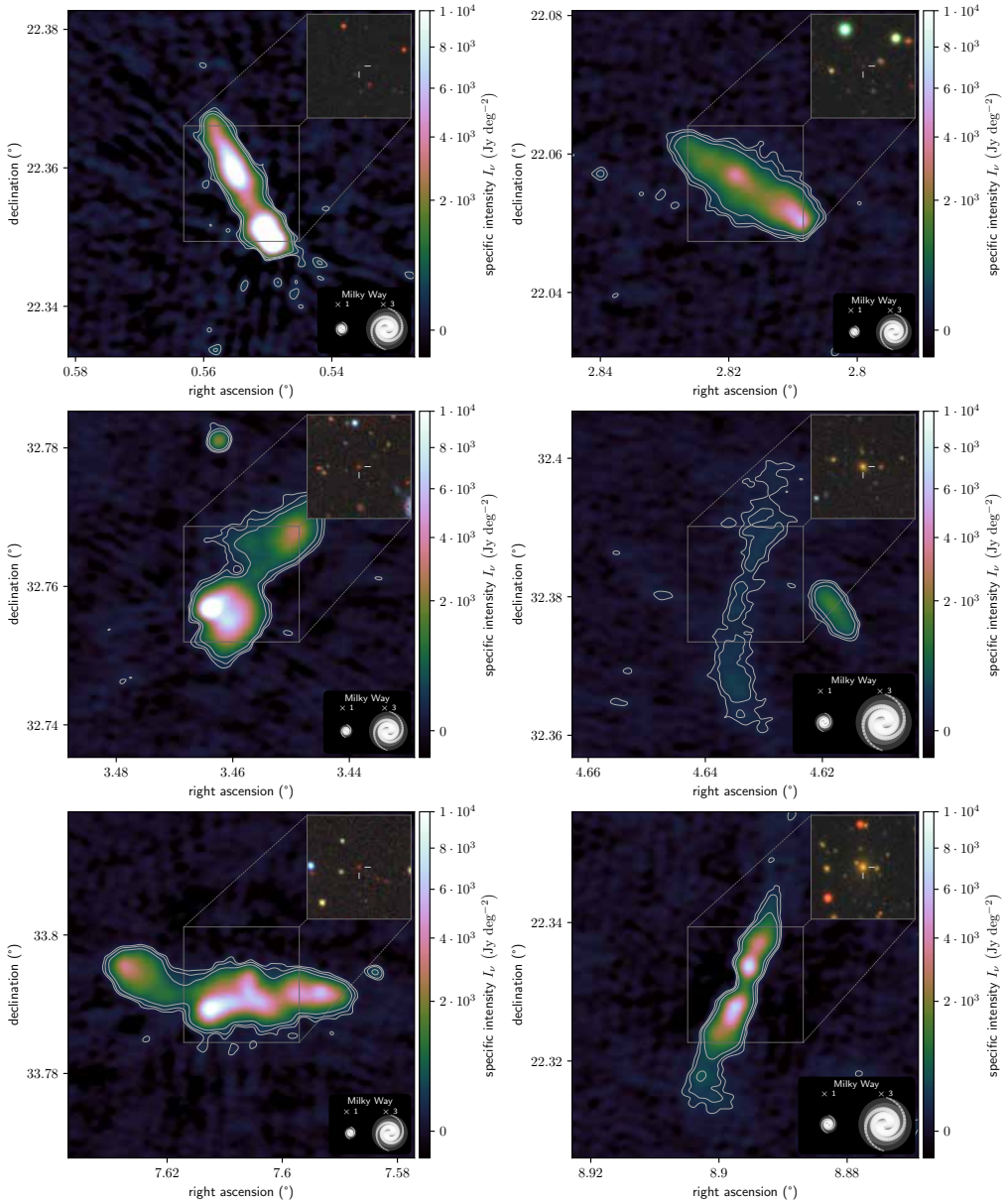


Figure 8.4: LoTSS DR2 cutouts at central observing frequency $\nu_{\rm obs} = 144$ MHz and resolution $\theta_{\rm FWHM} = 6''$, centred around the hosts of newly discovered giants. Each cutout covers a solid angle of $3' \times 3'$. Contours signify 3, 5, and 10 sigma-clipped standard deviations above the sigma-clipped median. For scale, we show the stellar Milky Way disk (with a diameter of 50 kpc) generated using the Ringermacher & Mead (2009) formula, alongside a 3 times inflated version. Each DESI Legacy Imaging Surveys DR9 (g, r, z) inset shows the central $1' \times 1'$ square region. As all giants obey $\varphi \geq 1.3'$, they must — if not oriented along one of the square's diagonals — necessarily protrude from this region. Rowwise from left to right, from top to bottom, these giants are ILTJ000212.45+222116.2, ILTJ001115.77+220316.6, ILTJ001350.25+324530.8, ILTJ001831.84+322247.7, ILTJ003025.90+334729.2, and ILTJ003534.45+221937.8.

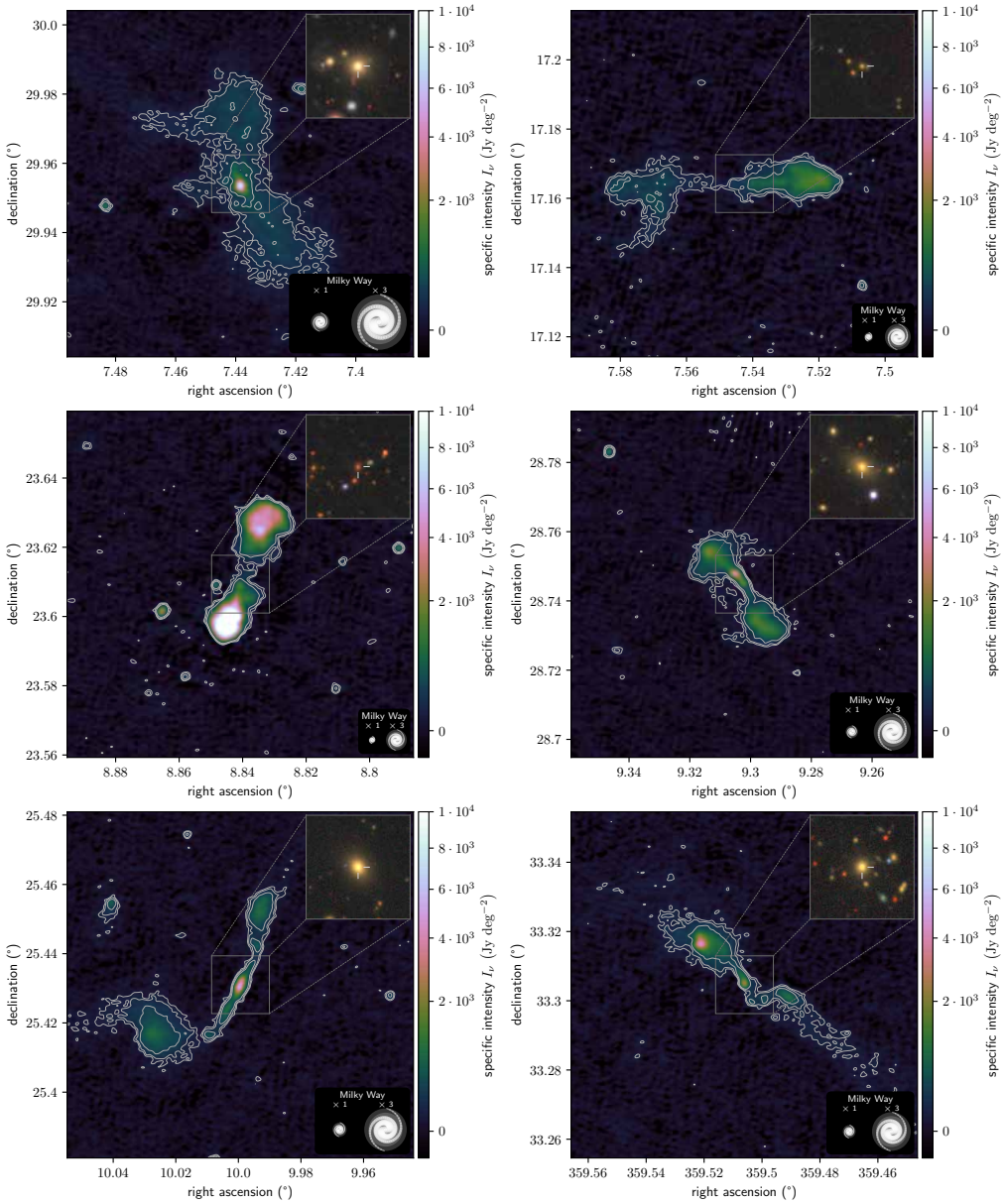


Figure 8.5: LoTSS DR2 cutouts at central observing frequency $\nu_{\mathrm{obs}} = 144$ MHz and resolution $\theta_{\mathrm{FWHM}} = 6''$, centred around the hosts of newly discovered giants. Each cutout covers a solid angle of $6' \times 6'$. Contours signify 3, 5, and 10 sigma-clipped standard deviations above the sigma-clipped median. For scale, we show the stellar Milky Way disk (with a diameter of 50 kpc) generated using the Ringermacher & Mead (2009) formula, alongside a 3 times inflated version. Each DESI Legacy Imaging Surveys DR9 (g, r, z) inset shows the central $1' \times 1'$ square region. As all giants obey $\varphi \geq 1.3'$, they must — if not oriented along one of the square's diagonals — necessarily protrude from this region. Rowwise from left to right, from top to bottom, these giants are ILTJ002943.72+295700.3, ILTJ003010.58+170948.6, ILTJ003521.87+233625.9, ILTJ003712.91+284436.8, ILTJ004002.30+252550.9, and ILTJ235802.49+331838.5.

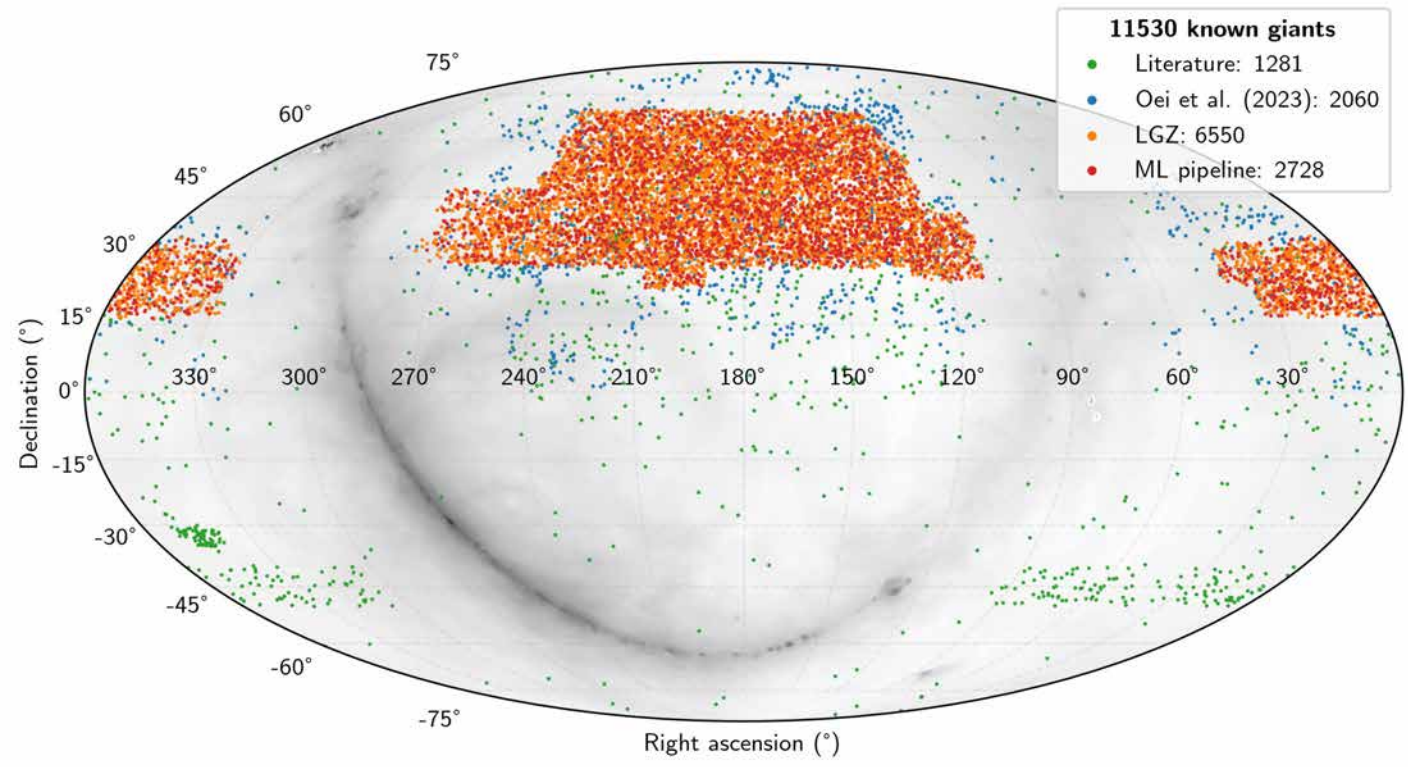


Figure 8.6: With 11, 524 unique sources, we present the largest collection of giants discovered as of yet. The ML pipeline and LGZ samples (red and orange markers) are strictly confined to the LoTSS DR2 area, while the sample by Oei et al. (2023a) extends to yet-to-be-released LoTSS pointings processed with the DR2 pipeline.

8.4.8 Merging our sample with the GRG sample in the literature

To complete our GRG sample, we iteratively added giants from the literature, going from the newest to the oldest publication. This approach follows from the assumption that newer publications are generally based on more sensitive and higher-resolution observations, leading to more accurate angular length estimates. In an effort to avoid having duplicate giants in the final sample, we only added giants when their host galaxies were more than $1'$ away from all host galaxies of the already aggregated giants.

The LGZ–ML pipeline GRG sample contains $9,272$ giants. We added $1,432$ out of the $2,193$ giants presented by Oei et al. (2023a), 41 out of the 69 giants presented by Simonte et al. (2022), 62 out of the 62 giants presented by Gürkan et al. (2022), 163 out of the 263 giants presented by Mahato et al. (2022), 178 out of the 178 giants presented by Andernach et al. (2021), 0 out of the 1 giants presented by Masini et al. (2021), 2 out of the 2 giants presented by Delhaize et al. (2021), 1 out of the 2 giants presented by Bassani et al. (2021), 1 out of the 4 giants presented by Tang et al. (2020), 372 out of the 694 giants presented by Dabhade et al. (2020a), and 0 out of the 6 giants presented by Ishwara-Chandra et al. (2020). These additions result in the final catalogue containing $11,524$ unique giants. This is the first catalogue of giants to contain more than 10^4 specimen.

Figure 8.6 shows a Mollweide view of the sky with the locations of both the newly confirmed giants and the giants from the literature. Almost all discovered giants stay clear of the Galactic Plane, where radio emission from the Milky Way — of which we show the specific intensity function at $\nu_{\rm obs} = 150$ MHz in greyscale (Zheng et al., 2017) — makes calibration and imaging harder. In addition, optical host identification is much harder near the Galactic Plane. The default field of view set-up of both our ML pipeline (Sect. 8.4.4) and of LGZ favours the discovery of giants with angular lengths of a few arcminutes at most. By contrast, the GRG campaign of Oei et al. (2023a) featured a 'fuzzy' $\sim 5'$ lower threshold to allow for an exhaustive manual search with an interactive and dynamic field of view (using Aladin; Bonnarel et al., 2000). In Fig. 8.7, we demonstrate that these design choices lead to GRG samples with markedly different angular length distributions.

As a result, the samples complement each other: the sample of Oei et al. (2023a) is more complete at lower redshifts and higher projected lengths, while the LGZ and ML pipeline samples are more complete at higher redshifts and lower projected lengths. Figure 8.8 demonstrates this point, while Table 8.1 presents the corresponding statistics of the GRG samples.

For comparison of the 3σ lengths of the ML pipeline and LGZ giants to those in

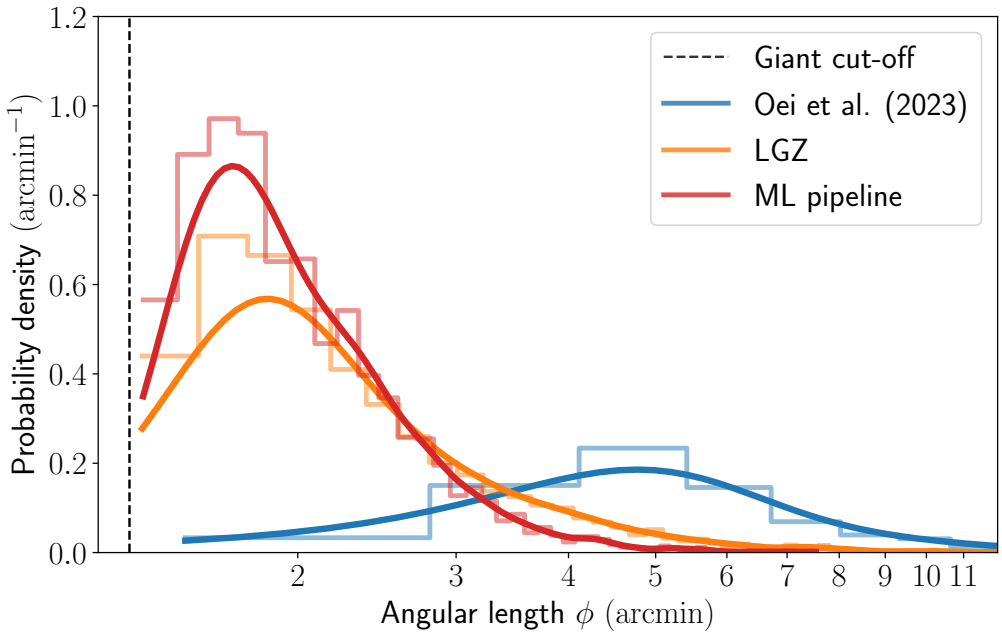


Figure 8.7: Observed distributions for angular length φ, showing that our three LoTSS DR2 search methods target different ranges of φ. The largest angular lengths detected by Oei et al. (2023a), LGZ, and the ML pipeline are 132′, 43′, and 8′ respectively, but we limit the horizontal axis to 12′ for interpretability. The vertical line marks the minimum angular length that giants can attain: $\varphi_{\rm GRG}(l_{\rm p,GRG} = 0.7\ {\rm Mpc}) = 1.3'$.

Table 8.1: Statistics of the GRG (sub)samples that we discovered, confirmed, or aggregated. From left to right, the columns provide the number of giants in each sample, N, and the 10th, the median, and the 90th percentile of the angular length φ, redshift z, and projected proper length $l_{\rm p}$.

Sample	N	$\varphi_{\rm 10th}$ (′)	$\varphi_{\rm median}$ (′)	$\varphi_{\rm 90th}$ (′)	$z_{\rm 10th}$	$z_{\rm median}$	$z_{\rm 90th}$	$l_{\rm p,10th}$ (Mpc)	$l_{\rm p,median}$ (Mpc)	$l_{\rm p,90th}$ (Mpc)
ML pipeline	2,722	1.50	1.97	3.09	0.44	0.87	1.28	0.72	0.87	1.29
LGZ	6,550	1.56	2.19	4.40	0.31	0.75	1.19	0.73	0.91	1.57
Known giants	11,524	1.57	2.33	5.70	0.23	0.72	1.19	0.73	0.94	1.68

other surveys, we inform the reader that the central frequency and the average surface brightness threshold of the observations that we use are $\nu_{\rm obs} = 144$ MHz and $b_{\nu,\rm th} = 25$ Jy deg^{-2} respectively.

8.4.9 Bayesian parameter estimation

After having refined our statistical GRG framework (Sect. 8.2), and after having assembled the largest sample of giants yet (Sects. 8.4.1–8.4.8), we combined both advances to perform inference of the length distribution, number density, and lobe volume-filling fraction of giants.

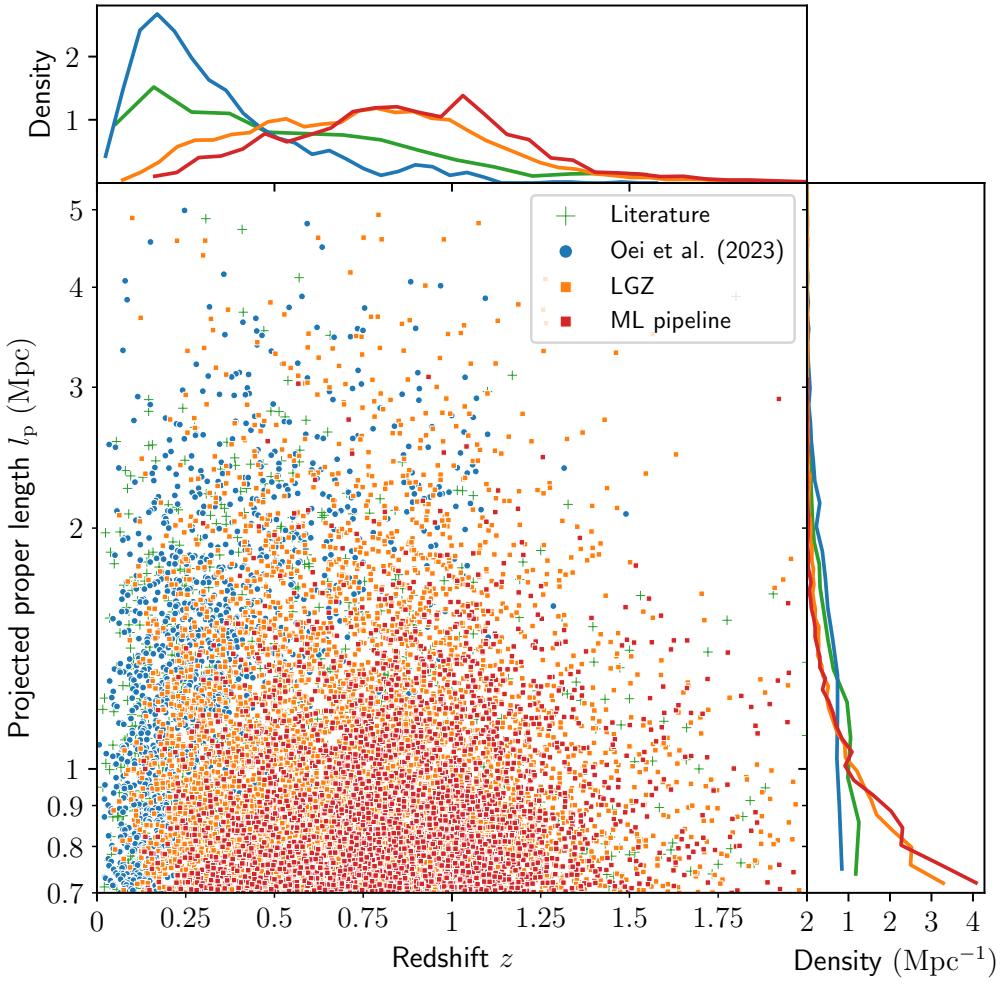


Figure 8.8: Our sample of LGZ giants (orange squares) and ML pipeline giants (red squares) effectively complements the sample of giants with large angular lengths (blue dots) from the manual search of Oei et al. (2023a). The remaining giants (green pluses) are from earlier literature, as specified in Sect. 8.4.8.

Given that our goal has been to infer properties of the full population of giants, rather than just of those currently observed, we included two main selection effects in our forward modelling. As detailed in Sect. 8.2.4, we parametrised surface brightness selection with three parameters, which are free parameters of the model. As detailed in Sect. 8.2.4, a second cause of selection is the imperfect operation of our three LoTSS DR2 search methods, all of which fail to identify a significant fraction of giants with lobe surface brightnesses *above* the survey noise level. We modelled this identification selection $p_{\mathrm{obs,ID}}$ with a set of logistic functions, regressed to GRG data.

We now provide details of this process.

Identification probability functions

To estimate $p_{\mathrm{obs,ID}}(l_{\mathrm{p}}, z)$ from data, we first selected all giants detected by the joint efforts of our machine learning pipeline, LGZ, and the manual, visual search of Oei et al. (2023a). Next, we retained only those giants that are located in regions of the sky that have been scanned by all three searches. This overlap region in principle corresponds to the full LoTSS DR2 coverage — were it not for the fact that the search of Oei et al. (2023a) skipped over the LoTSS DR1, which had already been scanned by Dabhade et al. (2020b). Thus, the actual overlap region amounts to the LoTSS DR2 coverage with a spherical quadrangle removed, whose minimum and maximum right ascensions are $\alpha_{\mathrm{min}} = 160^\circ$ and $\alpha_{\mathrm{max}} = 230^\circ$ and whose minimum and maximum declinations are $\delta_{\mathrm{min}} = 45^\circ$ and $\delta_{\mathrm{max}} = 56^\circ$. In Appendix 8.A5, we provide an explicit decomposition of our assumed LoTSS DR2 coverage — and thus implicitly of the overlap region — in terms of disjoint spherical quadrangles.

Some of the retained giants have been detected only in the combined ML–LGZ search, others have been detected only in the Oei et al. (2023a) search, and yet others have been detected in both. We note that, had it operated flawlessly, the combined ML–LGZ search would have detected all sources claimed by Oei et al. (2023a) (or at least those that are genuine giants — which should be the vast majority). Thus, by mapping the (in)ability of the ML–LGZ search to detect the giants of Oei et al. (2023a) as a function of l_{p} and z, one can estimate the ML–LGZ search's identification probability function, $p_{\mathrm{obs,ID,1}}(l_{\mathrm{p}}, z)$. More precisely, for each giant detected by Oei et al. (2023a), we evaluated whether it was also detected in the ML–LGZ search, and stored a corresponding Boolean (that is to say, either 1 or 0). We show these Booleans, at the (l_{p}, z) coordinates of the giants they belong to, as yellow (representing 1) and blue (representing 0) dots in the top-left panel of Fig. 8.9. Viewing the Boolean at (l_{p}, z) as a realisation of a Bernoulli RV with success probability $p = p_{\mathrm{obs,ID,1}}(l_{\mathrm{p}}, z)$, we recognise the inference of the identification probability function as a binary logistic regression problem with two explanatory variables. The background of Fig. 8.9's top-left panel shows the corresponding best fit.

By symmetry, this approach can be reversed to estimate the Oei et al. (2023a) search's identification probability function, $p_{\mathrm{obs,ID,2}}(l_{\mathrm{p}}, z)$. Thus, for each giant detected in the ML–LGZ search, we evaluated whether it was also detected by Oei et al. (2023a), and stored a corresponding Boolean. In the same way as before, we show these Booleans in the middle-left panel of Fig. 8.9. The panel's background shows the best logistic fit.

We combine the two identification probability functions, $p_{\mathrm{obs,ID,1}}(l_{\mathrm{p}}, z)$ and $p_{\mathrm{obs,ID,2}}(l_{\mathrm{p}}, z)$, in point-wise fashion as to obtain a single function $p_{\mathrm{obs,ID}}(l_{\mathrm{p}}, z)$. To do so, we follow the minimal combination rule of Eq. 8.20.

We remark that, by giving each Boolean in these logistic regressions an equal weight, the resulting functions are tuned to fit crowded regions of projected length–redshift parameter space best — at the expense of accuracy in sparser regions. To increase the accuracy of the functions for the parameter space at large, we performed a simple rebalancing step. First, we calculated the mean number density in the parameter space given by $l_{\mathrm{p}} \in [0.7, 5\ \mathrm{Mpc}] \times [0, 0.5] \ni z$. We then selectively subsampled the data in crowded regions, following the rule that the number density in each bin of width 0.5 Mpc and height 0.05 should not exceed twice the mean number density of the entire parameter space. We show the rebalanced data, alongside refitted logistic models, in the right column of Fig. 8.9. We report the rebalanced model coefficients in Table 8.2, and treat them as constants during the Bayesian inference.

Inference in practice

In this work, we constrain the parameters of Sect. 8.2's GRG population model via a projected length–redshift histogram. From our most extensive sample of giants, we select those with $0.7\ \mathrm{Mpc} =: l_{\mathrm{p,GRG}} < l_{\mathrm{p}} < 5.1\ \mathrm{Mpc}$ and $0 < z < z_{\mathrm{max}} := 0.5$ that lie in the LoTSS DR2 coverage as specified in Appendix 8.A5. We do not include the giants from Oei et al. (2023a) for which only a lower bound to the redshift is known. This selection retains $2,685$ out of $11,524$ giants. We use these giants to fill a histogram with bins of width $\Delta l_{\mathrm{p}} = 0.1\ \mathrm{Mpc}$ and $\Delta z = 0.02$. We have not systematically explored the effect of these bin size parameters on the resulting inference. However, the smaller one chooses the bins, the higher the numerical cost will be. On the other hand, if the bins are chosen much larger than the typical scales over which the underlying observed projected length–redshift distribution[13] varies, then some ability to extract parameter constraints will be lost.

To compute the posterior distribution over the six parameters $\theta = [\xi(l_{\mathrm{p,GRG}}), \Delta\xi, b_{\nu,\mathrm{ref}}, \sigma_{\mathrm{ref}}, \zeta, n_{\mathrm{GRG}}]$, we assumed a uniform prior and brute-force evaluated the likelihood function over a regular grid that covers a total of $2.1 \cdot 10^9$ parameter combinations.[14] In doing so, we applied the Poissonian likelihood trick described in Ap-

[13] With the 'underlying' observed projected length–redshift distribution, we mean the observed projected length–redshift distribution one would obtain in the limit of an infinite number of observed giants.

[14] This approach is feasible by virtue of the low numerical cost of each likelihood function evaluation. Its main advantage is its simplicity: there are no parameters to tune that govern the method's convergence behaviour. Once the model is expanded to include more parameters, or when selection ef-

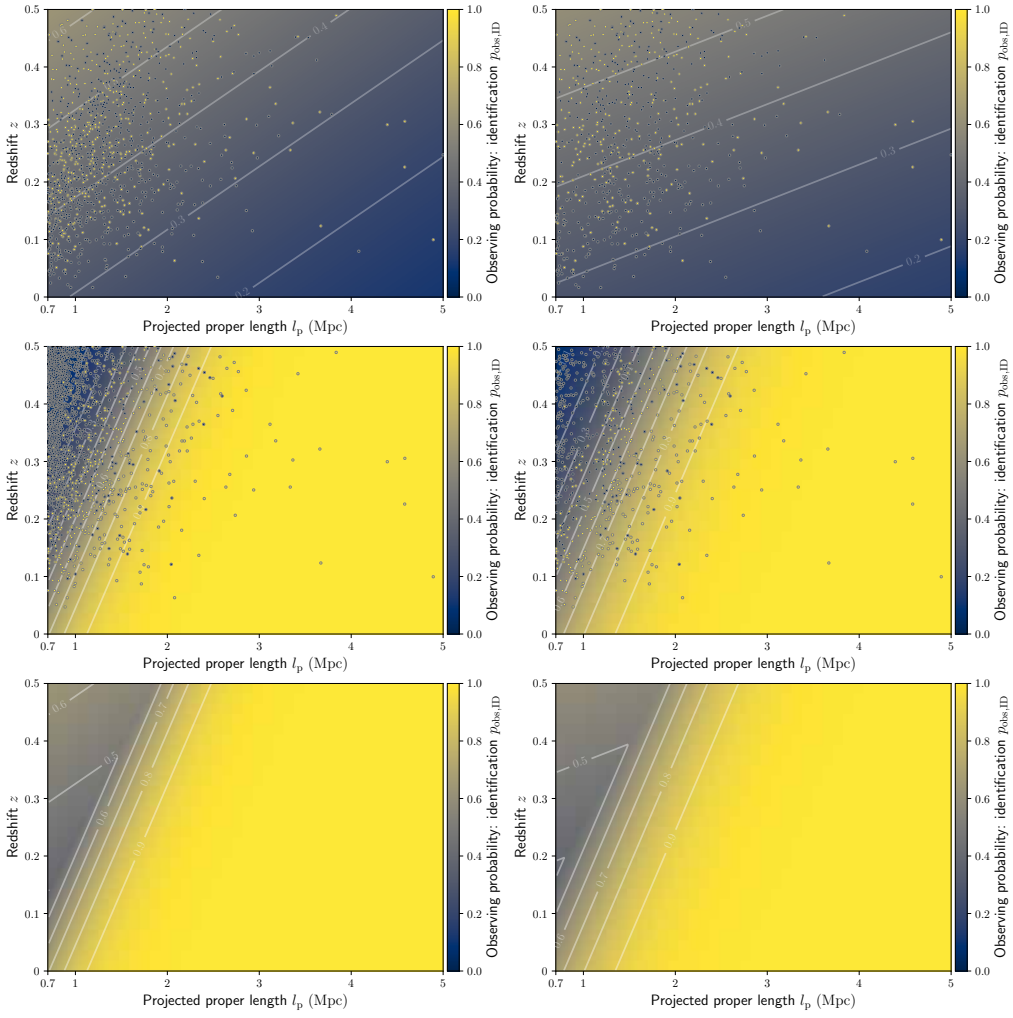


Figure 8.9: Overview of our determination of the probability to identify giants in the LoTSS DR2 with above-noise surface brightnesses, as a function of projected length and redshift — through machine learning or LOFAR Galaxy Zoo (top row), through the search of Oei et al. (2023a) (middle row), and through these methods in unison (bottom row). Each of the upper four panels shows a binary logistic regression following the theory of Sect. 8.2.4 and the practical considerations of Sect. 8.4.9. The left column shows results from all available data, whilst the right column shows results from rebalanced data. In our Bayesian inference, we use the latter results.

pendix 8.A2, which sped up our computations by one to two orders of magnitude. Table 8.2 provides the parameter ranges for which we evaluated the likelihood (which coincide with their prior distribution ranges), alongside all of the model's constants

fects with higher numerical cost are incorporated, more efficient (though more complicated) methods such as Markov chain Monte Carlo or nested sampling will become necessary.

and their assumed values. Because each likelihood function evaluation can be computed independently of the others, the problem is fully parallelisable. In practice, we distributed the $\sim 10^4$ core-hours Python calculation over $\sim$1500 virtual cores, which were spread across $\sim$20 nodes of a computer cluster. Next, we generated samples from the posterior distribution using rejection sampling (e.g. Rice, 2006). We subsequently used these samples to calculate probability distributions for derived quantities.[15]

8.5 Results

By combining an unparalleled sample of giant radio galaxies with a rigorous forward model, we have produced a posterior distribution over parameters that characterise the intrinsic population of giants. Figure 8.10 summarises the posterior over parameter hexads $\theta = [\xi(l_{\mathrm{p,GRG}}), \Delta\xi, b_{\nu,\mathrm{ref}}, \sigma_{\mathrm{ref}}, \zeta, n_{\mathrm{GRG}}]$ by means of its one- and two-dimensional marginal distributions. In this section, we analyse our newfound parameter constraints.

8.5.1 GRG length distribution

Radio galaxies enrich the IGM with magnetic fields, but giants — given their megaparsec-scale reach — appear uniquely capable of seeding the more remote regions of the Cosmic Web. Consequently, scientific interest in quantifying the length distribution of giants has arisen from the possibility that giants contribute significantly to cosmic magnetogenesis. The question at hand is deceivingly simple: how common are giants of various lengths?

As pointed out by Oei et al. (2023a), due to selection effects, the observed projected length distribution is not a reliable estimate of the *true* projected length distribution. Worse still, the relevant selection effects might not be quantitatively known a priori, requiring joint inference of the length distribution, and the selection effect parameters. Oei et al. (2023a) performed such a joint inference, and found that their data were consistent with an underlying population of giants with Pareto-distributed lengths, characterised by tail index $\xi = -3.4 \pm 0.5$. In the current work, we have relaxed the assumption of perfect Paretianity, and explore whether the data are consis-

[15] To calculate probability distributions over quantities that are a function of the parameters, such as the Local Universe GRG lobe VFF, $\mathcal{V}_{\mathrm{GRG-CW}}(z = 0)$, or the joint search completeness function C, we could in principle evaluate these quantities for each parameter combination of the aforementioned grid and weigh each grid point's result by the associated likelihood (or, equivalently, posterior probability). However, some derived quantities are costly to compute, so that excessive evaluations should be avoided.

Table 8.2: Parameters and constants of Sect. 8.2's GRG population forward model alongside their prior ranges and values, as used in the Bayesian inference presented in Sect. 8.5. The first six constants serve to define the quantitative meaning of the parameters and set the scope of the analysis. The other eleven constants are not arbitrary: they affect the likelihood function and posterior distribution for a given set of parameter definitions and scope.

Parameter	Uniform prior range	Explanation
$\xi(l_{\mathrm{p},1} = l_{\mathrm{p,GRG}})$	$[-3.5, -2]$	Sect. 8.2.3
$\Delta\xi$	$[-3.5, -1.5]$	Sect. 8.2.3
$b_{\nu,\mathrm{ref}}$	$[1, 100] \cdot \mathrm{Jy\ deg}^{-2}$	Sect. 8.2.4
σ_{ref}	$[0.5, 2]$	Sect. 8.2.4
ζ	$[-0.5, 0]$	Sect. 8.2.4
n_{GRG}	$[0, 50] \cdot (100\ \mathrm{Mpc})^{-3}$	Sect. 8.2.5

Constant	Value	Explanation
$l_{\mathrm{p,GRG}}$	0.7 Mpc	Sect. 8.1
$l_{\mathrm{p},1}$	0.7 Mpc	Sect. 8.2.3
$l_{\mathrm{p},2}$	5 Mpc	Sect. 8.2.3
l_{ref}	0.7 Mpc	Sect. 8.2.4
ν_{obs}	144 MHz	Sect. 8.2.4
z_{max}	0.5	Sect. 8.2.4
α	-1	Sect. 8.2.4
$b_{\nu,\mathrm{th}}$	$25\ \mathrm{Jy\ deg}^{-2}$	Sect. 8.2.4
$\beta_{0,1}$	-1.0	Sect. 8.2.4
$\beta_{0,2}$	-1.0	Sect. 8.2.4
$\beta_{l_{\mathrm{p}},1}$	$-0.1\ \mathrm{Mpc}^{-1}$	Sect. 8.2.4
$\beta_{l_{\mathrm{p}},2}$	$2.4\ \mathrm{Mpc}^{-1}$	Sect. 8.2.4
$\beta_{z,1}$	2.8	Sect. 8.2.4
$\beta_{z,2}$	-6.4	Sect. 8.2.4
Δl_{p}	0.1 Mpc	Sect. 8.4.9
Δz	0.02	Sect. 8.4.9
Ω	1.62 sr	Appendix 8.A5

tent with a curved power law PDF for the GRG projected proper length RV $L_{\mathrm{p}} \mid L_{\mathrm{p}} \geq l_{\mathrm{p,GRG}}$. The marginals of Fig. 8.10 suggest that they are — in fact, the data strongly favour curvature, with a tail index at $l_{\mathrm{p},1} := l_{\mathrm{p,GRG}} := 0.7$ Mpc of $\xi(l_{\mathrm{p,GRG}}) = -2.8 \pm 0.2$ and a total increase in tail index up to $l_{\mathrm{p},2} := 5$ Mpc of $\Delta\xi = -2.4 \pm 0.3$. Given the small relative uncertainty on the latter value, our data appear inconsistent with perfect Paretianity ($\Delta\xi = 0$). We note that our notion of 'data' is different from

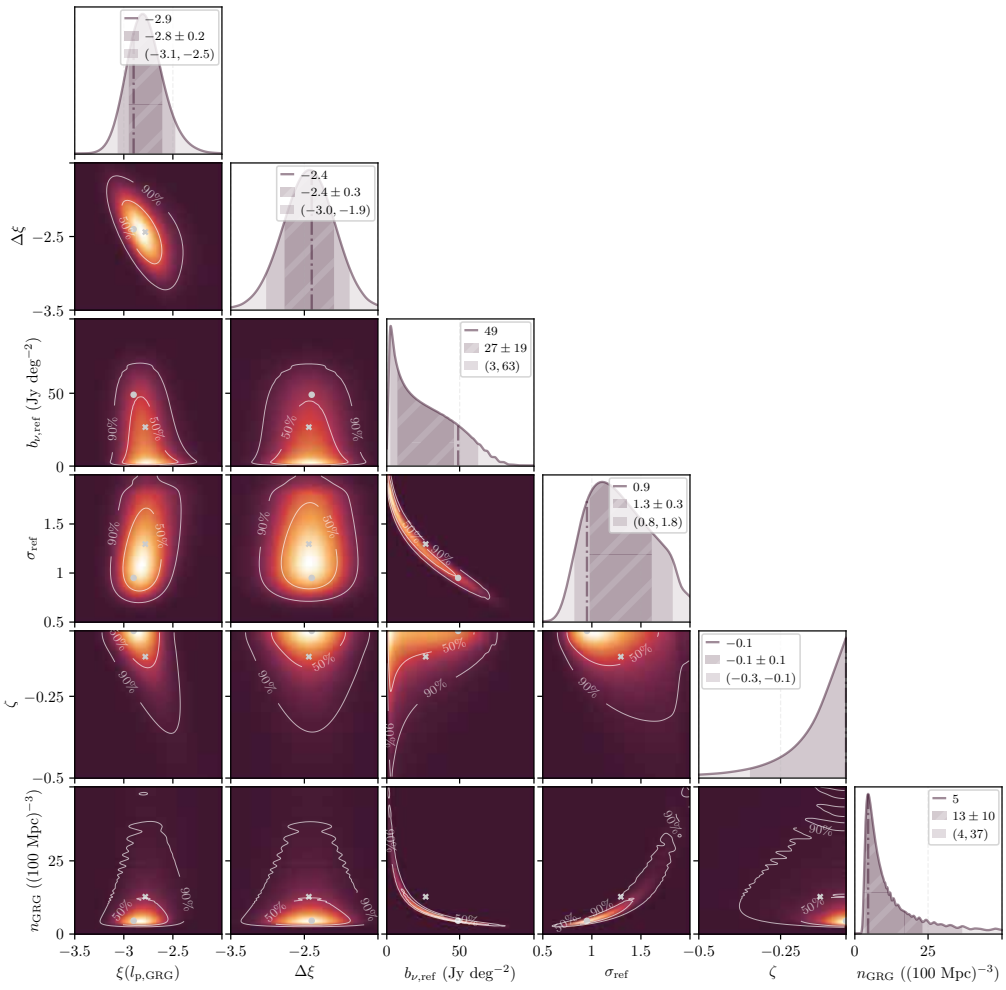


Figure 8.10: Likelihood function over $\theta = [\xi(l_{\mathrm{p,GRG}}), \Delta\xi, b_{\nu,\mathrm{ref}}, \sigma_{\mathrm{ref}}, \zeta, n_{\mathrm{GRG}}]$, based on 2685 projected lengths and redshifts of giants up to $z_{\mathrm{max}} = 0.5$. We show all two-parameter marginals of the likelihood function, with contours enclosing 50% and 90% of total probability. We mark the maximum likelihood estimate (MLE) values (grey dot) and the likelihood mean values (grey cross). The one-parameter marginals again show the MLE (dash-dotted line), a mean-centred interval of standard deviation–sized half-width (hashed region), and a median-centred 90% credible interval (shaded region).

that in Oei et al. (2023a): not only do we use more than a thousand additional giants, we also make more effective use of their redshift information. For further discussion, see Sect. 8.6.2.

It remains an open question whether giants can be understood as part of the ordinary radio galaxy population, or whether they evolve through qualitatively differ-

ent physical processes. As pointed out in Sect. 4.1.5 of Oei et al. (2023a), a curved power law PDF for $L_\mathrm{p} \mid L_\mathrm{p} \geq l_\mathrm{p,GRG}$ is consistent with a scenario in which giants share a broader length continuum with smaller radio galaxies. More specifically, if the broader radio galaxy length distribution is approximately lognormal, as appears justifiable on statistical grounds, then ξ should decrease throughout the distribution's right tail — that is, throughout the GRG range. Future research should determine whether such a unified non-giant RG–GRG scenario is also *quantitatively* consistent with the decrease in ξ we have inferred here. In addition, our inferences of $\xi(l_\mathrm{p,GRG})$ and $\Delta\xi$ are important in constraining Sect. 8.5.3's GRG lobe volume-filling fraction.

8.5.2 GRG number density

The extent to which giants have contributed to cosmic magnetogenesis depends on their intrinsic number density — which need not necessarily be a constant, but could have evolved over time. Observationally, giants are considered rare in comparison to smaller radio galaxies. However, because giants are presumably strongly affected by surface brightness selection, this current-day observed rarity might not translate to an intrinsic rarity. Excitingly, by forward modelling selection effects — and in particular surface brightness selection — we can constrain the intrinsic comoving GRG number density between $z = 0$ and $z = z_\mathrm{max}$, which we denote simply by n_GRG.

The bottom-right one-dimensional marginal of Fig. 8.10 shows a strongly skewed distribution for n_GRG, with a marginal mean $\mathbb{E}[n_\mathrm{GRG}] = 13 \pm 10\ (100\ \mathrm{Mpc})^{-3}$ and a 95% probability that $n_\mathrm{GRG} > 4\ (100\ \mathrm{Mpc})^{-3}$. These number densities are a factor of order unity higher than those of Oei et al. (2023a), who inferred a marginal mean $\mathbb{E}[n_\mathrm{GRG}] = 4.6 \pm 2.4\ (100\ \mathrm{Mpc})^{-3}$ and a 90% probability that $n_\mathrm{GRG} < 6.7\ (100\ \mathrm{Mpc})^{-3}$.

The joint marginal distribution of n_GRG and $b_{\nu,\mathrm{ref}}$ reveals a strong inverse relationship, whose origin is easy to grasp. Models in which giants are relatively rare (i.e. with low n_GRG) but with relatively mild surface brightness selection (i.e. with high $b_{\nu,\mathrm{ref}}$) are about as successful in reproducing the data-derived projected length–redshift histogram as models in which giants are relatively common (i.e. with high n_GRG) but with relatively severe surface brightness selection (i.e. with low $b_{\nu,\mathrm{ref}}$). The narrowness of the joint distribution also suggests that, if estimates of $b_{\nu,\mathrm{ref}}$ would reveal it to be $\gtrsim 10\ \mathrm{Jy\ deg^{-2}}$, it should be possible to break the degeneracy and accurately determine n_GRG.

Recent work (Oei et al., in prep.) suggests that the comoving number density of luminous, non-giant radio galaxies (LNGRGs), understood to have radio luminosities at 150 MHz of $l_\nu \geq 10^{24}\ \mathrm{W\ Hz^{-1}}$ and projected lengths $l_\mathrm{p} < l_\mathrm{p,GRG} := 0.7$ Mpc, is

$n_{\mathrm{LNGRG}} = 12 \pm 1\,(100\,\mathrm{Mpc})^{-3}$. Our work suggests that giants might be comparably common. If this is indeed the case, then the widespread belief that giants form a rare population of radio galaxies must be revised.

8.5.3 GRG lobe volume-filling fraction

The present-day volume-filling fraction of the lobes of giants in clusters and filaments of the Cosmic Web, $\mathcal{V}_{\mathrm{GRG-CW}}(z = 0)$, is not a parameter of our model, but rather a derived quantity. As briefly discussed in Sect. 8.4.9, we compute its probability distribution using the parameter hexads that we have obtained by rejection sampling from the posterior.

For each sampled hexad, we compute $\xi(l_{\mathrm{p}})$ using $\xi(l_{\mathrm{p,GRG}})$, $\Delta\xi$, and Eq. 8.9, then $f_{L_{\mathrm{p}} \mid L_{\mathrm{p}} \geq l_{\mathrm{p,GRG}}}(l_{\mathrm{p}})$ using Eq. 8.10, and finally $\mathcal{V}_{\mathrm{GRG-CW}}(z = 0)$ using n_{GRG} and Eq. 8.25. This last step also requires an estimate of $\mathbb{E}[\Upsilon_{\mathrm{p}} \mid L_{\mathrm{p}} \geq l_{\mathrm{p,GRG}}]$, the expectation value of the ratio between the combined lobe volumes and cubed projected lengths of giants. Problematically, only few accurate data currently exist to estimate this quantity. In Oei et al. (2022a), the authors estimated that Alcyoneus, which measures $l_{\mathrm{p}} = 4.99 \pm 0.04$ Mpc, boasts a combined lobe volume $V = 2.5 \pm 0.3$ Mpc^3. Similarly, Oei et al. (2023b) estimated that the giant generated by NGC 6185, which measures $l_{\mathrm{p}} = 2.45 \pm 0.01$ Mpc, has a combined lobe volume $V = 0.35 \pm 0.03$ Mpc^3. These cases give $\Upsilon_{\mathrm{p}} \mid L_{\mathrm{p}} \geq l_{\mathrm{p,GRG}} = 2.0 \pm 0.2\%$ and $\Upsilon_{\mathrm{p}} \mid L_{\mathrm{p}} \geq l_{\mathrm{p,GRG}} = 2.4 \pm 0.2\%$, respectively. We note that these cases concern giants generated by an elliptical and a spiral galaxy, respectively. The highly preliminary sample mean thus is $\langle \Upsilon_{\mathrm{p}} \mid L_{\mathrm{p}} \geq l_{\mathrm{p,GRG}} \rangle = 2.2 \pm 0.2\%$, which we treat as an approximation of $\mathbb{E}[\Upsilon_{\mathrm{p}} \mid L_{\mathrm{p}} \geq l_{\mathrm{p,GRG}}]$. As in Oei et al. (2023a), we assume that clusters and filaments comprise 5% of the Local Universe's volume (Forero-Romero et al., 2009): $\mathcal{V}_{\mathrm{CW}}(z = 0) = 5\%$. Propagating all uncertainties, we obtain the posterior distribution over $\mathcal{V}_{\mathrm{GRG-CW}}(z = 0)$ shown in Fig. 8.11.

This probability distribution inherits its skewness from the skewed marginal of n_{GRG}. We find a posterior mean and standard deviation (SD) of $\mathcal{V}_{\mathrm{GRG-CW}}(z = 0) = 1.1 \pm 0.9 \cdot 10^{-5}$. This result appears statistically consistent with that of Oei et al. (2023a), who found $\mathcal{V}_{\mathrm{GRG-CW}}(z = 0) = 5^{+8}_{-2} \cdot 10^{-6}$. While this appears low at first sight, we speculate that these numbers are consistent with a scenario in which giants contribute significantly to cosmic magnetogenesis. To see why, we first note that the number of giants that have ever existed might exceed those that exist now by an order of magnitude. Second, a large fraction of the giants that existed throughout cosmic history might have lived at $z \gtrsim 1$, when the Universe's volume was an order of magnitude smaller. As a consequence, the instantaneous volume-filling fraction

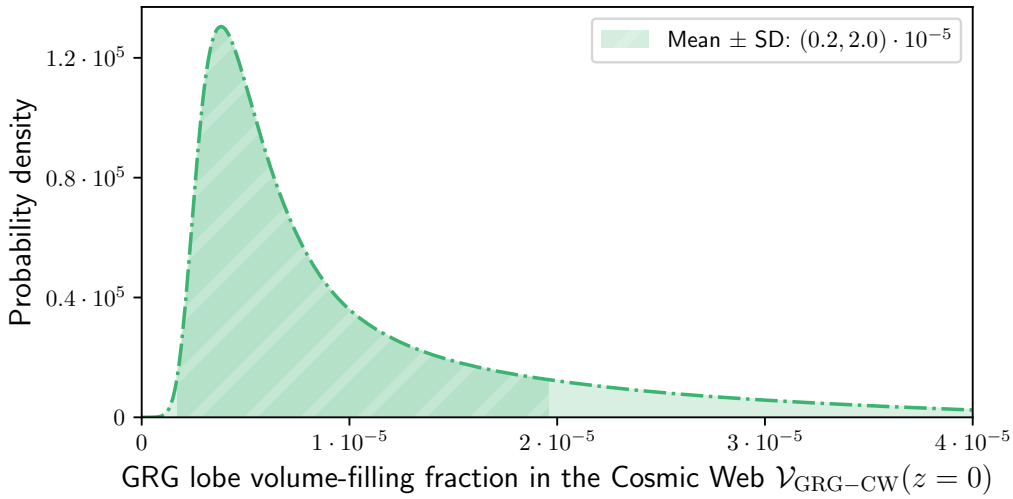


Figure 8.11: Posterior distribution for the instantaneous, current-day GRG lobe volume-filling fraction in clusters and filaments of the Cosmic Web, $\mathcal{V}_{\rm GRG-CW}(z = 0)$.

$\mathcal{V}_{\rm GRG-CW}(z)$ would have been an order of magnitude larger — at least, if the instantaneous comoving number density and distributions over proper length and shape remain roughly constant over time. Third, buoyant lobes might deposit magnetic fields in their wake, while diffusion might have spread the magnetic fields of GRG lobes further through the IGM. Taken together, these three effects could render the current-day VFF of magnetic fields that were once contained in the lobes of giants higher than $\mathcal{V}_{\rm GRG-CW}(z = 0)$ by several (i.e. three or more) orders of magnitude. This, in turn, suggests a significant astrophysical seeding potential. For instance, assuming four orders of magnitude, $\sim$10% of the volume of today's Cosmic Web could have been magnetised by giants.

We finally point out that giant-induced IGM magnetic fields could have strengths consistent with observational constraints. At the moment, the lowest magnetic field strengths measured in giant radio galaxy lobes, as inferred from images of Alcyoneus and the giant generated by NGC 6185 assuming the equipartition or minimum energy condition, are 400–500 nG (Oei et al., 2022a, 2023b). If such field strengths would be typical, and buoyancy and diffusion lowers the density of field lines by an order of magnitude, then the typical giant-induced IGM field strength would be $\sim$10 nG. This is in agreement with recent radio estimates and limits (e.g. Table 1 of Vazza et al., 2021a). We note that this argument ignores possibly significant amplification and decay mechanisms, such as turbulent amplification and decay.

8.6 Discussion

Below, we discuss how our ML pipeline and GRG population inference compare to earlier work.

8.6.1 Comparison with previous ML GRG search techniques

Proctor (2016) applied an ML approach to search for GRG candidates by looking for likely pairs of (unresolved) radio lobes with the required angular length in the NRAO VLA Sky Survey (NVSS; Condon et al., 1998). For this radio source component association problem, Proctor (2016) trained an oblique classifier (a type of decision tree ensemble; Murthy et al., 1993), using six source finder–derived features on $51,195$ pairs of radio components, 48 of which were verified giants. This method proved to be useful under the assumption that giants generally appear as an isolated pair of unresolved radio blobs, which is the case for NVSS with its $45''$ resolution and 450 μJy beam^{-1} sensitivity. Dabhade et al. (2020a) visually inspected the $1,600$ GRG candidates presented by Proctor (2016) and confirmed 151 giants, which implies a 9% precision for the GRG candidate predictions. However, Proctor (2016) expect that giants with resolved lobes — which rule-based source finders often incorrectly break down into multiple separate sources — require a different approach, and we would like the reader to note that virtually all GRG lobes in LoTSS are resolved.[16] It works in our favour that the convolution neural network in our ML pipeline (Sect. 8.4.4) was specifically designed to use the morphology of the resolved, extended emission as a cue for the radio source component association. Furthermore, as the source suggestions from our ML pipeline include optical host identifications, the candidates that we inspected not only have the required angular length but also have a host galaxy and corresponding redshift estimate assigned. This allows us to visually inspect only those radio sources that fulfil the projected proper length $l_{\mathrm{p,GRG}} := 0.7$ Mpc requirement. Overall, our ML pipeline has a precision of 47% for the GRG candidates that it suggests.

8.6.2 Comparison with previous inference strategies

Compared to the approach of Oei et al. (2023a), our approach makes better use of the redshift information available for each giant. More specifically, we use the redshifts to make a 'redshift-resolved' observed projected length histogram, while Oei et al. (2023a) only compared a 'redshift-collapsed' distribution of observed projected

[16] As all giants have angular lengths $\varphi \geq 1.3'$, they cover at least 13 LoTSS $6''$ beams. This suggests that a single GRG lobe will cover multiple beams, too.

lengths to forward model predictions of $L_{\mathrm{p,obs}} \mid L_{\mathrm{p,obs}} \geq l_{\mathrm{p,GRG}}$. Effectively, Oei et al. (2023a) thus used for each giant only *Boolean* redshift information, $\mathbb{I}(z_i < z_{\max})$: that is, a truth value indicating whether or not the giant with index i resides at a redshift below the maximum considered value.

In addition, our work changed the comoving number density of giants, n_{GRG}, from a derived quantity to a model parameter. This approach acknowledges the fact that the observed number of giants, either for a specific projected length–redshift bin or for the parameter space in its entirety, scales linearly with n_{GRG} (if the selection effects remain the same). Thus, there is intrinsic population information contained in the observed *number* of giants. However, by comparing predicted and observed *probability distributions* only, Oei et al. (2023a) did not exploit this fact.

8.6.3 Future work

With the advent of large-scale, sensitive, low-frequency sky surveys such as the LoTSS, the Evolutionary Map of the Universe survey (EMU; Norris et al., 2011), and the arrival of next-generation instruments such as the SKA (Dewdney et al., 2009) and the DSA-2000 (e.g. Hallinan et al., 2019; Connor et al., 2022) later this decade, opportunities shall arise to detect many more giants than have been found hitherto. It is therefore likely that automated approaches to giant finding and host association will become only more relevant in the future.

Regarding our own machine learning–based pipeline, there is significant room to improve both the radio component association and the host association. Visual inspection indicated a precision of 47% and the empirically determined $p_{\mathrm{obs,ID}}$ in Fig. 8.9 showed that even in combination with the LGZ sample, the ML pipeline recall does not surpass 70%. Sensible paths to improve the radio component association within the ML pipeline architecture include switching from rectangular bounding box–based object detection (the Fast R-CNN used in this article) to pixel-based instance segmentation and using a larger convolutional backbone (e.g. Liu et al., 2022; Wright et al., 2010) or a transformer-based backbone (e.g. Liu et al., 2021; Zhang et al., 2022; Li et al., 2022). Mostert et al. (2022) conclude that a larger convolutional neural network is not effective unless one also significantly increases the quantity of high-quality training data, and in general, transformers require even more training data than convolutional neural networks (e.g. Wang et al., 2022). To that extent, adding a filtered version[17] of the available LoTSS DR2 LGZ annotations (Hardcastle et al., 2023) to the training data can be considered. Furthermore, assembling a joined

[17] For example, by identifying a handful of very active and expert volunteers and increasing the weight of their votes.

dataset encompassing the (labelled) survey data of other low frequency radio telescopes can be considered. Pre-training on this dataset can benefit radio source component association, host identification and morphological classification tasks across the board.

Finally, there appear to be clear opportunities to make the population-based forward model presented in Sect. 8.2 more accurate. For example, at present, we have neglected photometric redshift uncertainties; however, the consequences of these uncertainties appear perfectly possible to forward model. One such currently ignored consequence is Eddington bias: as RGs with projected lengths $l_{\mathrm{p}} = 0.6$ Mpc are intrinsically more common than those with projected lengths $l_{\mathrm{p}} = 0.8$ Mpc, redshift error–induced projected length errors have the net effect of falsely raising the number of supposed giants with projected lengths near $l_{\mathrm{p,GRG}} := 0.7$ Mpc. This effect could contaminate the inference of $\xi(l_{\mathrm{p,GRG}})$. Somewhat more challenging, but plausibly of greater value, would be a further exploration of how surface brightness selection is effectively modelled. A major focus of such an exploration would be to analyse the surface brightness characteristics of hitherto discovered giants. As the masked cutouts of Fig. 8.3 suggest, the machine learning–based pipeline described in this work offers the exciting potential to amass — fully automatically — surface brightness properties for thousands of giants. The availability of such properties for a large fraction of observed giants also allows one to compare the forward model's predictions with an observed projected length–redshift–surface brightness histogram, rather than with an observed projected length–redshift histogram only. It is highly likely that adding another dimension to the data yields tighter parameter constraints. To make the identification probability functions of Fig. 8.9 more accurate, it appears promising to have an expert visually (and exhaustively, i.e. without imposing angular length thresholds) comb through a small representative region of LoTSS DR2 in search of giants. The resulting dataset would provide a better basis for determining the identification probability functions than the ML–LGZ or Oei et al. (2023a) datasets used in this work. We note that the Boötes LOFAR Deep Field search of Simonte et al. (2022) does not appear suited for this purpose, as the increased depth of this field renders it unrepresentative of LoTSS DR2 as a whole. Finally, the model could be expanded in an attempt to measure cosmological evolution of, for example, n_{GRG}. However, we note that adding additional parameters to the model necessitates adopting more efficient inference techniques, such as Markov chain Monte Carlo or nested sampling. The associated numerical gain would, in part, be negated by losing the speed-up associated to the likelihood trick of Appendix 8.A2.

Currently, a major uncertain factor in the determination of $\mathcal{V}_{\mathrm{GRG-CW}}(z = 0)$ is the value of $\mathbb{E}[\Upsilon_{\mathrm{p}} \mid L_{\mathrm{p}} \geq l_{\mathrm{p,GRG}}]$. To improve this situation, we recommend ex-

panding the capabilities and automating the parametric Bayesian lobe volume estimation method introduced by Oei et al. (2022a, 2023b). This method could then be applied to thousands of our ML pipeline's masked cutouts, such as the one in Fig. 8.3. This effort would increase the number of giants on which our estimate of $\mathbb{E}[\Upsilon_{\mathrm{p}} \mid L_{\mathrm{p}} \geq l_{\mathrm{p,GRG}}]$ is based by several (i.e. two or three) orders of magnitude.

8.7 Conclusions

In this work, we concatenated an existing crowd-sourced radio–optical catalogue, a new ML pipeline to automate radio–optical catalogue creation, and a Bayesian forward model to build a next-generation giant radio galaxy discovery and characterisation machine. Applying this setup to the LOFAR Two-metre Sky Survey, we uncovered thousands of previously unknown giants, confirmed thousands of GRG candidates, and constrained the properties of the underlying population.

1. The LoTSS is an on-going sensitive, high-resolution, low-frequency radio survey whose second data release (DR2) covers 27% of the Northern Sky. As the number of detected sources already ranges in the millions, it has become unfeasible (at least for small scientific teams) to conduct manual, visual searches for giants — in particular for those with angular lengths close to the lower limit of 1.3′.

2. To address this challenge, we scanned all 841 LoTSS DR2 pointings — which together cover more than five thousand square degrees of Northern Sky — with an ML pipeline that crucially includes the convolutional neural network of Mostert et al. (2022), designed for the association of radio components for highly resolved radio galaxies, and an adapted version of the automated optical host galaxy identification heuristic developed by Barkus et al. (2022). Used as a GRG detection system, our ML pipeline has a precision of 47%, a significant improvement over the 9% precision obtained using the previous state-of-the-art ML GRG detection model (Proctor, 2016; Dabhade et al., 2020a). We merged the resulting giant candidate sample with that of the LGZ citizen science campaign (Hardcastle et al., 2023), homogenised the angular lengths, and subjected the candidates to a visual, expert quality check. The result is a sample of more than eight thousand newly confirmed giants, of which a large fraction is considered genuine beyond reasonable doubt. More than 10^4 unique giants are now known to the literature.

3. We expand the population-based statistical forward model of Oei et al. (2023a) aimed at constraining the geometric properties of giants. In particular, by

modelling the PDF of the radio galaxy projected length RV L_p as a curved power law, we automatically also model the PDF of the *giant* radio galaxy projected length RV $L_\mathrm{p} \mid L_\mathrm{p} \geq l_\mathrm{p,GRG}$ as a curved power law. We assume that these projected length distributions do not undergo intrinsic evolution between cosmological redshift $z = z_\mathrm{max}$ and $z = 0$, and likewise assume an intrinsically constant comoving GRG number density throughout this redshift range. We model surface brightness selection by assuming a lognormal lobe surface brightness distribution at the survey's central frequency ν_obs, valid for radio galaxies of intrinsic proper length l_ref at redshift $z = 0$. We relate lobe surface brightness distributions for radio galaxies of other lengths and at other redshifts to this reference distribution. In addition, we model selection caused by the imperfect ability of search methods to identify all in principle identifiable giants. For this purpose, we use logistic functions of projected length l_p and redshift z.

4. We then sought to identify all model parameter hexads that can reproduce the projected length–redshift histogram of the joint ML–LGZ–Oei et al. (2023a) LoTSS DR2 GRG sample. Through a simple Poissonian likelihood and a uniform prior distribution, we constructed a posterior distribution over the model parameters. By confronting the model with an observed projected length–redshift histogram, rather than with an observed projected length distribution only (as has been done in Oei et al. (2023a)), we obtain tighter parameter constraints.

5. We find evidence for the claim that the projected lengths of giant radio galaxies follow a curved power law PDF, whose tail index equals $\xi(l_\mathrm{p,GRG}) = -2.8 \pm 0.2$ at $l_\mathrm{p,1} = l_\mathrm{p,GRG} := 0.7$ Mpc and increases by $\Delta\xi = -2.4 \pm 0.3$ (i.e. decreases by 2.4 ∓ 0.3) in the projected length interval leading up to $l_\mathrm{p,2} = 5$ Mpc. The predicted median lobe surface brightness at $\nu_\mathrm{obs} = 150$ MHz, $l_\mathrm{ref} = 0.7$ Mpc, and $z = 0$ equals $b_{\nu,\mathrm{ref}} = 30 \pm 20$ Jy deg^{-2}. This surface brightness level is lower than previously thought. Tight degeneracies resembling inverse relations exist between $b_{\nu,\mathrm{ref}}$ and the reference surface brightness dispersion measure σ_ref, and between $b_{\nu,\mathrm{ref}}$ and the GRG number density n_GRG. The latter relation suggests that giant radio galaxies might be more common than previously thought. At $n_\mathrm{GRG} = 13 \pm 10\ (100\ \mathrm{Mpc})^{-3}$, giant radio galaxies appear to be of an abundance comparable to that of luminous *non*-giant radio galaxies. Strikingly, we conclude that, at any moment in time, a significant fraction of the radio galaxy population is in a GRG phase. As an immediate consequence, the fraction of radio galaxies that end their lives as giants must

be even higher.

6. Finally, we generate a posterior distribution for the instantaneous volume-filling fraction of GRG lobes in clusters and filaments of the Cosmic Web, $\mathcal{V}_{\mathrm{GRG-CW}}(z = 0)$ — a key statistic required for determining the cosmic magnetogenesis potential of giants. We find $\mathcal{V}_{\mathrm{GRG-CW}}(z = 0) = 1.1 \pm 0.9 \cdot 10^{-5}$. If a giant population similar to that in the Local Universe has existed for most of the Universe's lifetime, and IGM mixing processes are significant, then it appears possible that magnetic fields originating from giants permeate significant ($\sim$10%) fractions of today's Cosmic Web.

Using modern automation and inference techniques — that still leave much room for future improvements — we have conducted the most detailed study yet of the abundance and geometry of giant radio galaxies. These cosmic colossi may provide a previously underappreciated contribution to astrophysical magnetogenesis.

The full GRG catalogue with host identifications and the Stokes-I cutouts containing the segmented giants will soon be available on Zenodo. M.S.S.L. Oei and R.J. van Weeren acknowledge support from the VIDI research programme with project number 639.042.729, which is financed by the Dutch Research Council (NWO). B. Barkus is grateful for support from the UK STFC. L. Alegre is grateful for support from the UK STFC via CDT studentship grant ST/P006809/1. M.J. Hardcastle acknowledges support from the UK STFC [ST/V000624/1]. We like to thank Huib Intema for enabling the cross-institute collaboration on the Leiden Observatory computer infrastructure. We like to thank Frits Sweijen for coding the very useful `https://github.com/tikk3r/legacystamps`. This research has made use of the Python `astropy` package (The Astropy Collaboration et al., 2018); the VizieR catalogue access tool (Ochsenbein et al., 2000), CDS, Strasbourg, France (DOI: 10.26093/cds/vizier); and the 'Aladin Sky Atlas' developed at CDS, Strasbourg Observatory, France (Bonnarel et al., 2000; Boch & Fernique, 2014). LOFAR data products were provided by the LOFAR Surveys Key Science project (LSKSP; `https://lofar-surveys.org/`) and were derived from observations with the International LOFAR Telescope (ILT). LOFAR (van Haarlem et al., 2013) is the Low Frequency Array designed and constructed by ASTRON. It has observing, data processing, and data storage facilities in several countries, which are owned by various parties (each with their own funding sources), and which are collectively operated by the ILT foundation under a joint scientific policy. The efforts of the LSKSP have benefited from funding from the European Research Council, NOVA, NWO, CNRS-INSU, the SURF Co-operative, the UK Science and Technology Funding Council and the Jülich Supercomputing Centre. This publication uses data generated via the Zooniverse.org platform, development of which is funded by generous support, including a Global Impact Award from Google, and by a grant from the Alfred P. Sloan Foundation.

Author contributions

Rafaël and Martijn together came up with the idea of the study. Rafaël built and ran the machine learning–accelerated pipeline, merged the results with those of the LGZ pipeline, and visually judged all GRG candidates. Rafaël also generated the final GRG catalogue. Martijn developed the forward model, ran it on the data, and inferred the GRG lobe VFF. Rafaël and Martijn jointly wrote the article.

8.A1 Curved power law PDF for L

In Sect. 8.2.3, we have started modelling the geometry of radio galaxies at the level of the projected proper length RV L_p. While algebraically easier — when curved power laws are considered, at least — this approach is more limited than starting the forward model at the level of the intrinsic proper length RV L. In this appendix, we calculate the distribution of L_p upon modelling L with a curved power law.

Let us assume that, for $l \geq l_\mathrm{min}$,

$$f_L(l) \propto \left(\frac{l}{l_\mathrm{min}}\right)^{\xi(l)}, \tag{8.35}$$

where $\xi(l) = al + b$. We now use the identity that for $f(x) = \left(\frac{x}{c}\right)^{ax+b}$, one finds

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x} = \left(\frac{x}{c}\right)^{ax+b}\left(a\ln\frac{x}{c} + a + \frac{b}{x}\right) = f(x)\left(a\ln\frac{x}{c} + a + \frac{b}{x}\right). \tag{8.36}$$

Therefore,

$$\frac{\mathrm{d}f_L(l)}{\mathrm{d}l} = f_L(l)\left(a\ln\frac{l}{l_\mathrm{min}} + a + \frac{b}{l}\right), \tag{8.37}$$

and

$$\frac{\mathrm{d}f_L(l_\mathrm{p}\eta)}{\mathrm{d}l_\mathrm{p}} = f_L(l_\mathrm{p}\eta)\left(a\ln\frac{l_\mathrm{p}\eta}{l_\mathrm{min}} + a + \frac{b}{l_\mathrm{p}\eta}\right)\eta. \tag{8.38}$$

Thus, finding the PDF of L_p requires calculating three different integrals over η:

$$\begin{aligned} f_{L_\mathrm{p}}(l_\mathrm{p}) = &-(1+b)\int_1^\infty \sqrt{1-\frac{1}{\eta^2}}\, f_L(l_\mathrm{p}\eta)\,\mathrm{d}\eta \\ &- l_\mathrm{p}a\left(1+\ln\frac{l_\mathrm{p}}{l_\mathrm{min}}\right)\int_1^\infty \sqrt{\eta^2-1}\, f_L(l_\mathrm{p}\eta)\,\mathrm{d}\eta \\ &- l_\mathrm{p}a\int_1^\infty \sqrt{\eta^2-1}\, f_L(l_\mathrm{p}\eta)\ln\eta\,\mathrm{d}\eta \quad \text{for } l_\mathrm{p} > l_\mathrm{min}. \end{aligned} \tag{8.39}$$

The PDF of $L_\mathrm{p} \mid L_\mathrm{p} \geq l_\mathrm{p,GRG}$ follows through Eq. 8.3.

8.A2 Likelihood trick

Thanks to its Poissonian form, there exists a particularly numerically efficient way of computing the likelihood presented in Sect. 8.2.8 as a function of n_{GRG}, for fixed values of the other parameters. Defining

$$A(\theta) := \sum_{i=1}^{N_{\mathrm{b}}} N_i \ln \lambda_i(\theta) \quad \text{and} \quad B(\theta) := \sum_{i=1}^{N_{\mathrm{b}}} \lambda_i(\theta), \tag{8.40}$$

one interested in the log-likelihood up to a constant only needs to compute

$$\ell(\theta) := \ln \mathcal{L}(\{N_i\} \mid \theta) + \sum_{i=1}^{N_{\mathrm{b}}} \ln (N_i!) = A(\theta) - B(\theta). \tag{8.41}$$

The quantity $B(\theta)$ has a simple interpretation: it is the total number of giants expected to be observed under θ within the entire projected length–redshift parameter space considered.

How does ℓ change upon changing n_{GRG}? When $n_{\mathrm{GRG}} \mapsto a \cdot n_{\mathrm{GRG}}$, $\lambda_i \mapsto a \cdot \lambda_i$, so that

$$\begin{aligned} \ell(n_{\mathrm{GRG}}) &\mapsto \sum_{i=1}^{N_{\mathrm{b}}} N_i \ln (a \cdot \lambda_i) - a \cdot \lambda_i \\ &= A(n_{\mathrm{GRG}}) - a \cdot B(n_{\mathrm{GRG}}) + \ln a \cdot \sum_{i=1}^{N_{\mathrm{bins}}} N_i. \end{aligned} \tag{8.42}$$

(In the notation $\ell(n_{\mathrm{GRG}})$, $A(n_{\mathrm{GRG}})$, and $B(n_{\mathrm{GRG}})$, we suppress the dependence on the other five parameters.) We conclude that, when n_{GRG} increases by a factor a, the A-term in ℓ remains the same, the B-term in ℓ becomes a factor a bigger, and an extra factor emerges: namely, the product of $\ln a$ and the total number of giants in the dataset.

The significance of this result is that, once A and B are known at some reference number density $n_{\mathrm{GRG,ref}}$, we can rapidly evaluate ℓ for any other number density. In this work, we implement this 'likelihood trick' by evaluating ℓ for two different values of n_{GRG} (and for many different values of the other parameters). We then solve for

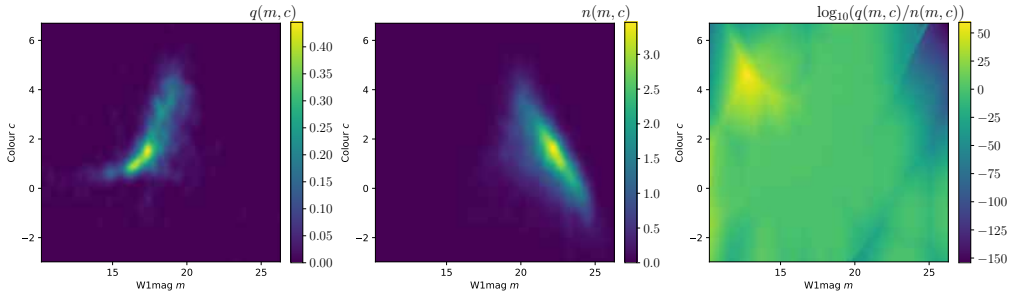


Figure 8.12: Unregularised KDE estimates for q in the left panel, n in the second panel, and q/n with logarithmic colour bar in the third panel. The KDE bandwidth of 0.2 stems from Barkus et al. (2022).

$A(n_{\mathrm{GRG,ref}})$ and $B(n_{\mathrm{GRG,ref}})$, and use

$$\ell(n_{\mathrm{GRG}}) = A(n_{\mathrm{GRG,ref}}) - \frac{n_{\mathrm{GRG}}}{n_{\mathrm{GRG,ref}}} \cdot B(n_{\mathrm{GRG,ref}}) + \ln \frac{n_{\mathrm{GRG}}}{n_{\mathrm{GRG,ref}}} \cdot \sum_{i=1}^{N_{\mathrm{b}}} N_i. \tag{8.43}$$

8.A3 PyBDSF parameters

As described in Sect. 8.4.1, the GRG detection pipeline uses PyBDSF for the initial radio blob detection. For reproducibility, we provide the specific parameters used, which we adopted from Shimwell et al. (2022):

```
bdsf.process_image(<filename>,thresh_isl=4.0,
thresh_pix=5.0, rms_box=(150,15), rms_map=True,
mean_map='zero', ini_method='intensity',
adaptive_rms_box=True, adaptive_thresh=150,
rms_box_bright=(60,15), group_by_isl=False,
group_tol=10.0, output_opts=True, atrous_do=True,
atrous_jmax=4, flagging_opts=True,
flag_maxsize_fwhm=0.5, advanced_opts=True,
blank_limit=None, frequency=143.65e6)
```

8.A4 Adaptations of the radio ridgeline based host galaxy identification

Here we elaborate on two small adaptations of the radio-optical crossmatch method introduced by Barkus et al. (2022). First, we explicitly regularised $q(m,c)$ and $n(m,c)$.

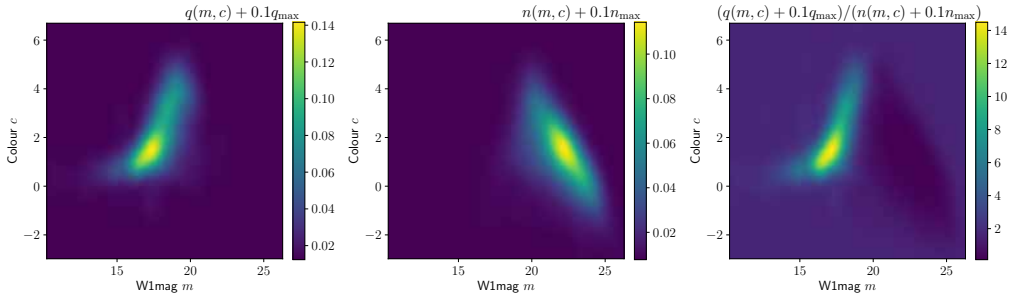


Figure 8.13: Regularised KDE estimates for q in the left panel, n in the second panel, and q/n with logarithmic colour bar in the third panel. The KDE bandwidth of 0.4 stems from 10-fold cross-validation.

Figure 8.12 shows that the unregularised forms of q and n can take on extreme values in the LR (eq. 8.31) in sparsely sampled regions of the (m, c)–parameter space. The 2D KDE that models $q(m, c)$ was fitted on the m and c values of all 905 sources with an angular length $\varphi > 1'$ from 40 randomly picked LoTSS DR2 pointings. The 2D KDE that models $n(m, c)$ was fitted on the m and c values of 10, 000 sources that were randomly sampled from the entire combined infrared–optical catalogue. By simply adding a small constant factor to $q(m, c)$ and $n(m, c)$ we get more robust LR values, see Fig. 8.13. We added a constant factor $0.1 \cdot q_{\max}$ and $0.1 \cdot n_{\max}$ for q and n respectively, where $q_{\max}$ is the maximum of the KDE for q and $n_{\max}$ is the maximum of the KDE for n. We set the KDE bandwidths to 0.4 following a 10-fold cross-validation.

Second, we changed the form of $f(r)$. Theoretically, we might expect both the distance between the 'true' optical counterpart and the radio ridgeline $r_{\text{opt,ridge}}$ and the distance between the 'true' optical counterpart and the radio centroid $r_{\text{opt,centroid}}$ to be Rayleigh distributed.[18] However, as Fig. 8.14 demonstrates, the lognormal distribution clearly provides the best empirical fit to the distances. The figure shows a histogram of the distance measures for radio sources to their optical counterpart as manually identified through LGZ. Specifically, we plot the distances for the same 905 radio sources, with an angular length $\varphi > 1'$, from 40 randomly selected paintings as above. Thus we update $f(r)$ to be:

$$f(r_{\text{mean}}) = \frac{1}{r_{\text{mean}}\sigma\sqrt{2\pi}} e^{\left(-\frac{(\ln r_{\text{mean}} - \mu)^2}{2\sigma^2}\right)}, \tag{8.44}$$

[18] In two dimensions, the Euclidean distance between the origin and a point whose Cartesian coordinates are independent, zero-mean, and equal-variance normal random variables, is Rayleigh distributed. This motivates modelling the angular distance between the optical counterpart and the radio centroid with a Rayleigh distribution. The appropriate value of the distribution's parameter likely depends (positively) on the angular length of the radio source considered; as such, one would not expect a single Rayleigh distribution to work for all radio sources.

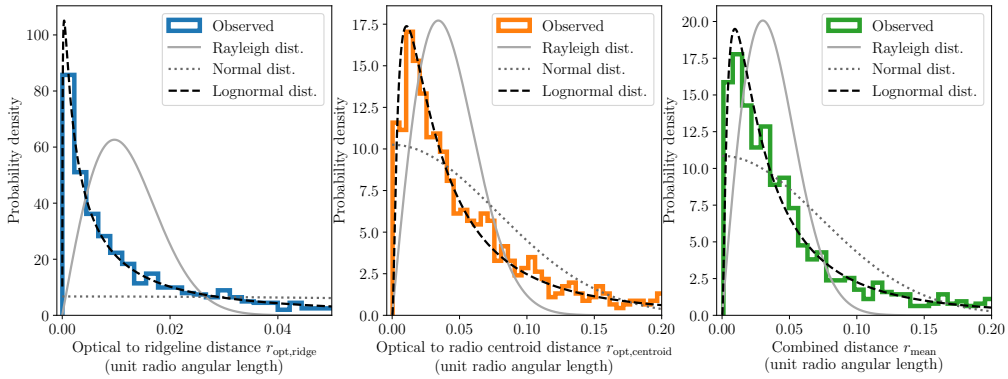


Figure 8.14: Each panel shows the histogram of a different distance measure between 905 radio sources with $\varphi > 1'$ and their optical host. The grey, dark grey, and black lines show empirical fits to these histograms for Rayleigh, normal, and lognormal distributions respectively. The tails of the histograms are long, for visualisation purposes we only plot the x-axis up to 0.05 and 0.20.

where we empirically determine σ and μ using our sample of 905 radio-sources,

$$\mu = \frac{\sum_i \ln r_{\mathrm{mean},i}}{n} = -3.37 \tag{8.45}$$

and

$$\sigma^2 = \frac{\sum_i (\ln r_{\mathrm{mean},i} - \mu)^2}{n} = 1.28, \tag{8.46}$$

with $n = 905$ the size of our sample.

8.A5 Sky coverages

As an extension of Sect. 8.4.9, this appendix details the sky coverages of our analyses. In particular, Table 8.3 provides a decomposition — in terms of disjoint spherical quadrangles — of the sky coverage common between the ML pipeline, LGZ, and the combined manual search of Dabhade et al. (2020b) and Oei et al. (2023a). For simplicity, and as an acknowledgement of the wiggle room inherent to defining this joint sky coverage, we chose integer coordinates. Together, these spherical quadrangles cover $\Omega = 5327.9\ \mathrm{deg}^2 = 1.62\ \mathrm{sr}$ (25.8%) of the Northern Sky. We shall refer to this coverage simply as the 'LoTSS DR2 coverage'.

The ML–LGZ–Oei et al. (2023a) overlap region amounts to the LoTSS DR2 coverage with the LoTSS DR1 spherical quadrangle removed. The minimum and maximum right ascensions of this quadrangle are $\alpha_{\mathrm{min}} = 160°$ and $\alpha_{\mathrm{max}} = 230°$, while its

minimum and maximum declinations are $\delta_{\rm min} = 45^\circ$ and $\delta_{\rm max} = 56^\circ$. This smaller overlap region covers 4838.9 deg^2 (23.5%) of the Northern Sky. It is the sky coverage relevant to estimating the identification probability functions of Sect. 8.4.9 and Fig. 8.9: $p_{\rm obs,ID,1}(l_{\rm p}, z)$, $p_{\rm obs,ID,2}(l_{\rm p}, z)$, and $p_{\rm obs,ID}(l_{\rm p}, z)$.

Table 8.3: Sky coordinates and solid angles of disjoint spherical quadrangles whose union forms the LoTSS DR2 sky coverage — over which we have performed our inference. For each spherical quadrangle, we provide the minimum and maximum right ascension, $\alpha_{\rm min}$ and $\alpha_{\rm max}$, the minimum and maximum declination, $\delta_{\rm min}$ and $\delta_{\rm max}$, and its solid angle, Ω. We list the largest quadrangles first. The second and third object touch along the 360°–0° right ascension coordinate discontinuity, and could be viewed as a single quadrangle.

$\alpha_{\rm min}$ (°)	$\alpha_{\rm max}$ (°)	$\delta_{\rm min}$ (°)	$\delta_{\rm max}$ (°)	Ω (deg^2)
120	253	28	69	3536.7
0	35	16	35	597.5
338	360	16	35	375.6
253	269	28	47	240.1
109	120	25	41	147.1
269	277	31	47	99.2
330	338	17	30	95.2
191	210	23	28	85.7
35	41	24	32	42.3
253	260	58	69	34.3
120	131	25	28	29.5
277	281	41	47	17.3
327	330	17	20	8.5
277	280	32	35	7.5
117	120	53	57	6.9
260	264	66	69	4.6

A

Gaussian random field ionosphere model extension: the curved Earth

In Chapter 2, we considered, at one instant of time, a thick, single-layered ionosphere whose free electron density (FED) n_e is a Gaussian random field. For algebraic and numerical simplicity, we took the layer to be parallel to a *flat* Earth. For this configuration, we calculated the differential total electron content (DTEC) $\Delta\tau$ covariance function, and showed its superiority over ad-hoc covariance functions in DTEC Gaussian process regression (GPR).

For radio interferometers with long baselines, this 'flat Earth' approximation becomes coarse. To avoid associated errors in DTEC GPR, this appendix generalises Chapter 2's model to take proper account of the (approximate) sphericity of the Earth.[1]

We now assume the ionosphere to be a *spherical shell* with thickness b, centered around some point $\mathbf{x}_c \in \mathbb{R}^3$. We require the planet's rotational axis to also pass through this point. Although we specify the ionosphere's geometry, in general we need not impose a requirement on Terra's shape itself — it might be perfectly spherical, slightly ellipsoidal, or spherical with plate tectonics–induced surface height variations, for ex-

[1] The model thus becomes incompatible with flat Earth beliefs — whose modern adherents constitute a *global* movement depicting, in my opinion, the human intellect in a rather *unflattering* light.

ample. More relevant to the model are the locations of the antennae — or, in the LOFAR context, *stations* — $\mathbf{x}_i \in \mathbb{R}^3$.[2] The height of the ionosphere is defined such that the middle of the layer lies a distance a above the reference station at $\mathbf{x}_0 \in \mathbb{R}^3$. Let A denote the height of the middle of the ionospheric shell above $\mathbf{x}_c$; thus

$$A := ||\mathbf{x}_0 - \mathbf{x}_c||_2 + a. \tag{A.1}$$

A point $\mathbf{y} \in \mathbb{R}^3$ lies in the shell as long as

$$A - \frac{b}{2} < ||\mathbf{y} - \mathbf{x}_c||_2 < A + \frac{b}{2}. \tag{A.2}$$

Let us suppose that $\mathbf{y}$ lies on a line through $\mathbf{x}_i$ heading in skybound direction $\hat{\mathbf{k}} \in \mathbb{S}^2$. A parametrisation of this line is $\mathbf{y}(s) = \mathbf{x}_i + \hat{\mathbf{k}}s$, with $s \in \mathbb{R}$. The line segment that lies fully within the shell is demarcated by the two values $s_i^{\pm} \in \mathbb{R}_{>0}$ such that

$$\begin{aligned} ||\mathbf{y}(s_i^{\pm}) - \mathbf{x}_c||_2 &= A \pm \frac{b}{2}, \text{ or} \\ ||\mathbf{x}_i + \hat{\mathbf{k}}s_i^{\pm} - \mathbf{x}_c||_2 &= A \pm \frac{b}{2}. \end{aligned} \tag{A.3}$$

Squaring both sides, and writing the result as a second-degree polynomial equation in $s_i^{\pm}$, we find

$$\left(s_i^{\pm}\right)^2 + 2\hat{\mathbf{k}} \cdot (\mathbf{x}_i - \mathbf{x}_c)\, s_i^{\pm} + ||\mathbf{x}_i - \mathbf{x}_c||_2^2 - \left(A \pm \frac{b}{2}\right)^2 = 0. \tag{A.4}$$

By the quadratic formula, we find

$$\begin{aligned} s_i^{\pm} &= -\hat{\mathbf{k}} \cdot (\mathbf{x}_i - \mathbf{x}_c) \boxed{\pm} \sqrt{\left(\hat{\mathbf{k}} \cdot (\mathbf{x}_i - \mathbf{x}_c)\right)^2 - ||\mathbf{x}_i - \mathbf{x}_c||_2^2 + \left(A \pm \frac{b}{2}\right)^2} \\ &= \left|\hat{\mathbf{k}} \cdot (\mathbf{x}_i - \mathbf{x}_c)\right| \left(-\operatorname{sgn}\left(\hat{\mathbf{k}} \cdot (\mathbf{x}_i - \mathbf{x}_c)\right) \boxed{\pm} \sqrt{1 + \frac{\left(A \pm \frac{b}{2}\right)^2 - ||\mathbf{x}_i - \mathbf{x}_c||_2^2}{\left(\hat{\mathbf{k}} \cdot (\mathbf{x}_i - \mathbf{x}_c)\right)^2}} \right). \end{aligned} \tag{A.5}$$

A prima facie, the second line seems more complicated than the first. However, the expression is insightful.

[2] Naturally, we assume the stations to lie underneath the ionosphere — in other words, *enclosed* by the ionospheric shell.

First, we remark that $A \pm \frac{b}{2} > ||\mathbf{x}_i - \mathbf{x}_c||_2$ for all $\mathbf{x}_i$, as we have assumed the stations to lie underneath the ionosphere.

Furthermore, we remark that for all practical situations, $\hat{\mathbf{k}} \cdot (\mathbf{x}_i - \mathbf{x}_c) > 0$: the vector pointing from the centre of the shell to the station with index i lies in the same *hemisphere* as the observing direction. The reason is that Terra is approximately spherical, and that observers always make sure that sources stay above the local horizon with sufficient elevation during observing runs. In case of a perfectly spherical Earth with $\mathbf{x}_c$ as its centre, and stations put on top of its smooth surface, this has a neat geometrical interpretation. Under this idealisation, $\hat{\mathbf{k}} \cdot (\mathbf{x}_i - \mathbf{x}_c) > 0$ whenever the source is above the local horizon — and thus visible! As a consequence, $-\mathrm{sgn}(\hat{\mathbf{k}} \cdot (\mathbf{x}_i - \mathbf{x}_c)) = -1$ in practice.

Third, because $\hat{\mathbf{k}}$ points from Terra towards the *Great Unknown*, both $s_i^{\pm}$ are positive.

In order for the second line of Eq. A.5 to yield positive values, we need the $\boxed{+}$ of $\boxed{\pm}$. The equation reduces to

$$s_i^{\pm} = \left|\hat{\mathbf{k}} \cdot (\mathbf{x}_i - \mathbf{x}_c)\right| \left(\sqrt{1 + \frac{\left(A \pm \frac{b}{2}\right)^2 - ||\mathbf{x}_i - \mathbf{x}_c||_2^2}{\left(\hat{\mathbf{k}} \cdot (\mathbf{x}_i - \mathbf{x}_c)\right)^2}} - 1 \right). \tag{A.6}$$

Using the *first* line of Eq. A.5, we find the path length through the shell — for a station at $\mathbf{x}_i$ observing in direction $\hat{\mathbf{k}}$ — to be

$$\Delta s_i := s_i^+ - s_i^- \tag{A.7}$$

$$= \sqrt{\left(\hat{\mathbf{k}} \cdot (\mathbf{x}_i - \mathbf{x}_c)\right)^2 - ||\mathbf{x}_i - \mathbf{x}_c||_2^2 + \left(A + \frac{b}{2}\right)^2}$$

$$- \sqrt{\left(\hat{\mathbf{k}} \cdot (\mathbf{x}_i - \mathbf{x}_c)\right)^2 - ||\mathbf{x}_i - \mathbf{x}_c||_2^2 + \left(A - \frac{b}{2}\right)^2}. \tag{A.8}$$

Now we introduce time-dependence. While stations track a source of fixed $\hat{\mathbf{k}}$ (which we could identify with a tuple containing a right ascension and a declination), Δs_i will change (for each station indexed by i). This is because in reference frames which are *not* moving and spinning along with Terra, both the planet's core and the station positions are time-dependent: $\mathbf{x}_c = \mathbf{x}_c(t)$ and $\mathbf{x}_i = \mathbf{x}_i(t)$. Especially for sources in Sol's planetary system and in Via Lactea, the motion of Terra around Sol induces a parallax effect that causes no source to truly retain a fixed $\hat{\mathbf{k}}$; however, here we neglect this fact — most radio sources are extragalactic anyways. In the same spirit, we also

neglect the effect of *aberration of light*.

We end up with

$$\Delta s_i(\hat{\mathbf{k}}, t) = \sqrt{\left(\hat{\mathbf{k}} \cdot (\mathbf{x}_i(t) - \mathbf{x}_c(t))\right)^2 - ||\mathbf{x}_i - \mathbf{x}_c||_2^2 + \left(A + \frac{b}{2}\right)^2} - \sqrt{\left(\hat{\mathbf{k}} \cdot (\mathbf{x}_i(t) - \mathbf{x}_c(t))\right)^2 - ||\mathbf{x}_i - \mathbf{x}_c||_2^2 + \left(A - \frac{b}{2}\right)^2}. \tag{A.9}$$

Note that within each square root, the second and third term are time-independent, as they are left unshaken by Terra's rotation around her axis.

In a shell-centred reference frame, Eq. A.9 simplifies. In such frames, we have $\mathbf{x}_c(t) = 0$, so that, upon invoking Eq. A.1, we find

$$\Delta s_i(\hat{\mathbf{k}}, t) = \sqrt{\left(\hat{\mathbf{k}} \cdot \mathbf{x}_i(t)\right)^2 - ||\mathbf{x}_i||_2^2 + \left(||\mathbf{x}_0||_2 + a + \frac{b}{2}\right)^2} - \sqrt{\left(\hat{\mathbf{k}} \cdot \mathbf{x}_i(t)\right)^2 - ||\mathbf{x}_i||_2^2 + \left(||\mathbf{x}_0||_2 + a - \frac{b}{2}\right)^2}. \tag{A.10}$$

If Terra has a spherical surface cocentric with the ionospheric shell, and all stations lie at a distance R away from the centre, then Eq. A.10 further simplifies into

$$\Delta s_i(\hat{\mathbf{k}}, t) = \sqrt{\left(R \cos \varphi_i(\hat{\mathbf{k}}, t)\right)^2 - R^2 + \left(R + a + \frac{b}{2}\right)^2} - \sqrt{\left(R \cos \varphi_i(\hat{\mathbf{k}}, t)\right)^2 - R^2 + \left(R + a - \frac{b}{2}\right)^2}. \tag{A.11}$$

where we acknowledge $\varphi_i = \varphi_i(\hat{\mathbf{k}}, t)$ is the local zenith angle for direction $\hat{\mathbf{k}}$ at $\mathbf{x}_i(t)$. By taking out R, we make more explicit that Δs_i has dimensions of *length*, and find

$$\Delta s_i(\hat{\mathbf{k}}, t) = R\left(\sqrt{\cos^2 \varphi_i(\hat{\mathbf{k}}, t) - 1 + \left(1 + \frac{a}{R} + \frac{1}{2}\frac{b}{R}\right)^2} - \sqrt{\cos^2 \varphi_i(\hat{\mathbf{k}}, t) - 1 + \left(1 + \frac{a}{R} - \frac{1}{2}\frac{b}{R}\right)^2}\right). \tag{A.12}$$

If this expression is correct, then for fixed a and b we should regain Chapter 2's flat

geometry when $R \to \infty$. To test this, consider the Taylor polynomial P of degree 1 of $\Delta s_i(\hat{\mathbf{k}}, t)$ around $\xi := \frac{b}{R} = 0$:

$$P_i\left(\hat{\mathbf{k}}, t, \xi\right) := \Delta s_i(\hat{\mathbf{k}}, t, \xi = 0) + \frac{\partial \Delta s_i\left(\hat{\mathbf{k}}, t, \xi\right)}{\partial \xi}|_{\xi=0} \cdot \xi. \tag{A.13}$$

Now $\Delta s_i(\hat{\mathbf{k}}, t, \xi = 0) = 0$, and

$$\begin{aligned}\frac{\partial \Delta s_i(\hat{\mathbf{k}}, t, \xi)}{\partial \xi} = \frac{R}{2}\Big(&\frac{1 + \frac{a}{R} + \frac{1}{2}\xi}{\sqrt{\cos^2\varphi_i - 1 + \left(1 + \frac{a}{R} + \frac{1}{2}\xi\right)^2}} + \\ &\frac{1 + \frac{a}{R} - \frac{1}{2}\xi}{\sqrt{\cos^2\varphi_i - 1 + \left(1 + \frac{a}{R} - \frac{1}{2}\xi\right)^2}}\Big).\end{aligned} \tag{A.14}$$

Thus,

$$\begin{aligned}P_i\left(\hat{\mathbf{k}}, t, \xi\right) &= \frac{\partial \Delta s_i\left(\hat{\mathbf{k}}, t, \xi\right)}{\partial \xi}|_{\xi=0} \cdot \xi \\ &= \frac{R\left(1 + \frac{a}{R}\right)}{\sqrt{\cos^2\varphi_i - 1 + \left(1 + \frac{a}{R}\right)^2}} \cdot \frac{b}{R} \\ &= \frac{b\left(1 + \frac{a}{R}\right)}{\sqrt{\cos^2\varphi_i - 1 + \left(1 + \frac{a}{R}\right)^2}}.\end{aligned} \tag{A.15}$$

Using now also that $a \ll R$, we recover

$$P_i\left(\hat{\mathbf{k}}, t\right) \to \frac{b}{\sqrt{\cos^2\varphi_i}} = \frac{b}{|\cos\varphi_i|} = b\sec\varphi_i \tag{A.16}$$

for $0 \leq \varphi_i \leq 90°$.

What are typical values of $\frac{a}{R}$ and $\frac{b}{R}$ for Terra? Because $R \approx 6400$ km, and $a \approx 320$ km, we have $\frac{a}{R} \approx \frac{1}{20}$. Also $\frac{b}{R}$ is on the same scale, because $b \approx 320$ km is possible. So in practical scenarios, the regimes $\frac{a}{R} \ll 1$ and $\frac{b}{R} \ll 1$ do not fully hold.

The DTEC expectation value for stations with indices i and j under an ionosphere

with a constant FED mean μ, is

$$\mathbb{E}\left[\tau_{ij}\left(\hat{\mathbf{k}}, t\right)\right] = \mu\left(\Delta s_i(\hat{\mathbf{k}}, t) - \Delta s_j(\hat{\mathbf{k}}, t)\right). \tag{A.17}$$

In the special case that Terra's radius is very large compared to the ionosphere's height above the surface and thickness, we recover

$$\mathbb{E}\left[\tau_{ij}\left(\hat{\mathbf{k}}, t\right)\right] = \mu b\left(\sec\varphi_i(\hat{\mathbf{k}}, t) - \sec\varphi_j(\hat{\mathbf{k}}, t)\right), \tag{A.18}$$

as in Chapter 2.

B

Flux scale–induced spectral index uncertainties for high-SNR sources

Abstract

Context A standard problem in radio astronomy is the estimation of the spectral index α of a source with a power-law spectrum, based on images at two different frequencies. For sources imaged at high signal-to-noise ratio (SNR), the dominant flux density uncertainty — or specific intensity uncertainty — is typically the image's flux *scale* uncertainty. The LOFAR Two-metre Sky Survey (LoTSS) DR2, for example, suffers from a $\sim$10% flux scale uncertainty.

Aims We analytically and numerically characterise the effect of flux scale uncertainties on spectral index measurements $\hat{\alpha}$ for high-SNR sources.

Methods We derive an expression for the spectral index error $\varepsilon := \hat{\alpha} - \alpha$, whose distribution is, under a plausible assumption, exactly Gaussian. We derive expressions for its mean μ_ε and standard deviation σ_ε. The mean is non-zero: flux scale errors make the standard spectral index formula a *biased* estimator of the true spectral index. The magnitude of the relative bias $\frac{\mu_\varepsilon}{\sigma_\varepsilon}$ is independent of the images' frequencies.

Results For a range of realistic flux scale uncertainties, we numerically calculate the spectral index bias μ_ε, uncertainty σ_ε, and relative bias $\frac{\mu_\varepsilon}{\sigma_\varepsilon}$. The bias is insignificant compared to the uncertainty. The uncertainty itself, however, *is* significant.

Conclusions For high-SNR sources, flux scale uncertainties are the main obstacle to accurately measuring spectral indices. Conveniently, under reasonable conditions, the distribution of the measured spectral index $\hat{\alpha}$ approaches perfect Gaussianity as the source SNR increases. Generally speaking, $\hat{\alpha}$ is a biased estimator of α — remarkably even when the flux scale corrections applied to the imagery are on average correct. However, the bias appears small under realistic conditions: $\mu_\varepsilon \sim 10^{-3}$–$10^{-2}$ for LoLSS–LoTSS spectral index measurements. We offer a simple formula to remove this bias, actionable once flux scale uncertainties are quantified.

B.1 Introduction

Modern radio astronomical surveys, such as the LOFAR (van Haarlem et al., 2013) Two-metre Sky Survey (LoTSS) DR2 (Shimwell et al., 2022), go sufficiently deep and cover large enough areas of sky to detect millions of sources with high significance (i.e. $> 5\sigma$, where σ is the image background noise). Flux density uncertainties of such *high-σ* or *high-SNR* sources, as we will call them here, are no longer dominated by additive background noise of thermal nature. Instead, *multiplicative* noise becomes dominant. As an example, for the LoTSS DR2, flux scale uncertainty limits the accuracy of flux densities and specific intensities to 10% — regardless of a source's brightness.

Similar surveys at other frequencies are planned or currently ongoing, which will enable spectral index measurements of millions of high-σ sources in the foreseeable future. Such measurements are scientifically interesting for a myriad of reasons. One such reason is the search for ultra-steep-spectrum (USS) radio galaxies (RGs), which have been found to reside preferentially at high redshifts. Such RGs could help to probe the physics of the Epoch of Reionisation (e.g. Miley & De Breuck, 2008; Saxena et al., 2018). Another reason to amass spectral index measurements is to boost the search for MHz-peaked spectrum (MPS) and GHz-peaked spectrum (GPS) sources, which likely represent youthful RGs, RGs in dense environments, or transients. As such, they serve as important observational probes to test models of RG evolution (e.g. O'Dea et al., 1991; O'Dea & Saikia, 2021).

In this work, we analytically and numerically characterise the spectral index error introduced by multiplicative noise sources in the high-SNR limit.

B.2 Analytical results

The simplest synchrotron-emitting sources have a power-law spectrum: $F_\nu \propto \nu^\alpha$. Let us now consider such a source. In particular, let $F_\nu(\nu_1)$ and $F_\nu(\nu_2)$ denote the ground-

truth flux densities of the source at observing frequencies ν_1 and ν_2 respectively, and let α be the ground-truth spectral index. Then

$$\frac{F_\nu(\nu_1)}{F_\nu(\nu_2)} = \left(\frac{\nu_1}{\nu_2}\right)^\alpha, \text{ or } \alpha = \frac{\ln \frac{F_\nu(\nu_1)}{F_\nu(\nu_2)}}{\ln \frac{\nu_1}{\nu_2}} \text{ provided that } \nu_1 \neq \nu_2. \tag{B.1}$$

We usually approximate α through the estimator $\hat{\alpha}$, computed with the above formula, but with the ground-truth flux densities replaced by *observed* flux densities $\hat{F}_\nu(\nu_1)$ and $\hat{F}_\nu(\nu_2)$. For high-σ sources, the observed flux densities differ from the ground-truth flux densities by a *multiplicative* factor, because *additive* image noise is negligible (by definition): $\hat{F}_\nu(\nu_1) \approx F_\nu(\nu_1) \cdot C_1$ and $\hat{F}_\nu(\nu_2) \approx F_\nu(\nu_1) \cdot C_2$. Then

$$\hat{\alpha} := \frac{\ln \frac{\hat{F}_\nu(\nu_1)}{\hat{F}_\nu(\nu_2)}}{\ln \frac{\nu_1}{\nu_2}} \approx \frac{\ln \frac{F_\nu(\nu_1)\cdot C_1}{F_\nu(\nu_2)\cdot C_2}}{\ln \frac{\nu_1}{\nu_2}} = \frac{\ln \frac{F_\nu(\nu_1)}{F_\nu(\nu_2)} + \ln \frac{C_1}{C_2}}{\ln \frac{\nu_1}{\nu_2}} =: \alpha + \varepsilon. \tag{B.2}$$

We read off that the spectral index error is

$$\varepsilon = \frac{\ln C_1 - \ln C_2}{\ln \frac{\nu_1}{\nu_2}}. \tag{B.3}$$

What is the distribution of ε?

The multiplicative factors C_1 and C_2 might deviate from 1 because of a myriad of effects, such as errors in beam and ionospheric calibration. Just as the central limit theorem (CLT) predicts that the normalised sum (i.e. the average) of many independent random variables (RVs) follows a normal distribution, the *multiplicative* CLT predicts that the normalised *product* of many independent RVs follows a *lognormal* distribution. This justifies the assumption that $C_i \sim \text{Lognormal}(\mu_i, \sigma_i^2)$ for $i \in \{1, 2\}$. This in turn implies $\ln C_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$, so that — elegantly — Eq. B.3's spectral index error has a normal distribution! The parameters μ_i and σ_i^2 are the mean and variance of the corresponding normal distribution: $\mathbb{E}[\ln C_i] = \mu_i$ and $\mathbb{V}[\ln C_i] = \sigma_i^2$. They relate to the mean and variance of C_i through

$$\mu_i = \ln \frac{\mu_{C_i}^2}{\sqrt{\mu_{C_i}^2 + \sigma_{C_i}^2}}, \quad \sigma_i^2 = \ln\left(1 + \frac{\sigma_{C_i}^2}{\mu_{C_i}^2}\right). \tag{B.4}$$

The most optimistic case occurs when $\mu_{C_i} = 1$, so that $\mathbb{E}\left[\hat{F}_\nu(\nu_i)\right] \approx F_\nu(\nu_i)\,\mathbb{E}[C_i] = F_\nu(\nu_i)\,\mu_{C_i} = F_\nu(\nu_i)$: the observed flux density is an unbiased estimator of the ground-

truth flux density. Although C_i might vary per pointing, teams will generally try to ensure $\mu_{C_i} = 1$, and we will consider this case from here onwards:

$$\mu_i \left(\mu_{C_i} = 1\right) = -\frac{1}{2} \ln \left(1 + \sigma_{C_i}^2\right), \;\; \sigma_i^2 \left(\mu_{C_i} = 1\right) = \ln \left(1 + \sigma_{C_i}^2\right), \tag{B.5}$$

so that $\mu_i = -\frac{1}{2}\sigma_i^2$.

The mean of the spectral index error, which we will call the *(spectral index) bias*, is

$$\mathbb{E}\left[\varepsilon\right] = \frac{\mathbb{E}\left[\ln C_1\right] - \mathbb{E}\left[\ln C_2\right]}{\ln \frac{\nu_1}{\nu_2}} = \frac{\ln \left(1 + \sigma_{C_1}^2\right) - \ln \left(1 + \sigma_{C_2}^2\right)}{2 \ln \frac{\nu_2}{\nu_1}}, \tag{B.6}$$

while the variance of the spectral index error, assuming that C_1 and C_2 are independent, is

$$\mathbb{V}\left[\varepsilon\right] = \frac{\mathbb{V}\left[\ln C_1\right] + \mathbb{V}\left[\ln C_2\right]}{\ln^2 \frac{\nu_1}{\nu_2}} = \frac{\ln \left(1 + \sigma_{C_1}^2\right) + \ln \left(1 + \sigma_{C_2}^2\right)}{\ln^2 \frac{\nu_2}{\nu_1}}. \tag{B.7}$$

Because ε is Gaussian, $\mathbb{E}\left[\varepsilon\right]$ and $\mathbb{V}\left[\varepsilon\right]$ fully characterise its distribution.

We recall that $\sqrt{x^2} = |x| = x \cdot \operatorname{sgn}(x)$ for all $x \in \mathbb{R}$ and $\operatorname{sgn}(x) = \operatorname{sgn}^{-1}(x)$ for all $x \in \mathbb{R}_{\neq 0}$, where sgn is the signum function. The standard deviation (SD) of the spectral index error, or more succinctly the *spectral index uncertainty*,[1] is

$$\sigma_\varepsilon := \sqrt{\mathbb{V}\left[\varepsilon\right]} = \frac{\sqrt{\ln \left(1 + \sigma_{C_1}^2\right) + \ln \left(1 + \sigma_{C_2}^2\right)}}{\left|\ln \frac{\nu_2}{\nu_1}\right|}. \tag{B.8}$$

By combining Eqs. B.6 and B.8, we find the *relative bias* to be

$$\frac{\mu_\varepsilon}{\sigma_\varepsilon} = \frac{\ln \left(1 + \sigma_{C_1}^2\right) - \ln \left(1 + \sigma_{C_2}^2\right)}{2\sqrt{\ln \left(1 + \sigma_{C_1}^2\right) + \ln \left(1 + \sigma_{C_2}^2\right)}} \cdot \operatorname{sgn}\left(\nu_2 - \nu_1\right). \tag{B.9}$$

We remark that the absolute value of the relative bias, $\left|\frac{\mu_\varepsilon}{\sigma_\varepsilon}\right|$, is independent of the measurement frequencies.

[1] Because $\varepsilon := \hat{\alpha} - \alpha$ and α is a constant, $\mathbb{V}[\varepsilon] = \mathbb{V}[\hat{\alpha}]$. Consequently, $\sigma_\varepsilon := \sqrt{\mathbb{V}[\varepsilon]} = \sqrt{\mathbb{V}[\hat{\alpha}]} =: \sigma_{\hat{\alpha}}$. As it is apt to call $\sigma_{\hat{\alpha}}$ the *spectral index uncertainty*, this name also applies to σ_ε.

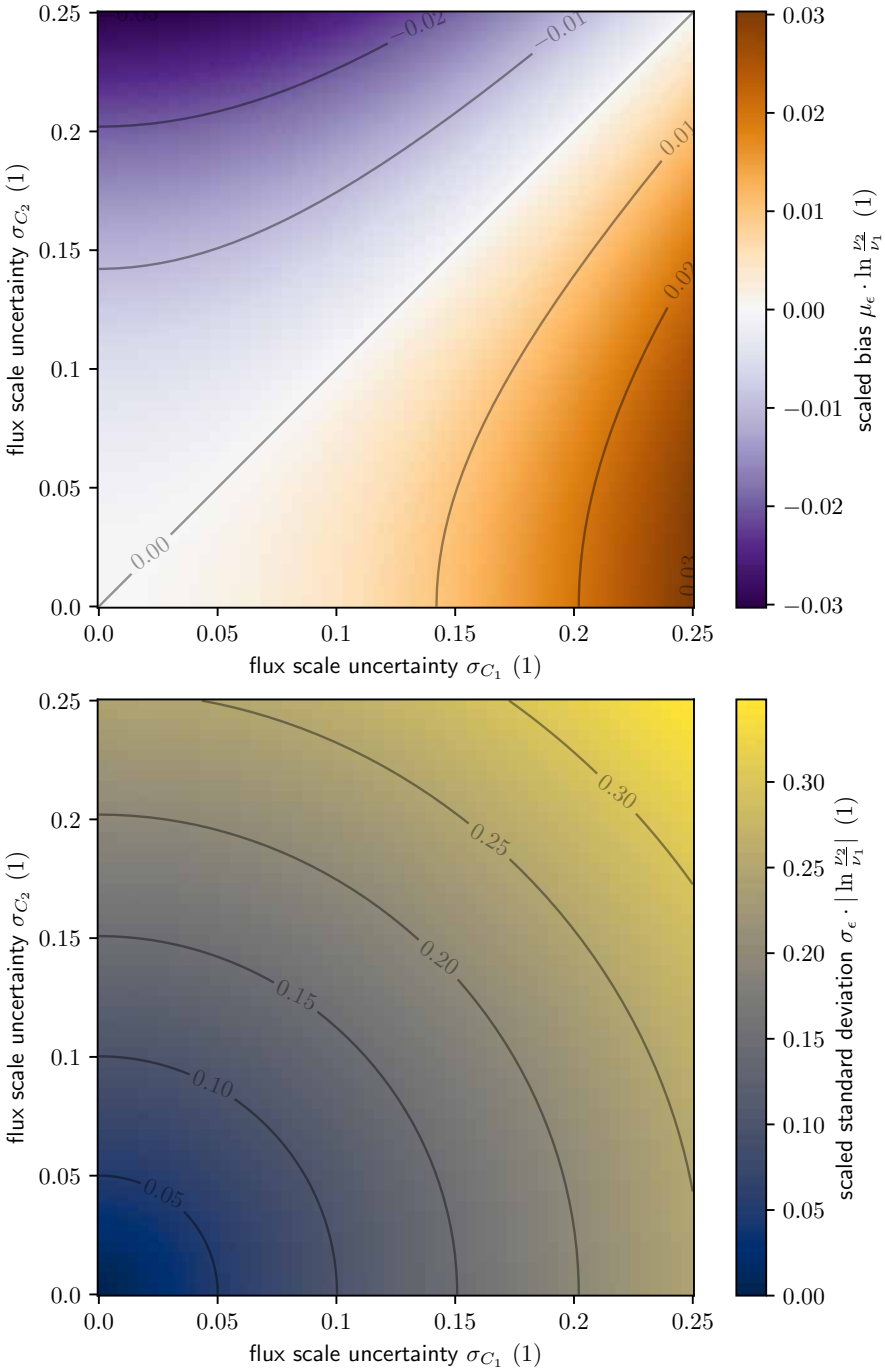


Figure B.1: Mean (top panel) and standard deviation (bottom panel) of the spectral index error ε for high-SNR sources, as a function of the flux scale uncertainties σ_{C_1} and σ_{C_2}. We scale both mean and standard deviation by a frequency-dependent factor. (This factor drops out if ν_2 is a factor e larger than ν_1. This condition is approximately satisfied for LoLSS–LoTSS spectral index measurements, where $\nu_1 = 54$ MHz and $\nu_2 = 144$ MHz.)

B.3 Numerical results

Having derived the statistical properties of the spectral index error ε that arises in spectral index measurements of bright (i.e. high-SNR) sources, we proceed to evaluating μ_ε, σ_ε, and $\frac{\mu_\varepsilon}{\sigma_\varepsilon}$ numerically.

B.3.1 Spectral index bias

From Eq. B.6, we see that the inferred spectral index $\hat{\alpha}$ is biased *positive* if the flux scale uncertainty for the lower-frequency image is *larger* than for the higher-frequency image. Similarly, $\hat{\alpha}$ is biased *negative* if the flux scale uncertainty for the lower-frequency image is *smaller* than for the higher-frequency image. We explicitly calculate the spectral index bias for the top panel of Fig. B.1. More precisely, we show $\mu_\varepsilon \cdot \ln \frac{\nu_2}{\nu_1}$, with the intent of reporting results that are independent of the particular pair of frequencies (ν_1, ν_2) used. We find that the bias is small: it is $\sim 10^{-2}$ if the σ_{C_i} differ by $\sim 10^{-1}$, and it is $\sim 10^{-3}$ if the σ_{C_i} differ by $\sim 10^{-2}$.

B.3.2 Spectral index uncertainty

From Eq. B.8, we see that the spectral index uncertainty σ_ε is invariant under exchange of σ_{C_1} and σ_{C_2}: the values of both flux scale uncertainties matter, but not to which image they belong. We explicitly calculate the spectral index uncertainty for the bottom panel of Fig. B.1. We multiply the results by the absolute value of the same frequency-dependent factor as in Sect. B.3.1. For LoTSS-like flux scale uncertainties, spectral index uncertainties for high-SNR sources are significant: $\sigma_\varepsilon \sim 10^{-1}$ for $\sigma_{C_i} \sim 10^{-1}$. Such uncertainties can interfere with drawing scientific conclusions from the estimated spectral indices.

B.3.3 Spectral index relative bias

We explicitly calculate the relative bias in Fig. B.2. For typical conditions, the bias appears to be a sub-dominant problem, being just a few percent of the spectral index uncertainty.

B.4 Conclusion

As radio surveys become more sensitive and ionospheric calibration methods improve, more source flux densities and specific intensities will be measured at high significance with respect to the image noise. For this growing population of high-SNR

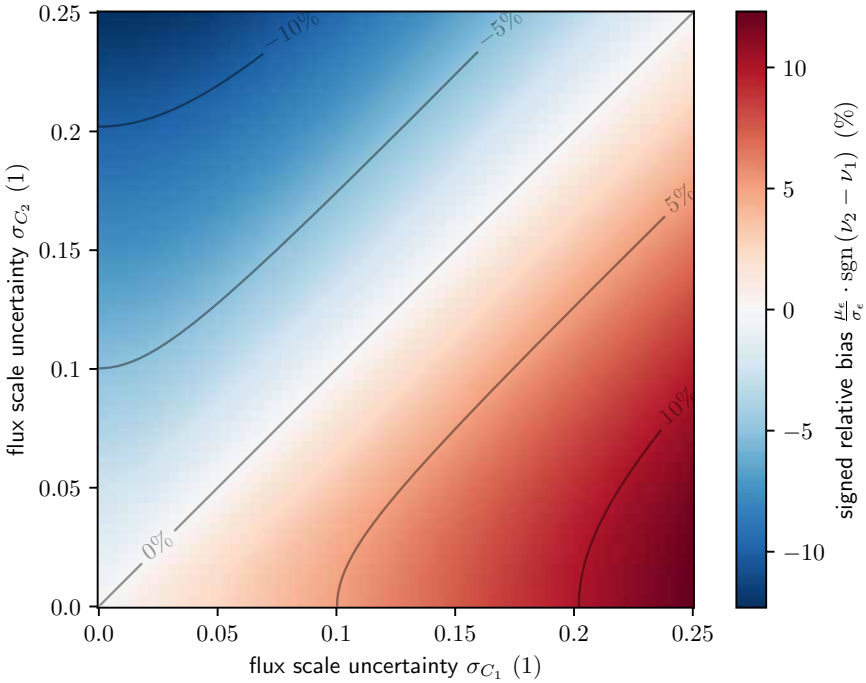


Figure B.2: Spectral index bias as a fraction of the spectral index uncertainty for high-SNR sources, as a function of the flux scale uncertainties σ_{C_1} and σ_{C_2}. The bigger the difference in flux scale uncertainty, the larger the relative bias.

sources, flux *scale* uncertainties typically dominate the flux density and specific intensity uncertainty budget. In this short work, we have considered — for such high-SNR sources — the statistical distribution of the measured spectral index $\hat{\alpha}$ and its error $\varepsilon := \hat{\alpha} - \alpha$, where α is the ground truth spectral index. We assume the standard definition for $\hat{\alpha}$: i.e. the logarithm of the ratio between the observed flux densities divided by the logarithm of the ratio between the observing frequencies. Under the assumptions that the image background noise is of negligible importance, and that the flux scale factors C_i are lognormally distributed — as suggested by the multiplicative CLT — $\hat{\alpha}$ is *exactly* Gaussian. We show that $\hat{\alpha}$ is a biased estimator of α, even when the flux scale corrections applied to the imagery are on average correct (i.e. $\mu_{C_i} = 1$). Fortunately, the spectral index bias appears small under realistic conditions. More precisely, if $\sigma_{C_2} - \sigma_{C_1} \sim 10^{-2}$–$10^{-1}$, as expected for LoLSS–LoTSS spectral index measurements, the bias $\mu_\varepsilon \sim 10^{-3}$–$10^{-2}$. More problematically, the spectral index uncertainty does not appear small: $\sigma_\varepsilon \sim 10^{-1}$. Thus, flux scale uncertainties $\sigma_{C_i} \sim 10^{-1}$ can impede the physical interpretation even of sources that are detected at very high significance with respect to the image noise (e.g. with *infinite* SNR). When the σ_{C_i} are known, Eq. B.6 offers a simple formula to remove spectral index bias.

Wat is wetenschap? Wetenschap is de titanische poging van het menselijk intellect zich uit zijn kosmische isolement te verlossen door te begrijpen.

Willem F. Hermans, Nederlands schrijver, in *Nooit Meer Slapen* (1966)

9 Populairwetenschappelijke samenvatting

Dit proefschrift beschrijft een reeks onderzoeken waarin we astronomische bronnen van radiogolven voorspellen, kalibratietechnieken voor waarnemingen door radiotelescopen beschrijven, en zulke waarnemingen analyseren en interpreteren. We richten ons in het bijzonder op radiogolven afkomstig van reusachtige straalstromen en schokgolven in het Kosmische Web. Om de inhoud en portee van het proefschrift helder over het voetlicht te brengen, duiden we eerst kort enkele belangrijke begrippen.

9.1 Licht en elektromagnetische golven

Thomas Young (1773–1829) was een Britse *uomo universale* die belangrijke bijdragen heeft geleverd aan een diverse reeks vakgebieden, waaronder de egyptologie, fysiologie, klassieke mechanica en optica. Vanwege Youngs buitengewoon brede wetenschappelijke deskundigheid wordt hij regelmatig beschouwd als *'de laatste persoon die alles wist'*. In 1803 beschreef Young voor de *Royal Society* een baanbrekend interferentie-experiment, dat hij zelf later als zijn belangrijkste wetenschappelijke bijdrage zou beschouwen. Het experiment, waarin het licht van één bron langs twee verschillende paden naar een scherm reist, toonde aan dat zichtbaar licht niet domweg bestaat uit een stroom deeltjes, zoals eminente fysici als Isaac Newton en Siméon Denis Poisson vermoedden, maar irreducibele golfeigenschappen bezit. Het was een belangrijke

stap op weg naar de acceptatie van licht als golfverschijnsel, zoals onder anderen Christiaan Huygens en Augustin-Jean Fresnel voorstonden. Niet veel later — in 1818 — bevestigde Dominique François Jean Arago van de Franse *Académie royale des sciences* Fresnels golftheorie in een *experimentum crucis*, nadat Poisson, gericht op het ontkrachten van Fresnels werk, er een ogenschijnlijk 'absurde' voorspelling uit destilleerde — namelijk dat zich in het midden van de schaduw van een circulair, ondoorzichtig voorwerp niettemin een zwak lichtvlekje zou vormen: de *tache de Fresnel*. Ironisch genoeg zou Poissons voorspelling zijn eigen ongelijk bewijzen. Arago's experimentele demonstratie van het bestaan van de *tache* overtuigde het leeuwendeel van de wetenschappelijke gemeenschap van het golfkarakter van licht. Experimentele én theoretische vooruitgang in het begrip van elektriciteit en magnetisme in de daaropvolgende decennia leidden James Clerk Maxwell tot een theorie van het *elektromagnetisme,* die niet alleen de samenhang tussen reeds bekende resultaten blootlegde, maar ook het bestaan van elektromagnetische *golven* voorspelde.[1] Al een decennium voordat — in 1865 — Maxwells *A Dynamical Theory of the Electromagnetic Field* verscheen, hadden Wilhelm Eduard Weber en Rudolf Kohlrausch opgemerkt dat er een aan elektriciteit en magnetisme gerelateerde constante bestond — laten we haar voor 't gemak c noemen — die men, gezien dier dimensies, kon interpreteren als een *snelheid*. Aan de hand van een Leidse fles maten zij vervolgens dat $c = 3{,}1 \cdot 10^8\ \mathrm{m\,s^{-1}}$. Deze waarde bleek verdacht veel op de toen recent vastgestelde snelheid van het licht: $3{,}1 \cdot 10^8\ \mathrm{m\,s^{-1}}$, aldus Hippolyte Fizeau in 1848, en $3{,}0 \cdot 10^8\ \mathrm{m\,s^{-1}}$, aldus Léon Foucault in 1850. Het was echter Maxwell die het verband tussen c en de snelheid van het licht als eerste opmerkte, en concludeerde dat licht een voorbeeld van 'zijn' elektromagnetische golven moest zijn.

9.2 Radiogolven

Het voor ons zichtbare licht vormt maar een klein deel van het brede scala, of *spectrum*, aan elektromagnetische golven dat vanuit het Heelal op Aarde 'aanspoelt'. Toch is het niet vreemd dat onze ogen — en die van de meeste andere diersoorten — geoptimaliseerd zijn voor het waarnemen van het zichtbare deel van het elektromagnetisch spectrum. Het is namelijk dít 'venster' waarin onze ster, de Zon, het sterkst straalt. (Gevoeligheid voor een deel van het spectrum waarvoor het donker is op Aarde, is van weinig nut.)

De vorm van het Zonnespectrum heeft niet alleen zijn stempel gedrukt op onze anatomie, maar ook op de geschiedenis van de sterrenkunde. De Zon is een 'gewone' ster — en dus zijn ogen die gevoelig zijn voor Zonlicht, ook bijzonder geschikt om

[1] De vectoranalysevorm van de maxwellvergelijkingen, die tegenwoordig gebruikelijk is, is van de hand van Oliver Heaviside.

de straling van *andere* sterren op te vangen. De nachtelijke hemel toont het Aardse dierenoog daarom bovenal een overweldigende zee van sterren. Tot in de twintigste eeuw bestond bijgevolg het geloof dat de elektromagnetische straling uit het Heelal vrijwel volledig van sterren moest komen — en dat zichtbaarlichttelescopen dus de meest geschikte instrumenten zijn om het Heelal te bestuderen.

In de loop van de twintigste eeuw, gedreven door stormachtige ontwikkelingen in de quantummechanica en de relativiteitstheorie, voltrok zich een ommezwaai die deze antropocentrische visie van tafel veegde. Zo voorspelden Ralph Alpher, Robert Herman en George Gamow in 1948 het bestaan van een zwakke maar alomtegenwoordige gloed van microgolfstraling die vlak na de oerknal geproduceerd zou zijn.[2] Het idee van een oerknal — een 'geboorte' voor het Heelal, in een ver maar *eindig* verleden — was indertijd controversieel. Was het Heelal dan niet, naar platonisch ideaal, onveranderlijk en oneindig oud? De toevallige ontdekking van de kosmische microgolfachtergrond door Arno Allan Penzias en Robert Woodrow Wilson in 1964, die hun de Nobelprijs voor de Natuurkunde opleverde, maakte treffend duidelijk dat waarnemingen van elektromagnetische straling buiten het optische venster tot fundamentele doorbraken kunnen leiden. Andere vroege doorbraken die het belang van de radiosterrenkunde bestendigden, zijn de voorspelling van de 21 centimeterlijn van neutraal waterstof door Hendrik Christoffel van de Hulst in 1944 en de detectie van de lijn in 1951, en de eerste metingen van gepolariseerd licht afkomstig van de Krabnevel (zie Fig. 9.1) door zowel Mikheil Alexandres dze Vashakidze als Viktor Alekseyevich Dombrovskij in 1954, en door Jan Hendrik Oort en Theodore Walraven in 1955, waarmee zij de astrofysische realiteit van het synchrotronstralingsmechanisme aannemelijk maakten. Synchrotronstraling ontstaat wanneer geladen deeltjes, zoals elektronen en positronen (anti-elektronen), met bijna de lichtsnelheid door een magnetisch veld spiralen. Deeltjes bereiken zulke snelheden niet zomaar. In de buurt van synchrotronstralingsbronnen moet daarom een natuurlijke deeltjesversneller aanwezig zijn — en een energiebron waaruit de nieuwverworven bewegingsenergie afkomstig is. Spannend genoeg verraadt een detectie van synchrotronstraling in een stuk ruimte dus het bestaan van hoogenergetische, explosieve en anderszins exotische astrofysische processen. De deeltjesversnellers die in dit proefschrift besproken worden, zijn superzware zwarte gaten, en oergas dat na miljarden jaren reizen met hoge snelheid op het Kosmische Web stort.

[2] Deze gloed bestond destijds vooral uit infraroodstraling, maar veranderde in de loop der tijd — door de expansie van het Heelal — in microgolfstraling. Ongeveer 1% van de 'sneeuw' zichtbaar op de schermen van ouderwetse, analoge televisies is het gevolg van deze kosmische microgolfstraling.

Figure 9.1: De Krabnevel, het overblijfsel van een supernova die op 4 juli 1054 als nieuwe 'ster' aan de hemel door Chinese astronomen opgemerkt werd. In het midden van deze foto van de James Webb Space Telescope huist de Krabpulsar, een neutronenster die zo'n dertig maal per seconde om haar as wentelt. De Krabnevel heeft een sleutelrol gespeeld in de ontdekking van de astrofysische relevantie van het synchrotronstralingsmechanisme. *Van: CSA, ESA, NASA, STScI & Tea Temim*

9.3 Het Kosmische Web

Hoewel het huidige sterrenkundige wereldbeeld incompleet is, verschaft het een verbluffend vergezicht van de bredere werkelijkheid waar wij deel van uitmaken. In dit wereldbeeld is de Aarde een natte, tollende rots van bescheiden formaat die om een middelgrote ster cirkelt — een bal van plasma en licht die wij de Zon noemen, en die het leven op Aarde al zo'n vier miljard jaar lang mogelijk maakt. De Zon staat niet op zichzelf, maar is onderdeel van een kosmisch eiland van een paar honderd miljard sterren.[3] Dit eiland ziet er voor mogelijke verafgelegen waarnemers uit als een imposante, meerarmige draaikolk (zie het linkerpaneel van Fig. 9.2). Aan *onze* hemel verschijnt

[3] Dat zijn zo'n één tot vier sterren voor iedere mens die ooit heeft geleefd.

Figure 9.2: *Links:* Zichtbaarlichtfoto gemaakt door de Hubbleruimtetelescoop van het sterrenstelsel UGC12158, een Melkweganalogon. Als het mogelijk was om de Melkweg uit te vliegen in een richting loodrecht op het Galactisch Vlak, dan zou het uitzicht op ons sterrenstelsel waarschijnlijk gelijkaardig zijn. *Rechts:* De 2,6 kilometer hoge Chileense berg Cerro Paranal met de contouren van de *Very Large Telescope*, waarboven de Melkweg zich uitstrekt. *Van: ESA & NASA (links); Alberto Ghizzi Panizza, ESO (rechts)*

het eiland echter als een *band* van zwak wit licht (zie het rechterpaneel van Fig. 9.2), dat men om die reden in het Oudgrieks, en in ettelijke andere, veelal Europese talen, de naam *Melkweg* gegeven heeft.[4] Deze hoogst zoogdierlijke metafoor hebben we aangehouden: alle andere dergelijke eilanden, die alomtegenwoordig zijn in het moderne Heelal, worden *galaxieën* genoemd (of: sterrenstelsels) — naar het Oudgriekse *γαλαξίας*: 'melkachtig'. De galaxieën zijn niet volledig geïsoleerd, maar ingebed in een bijna onzichtbare, warme mist van waterstof — het meest eenvoudige van alle elementen, dat zich vlak na de oerknal vormde. Deze mist heet het *intergalactisch medium*, en vormt een uitgebreid netwerk van draden, of *filamenten*, dat zich door het hele Universum uitstrekt. Het netwerk als geheel, meestal met de onderliggende donkerematerietstructuur meegerekend, wordt het *Kosmische Web* genoemd — al lijkt een vergelijking met de neurale structuur van het brein (Vazza & Feletti, 2020) treffender dan een met een spinnenweb. Omdat het Heelal eindig oud is, en de aanvoer van bouwmateriaal naar een bouwplaats hoogstens met de lichtsnelheid kan verlopen, is er een maximum aan de grootte van bouwwerken die tot nu toe in het Heelal door de zwaartekracht gevormd (kunnen) zijn. Wat die maximale grootte verder inperkt, is het feit dat het Heelal uitdijt: de uitdijing poogt voortdurend de afstand tussen filamenten te vergroten, en verhindert zo dat ze samensmelten. Om deze redenen is er, tot op heden, geen structuur groter dan het Kosmische Web gevormd. Het Kosmische

[4] De Melkweg heeft veel andere, beeldschone namen. In de meest gesproken Indiase talen — het Hindi, het Bengaals en het Marathi — noemt men de Melkweg *de Ganges van de Hemel*; in het Noors en het IJslands *het Winterse Pad*; in het Welsh *het Fort van Gwydion* (een magiër uit de Keltische mythologie); in het Thais *de Weg van de Witte Olifant*; en in het Hawaïaans onder andere *de Vis die in Schaduwen Springt*.

Web is het *Einde der Grootsheid* — de indrukwekkende hiërarchie van sterrenkundige lengteschalen ten spijt.

9.4 Schokgolven in het Kosmische Web en ionosferische kalibratie

Een veelbelovende methode om de filamenten van het Kosmische Web in beeld te brengen omvat het opvangen van synchrotronstraling met radiotelescopen. Die straling is, in het geval van het Kosmische Web, afkomstig van geladen deeltjes die versneld zijn door hypersonische schokgolven in de filamenten. Die schokgolven vormen zich wanneer oeroude gaswolken uit de gigantische ruimtes tussen de filamenten met snelheden van, doorgaans, honderden kilometers per seconde op de filamenten vallen. Dit proces, waarin de filamenten 'gevoed' worden met nieuw materiaal, wordt voorspeld door computersimulaties geënt op het reconstrueren van de ontstaansgeschiedenis van het Kosmische Web. Schokgolven in het Kosmische Web treden bovendien op ongeacht de manier waarop zulke simulaties precies afgesteld worden — de voorspelling is dus 'robuust'. De synchrotronstraling van de schokgolven is echter zwak in vergelijking met de radiostraling geproduceerd door sterrenstelsels, en dus is er een aanmerkelijk risico dat ongerelateerde sterrenstelsels in het gefotografeerde stuk hemel de radiotelescoop 'verblinden'. Om de schokgolven toch te kunnen zien, moet de straling van de schokgolven en dat van de sterrenstelsels van elkaar gescheiden blijven. Het is daarom essentieel dat verstoringen door de Aardatmosfeer de radiostraling van de sterrenstelsels niet onverhoopt uitsmeren over delen van de foto waar schokgolven zouden kunnen liggen.

Het deel van de atmosfeer dat de propagatie van kosmische radiogolven het meest verstoort, is de ionosfeer. De ionosfeer bestaat 's nachts uit twee lagen en overdag uit vier lagen, die elk enkele tientallen tot honderden kilometers boven het Aardoppervlak hangen. Een klein deel van de atomen en moleculen in deze ionosferische lagen is in geïoniseerde toestand door botsingen met ultraviolette straling en röntgenstraling van de Zon. Het elektrisch veld van de inkomende radiogolven forceert de bij ionisatie vrijgekomen elektronen in oscillatie, die daardoor een gedeelte van de energie van de inkomende golf absorberen en, later, weer uitzenden. Het gevolg is onder meer een vertraging in de aankomsttijd van de radiogolven bij de telescoop.

De ionosferische verstoringen zijn het grootst voor radiogolven met grote golflengtes — en dus lage frequenties. Het wegkalibreren van ionosferische verstoringen is dan ook van groot belang voor de laagfrequente radiosterrenkunde, de tak van de astronomie die in dit proefschrift besproken wordt. Om een goede resolutie te behalen, bestaan laagfrequente radiotelescopen doorgaans uit een aantal kleinere eenheden (veelal antennes, schotels of 'stations') die geografisch verspreid zijn, maar samenwerken als één grotere telescoop. De nu gangbare methoden voor ionosferische kalibratie maken echter niet optimaal gebruik van het feit dat de ionosferische verstorin-

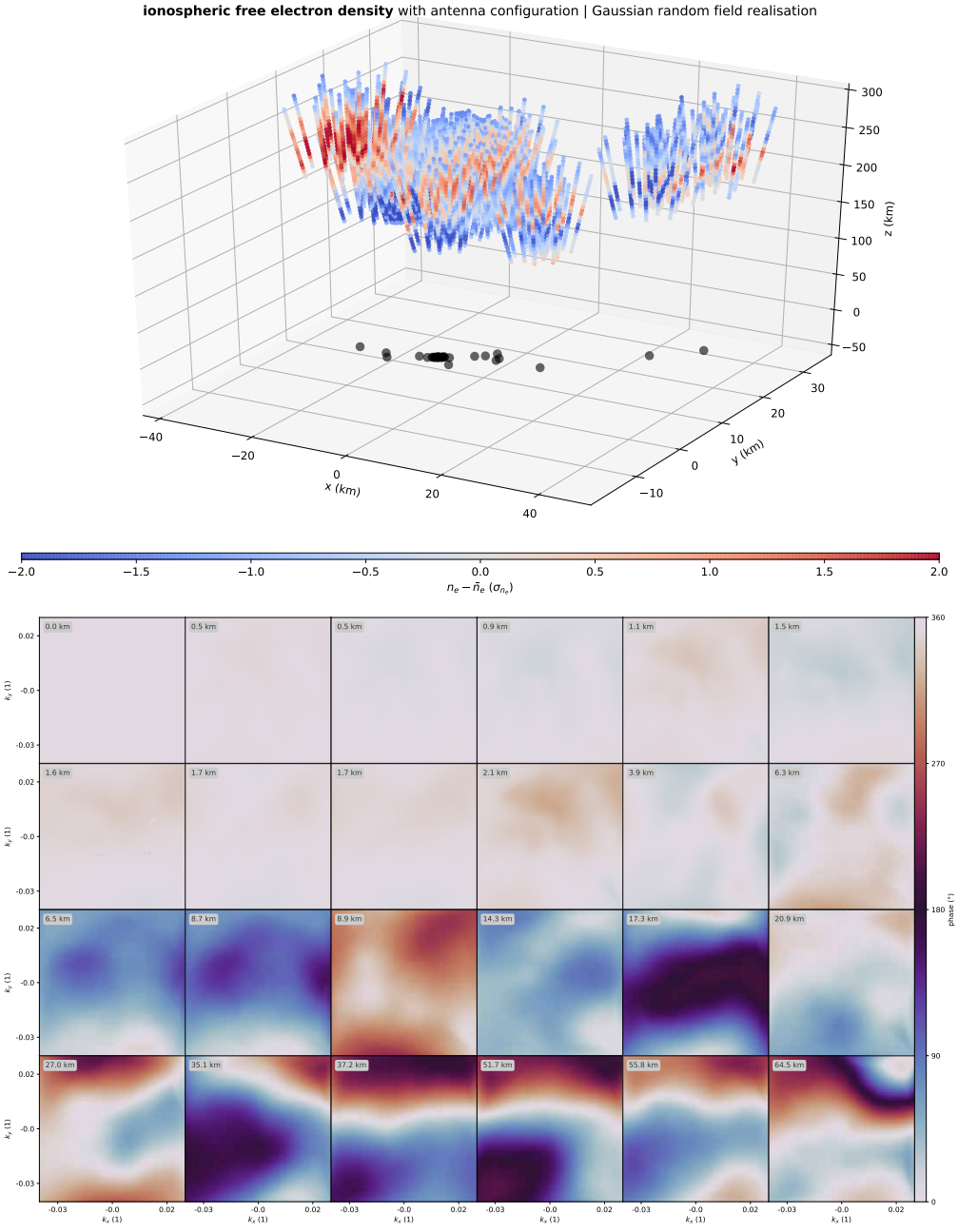


Figure 9.3: *Boven:* Ionosferische simulatie van een driedimensionaal gaussisch toevalsveld boven 24 Nederlandse LOFAR-stations *(zwarte stippen)*, waarvoor Hoofdstuk 2 de optimale kalibratiemethode presenteert. We tonen het veld $n_e(\mathbf{x}, t)$ alleen langs een paar zichtslijnen. *Onder:* Ionosferische faseverstoringen als functie van hemelrichting (k_x, k_y) voor ieder station, met de afstand tot het telescoopcentrum *(grijze legenda's)* aangegeven. *Van: Hoofdstuk 2*

gen waar verschillende stations mee te maken krijgen, flink op elkaar kunnen lijken — vooral als die stations vlak bij elkaar staan. Voor een gegeven station lijken bovendien de verstoringen van naburige richtingen aan de hemel op elkaar, alsook verstoringen van twee nauw opeenvolgende momenten in de tijd. In Hoofdstuk 2 introduceren we een nieuwe ionosferische kalibratiemethode die theoretisch optimaal is wanneer de ionosferische vrije-elektrondichtheid de structuur heeft van een gaussisch toevalsveld. Met behulp van een numerieke simulatie (zie Fig. 9.3) kwantificeren we de toename in kalibratiekwaliteit van de nieuwe methode ten opzichte van die van *ad hoc*–varianten.

9.5 Schokgolfvoorspellingen voor het Kosmische Web aan de noordelijke hemel

Een snelle inspectie van de foto's van huidige laagfrequente radiotelescopen suggereert dat de meeste schokgolven in het Kosmische Web nog niet individueel detecteerbaar zijn. Om het bestaan van de schokgolven in het Kosmische Web toch al aan te tonen, stellen wij (zoals andere auteurs) voor om een *statistisch* vergelijk te doen tussen de huidige beelden enerzijds en voorspelde schokgolfbeelden anderzijds. Voorgaand werk in deze richting heeft echter ofwel zeer simplistische voorspellingen gebruikt, ofwel voorspellingen zonder onzekerheidskwantificatie. In Hoofdstuk 3 beschrijven en presenteren we voorspellingen, gebaseerd op kosmologische magnetohydrodynamische simulaties en bayesiaanse reconstructies van de groteschaalstructuur van het Lokale Universum, mét onzekerheden. We doen voorspellingen voor de helft van de noordelijke hemel, tot op ruim twee miljard lichtjaar van de Melkweg. Zie Fig. 9.4. Het blijkt dat de grote variatie in lichtkracht die schokgolven in het Kosmische Web kunnen hebben, een grote voorspelonzekerheid induceert — zelfs wanneer we de radiostraling flink 'uitsmeren' over de hemel (over ruimtehoeken van enkele vierkante graden). De voorspellingen suggereren dat de omgevingen van het Herculescluster, het Comacluster, Abell 2199 en het Ursa Majorsupercluster mogelijk de helderste schokgolven huisvesten in het door ons geanalyseerde deel van het Kosmische Web.

Daarnaast introduceren we een eenvoudig statistisch-geometrisch model waarmee vastgesteld kan worden of de meeste schokgolfstraling in het Kosmische Web uit de ruimte tot op twee miljard lichtjaar afstand komt, of dat de totale bijdrage van verder weg gelegen schokgolven toch aanzienlijk is — en de hier gepresenteerde voorspellingen dus incompleet zijn. We laten zien dat, tot op goede benadering, slechts één parameter de gemiddelde afstand tot de schokgolfstraling bepaalt. Deze parameter kan aan de hand van kosmologische simulaties afgeschat worden.

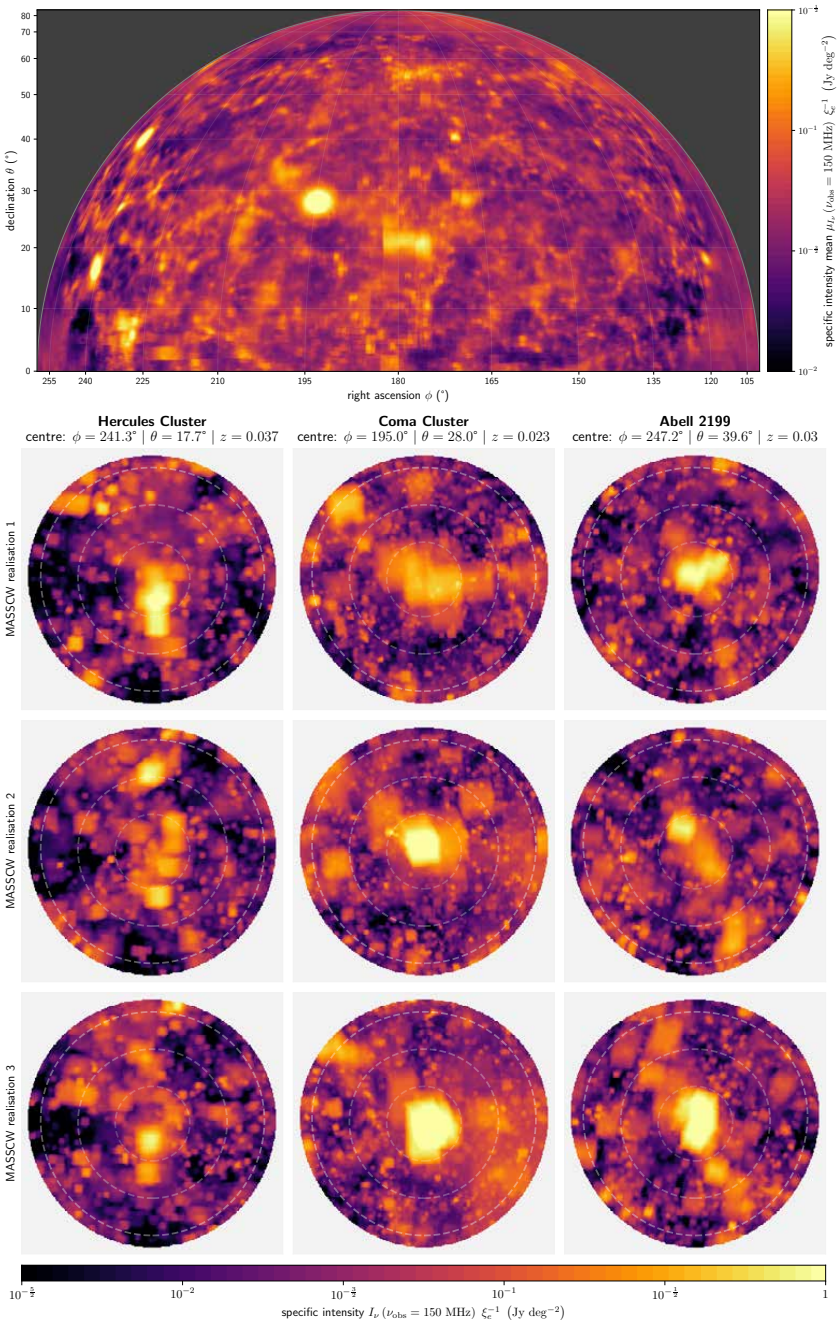


Figure 9.4: *Boven:* Gemiddelde van probabilistische Kosmische Webvoorspellingen van synchrotronstraling met een golflengte van twee meter voor de noordelijke hemel. Drie lageroodverschuivingssterrenstelsels vallen op; dit zijn (in volgorde van toenemende declinatie θ) het Herculescluster, het Comacluster en Abell 2199. *Onder:* Voor deze clusters *(kolommen)* tonen we realisaties *(rijen)* van hoe de omliggende radiohemel eruit zou kunnen zien *(stralen van gestreepte cirkels:* 3°, 6° *en* 9°*).* We modelleren slechts één stralingsmechanisme: emissie van clustersamensmeltingsschokgolven en accretieschokgolven die ontstaan tijdens de formatie van het Kosmische Web. *Van: Hoofdstuk 3*

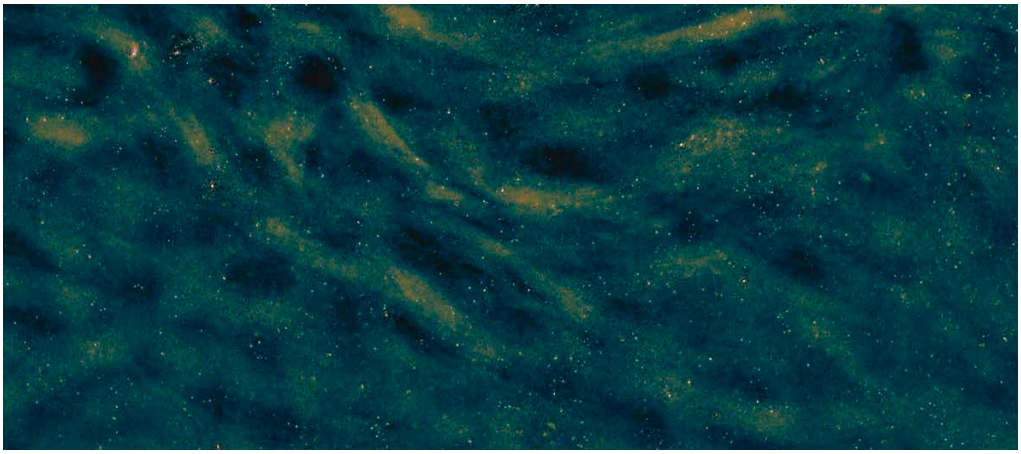


Figure 9.5: Door LoTSS-beelden te bewerken (zie Hoofdstukken 4 en 6), hebben we één van de meest, zo niet dé meest gevoelige reeks laagfrequenteradiofoto's gemaakt. Dit heeft geleid tot de toevallige ontdekking dat ons eigen sterrenstelsel de metergolfhemel met een complex patroon van gradenlange strepen vult — zelfs tot op hoge Galactische breedtegraad b. Hier tonen we een $\sim 10^2$ deg^2-groot mozaïek rond $b = 45°$. Enkele restanten van compacte bronnen zijn onverhoopt zichtbaar. *Van: Hoofdstuk 1*

9.6 Gevoelige LOFAR-beelden en een onverwachtse ontdekking

Met voorspellingen van radiostraling door schokgolven in het Kosmische Web in de hand wijden we ons aan het maken van échte radiobeelden van de noordelijke hemel. We benutten data van de LOFAR, een pan-Europese radiotelescoop met een Nederlands hart. Het tot op heden meest omvangrijke waarneemprogramma van de LOFAR, de LoTSS, is een campagne om de noordelijke hemel te fotograferen in het laagfrequente radiovenster rond 150 megahertz.[5] Het tweede grote pakket aan gepubliceerde LoTSS-data, de LoTSS DR2, omvat reeds gekalibreerde beelden van zo'n 30% van de noordelijke hemel. We bewerken deze beelden door zoveel mogelijk licht van sterrenstelsels uit de onderliggende data weg te poetsen, en vervolgens nieuwe foto's te maken met een lage resolutie. Hierdoor worden lichtzwakke patronen die zich uitstrekken over de hemel, beter zichtbaar. De bewerkte reeks beelden is de gevoeligste ooit gemaakt van de laagfrequente radiohemel met resoluties van rond de boogminuut. Spannend genoeg zou er daarom veel in te ontdekken kunnen zijn.

Het eerste dat opvalt, is een wanordelijk patroon van strepen, die vele graden lang kunnen zijn en zich tot op hoge Galactische breedtegraad manifesteren (zie Fig. 9.5). Deze strepen zijn waarschijnlijk afkomstig van de Melkweg — en niet van het Kosmische Web. Ze laten zien dat ons sterrenstelsel een tot op heden onderschatte ruisbron is in statistische zoektochten naar radiostraling van het Kosmische Web.

[5] Golven in dit venster hebben golflengtes van circa twee meter.

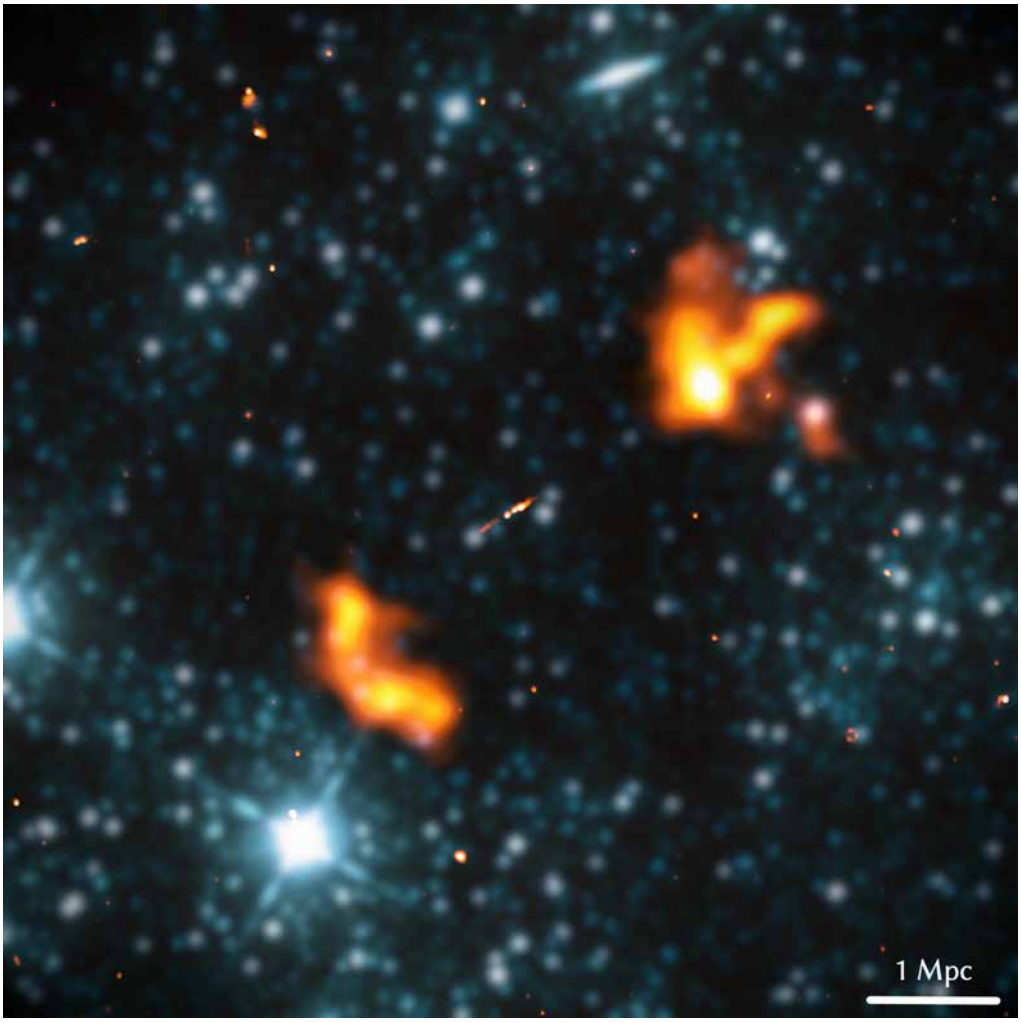


Figure 9.6: *Alcyoneus,* de grootste bekende structuur van galactische oorsprong, is gevormd door een krachtig actief zwart gat. Alcyoneus' ontdekking toont aan dat sterrenstelsels het omliggende Kosmische Web kunnen verhitten en magnetiseren tot op kosmologische afstanden (i.e. miljoenen lichtjaren). We combineren LOFAR-foto's bij golflengtes van twee meter *(oranje)* met een WISE-foto bij een golflengte van 3,4 μm *(blauw)*. *Van: Hoofdstuk 4*

Bij het bekijken van de nieuwe beeldenreeks vallen vervolgens twee pluimen op die zich aan weerszijden van een elliptisch sterrenstelsel bevinden (zie Fig. 9.6). Nadere analyse, beschreven in Hoofdstuk 4, wijst uit dat de pluimen gigantische plasmawolken representeren die de intergalactische ruimte zijn ingeblazen door een paar straalstromen met antiparallelle oriëntatie. Deze straalstromen zijn op hun beurt gegenereerd door een superzwaar zwart gat[6] in het centrum van het bijbehorende ster-

[6] In de sterrenkunde zijn superzware zwarte gaten gedefinieerd als zwarte gaten met massa's van ten

renstelsel. De minimale afstand tussen de pluimen, en daarmee ook de minimale lengte van de twee straalstromen, is zo'n 5 megaparsec (16 miljoen lichtjaar). Dit gigantische galactische uitvloeisel blijkt, ten minste in geprojecteerde zin, de grootste door een sterrenstelsel voortgebrachte structuur te zijn die tot op heden bekend is. We noemen het uitvloeisel *Alcyoneus*, naar de mythologische gigant die tijdens de gigantomachie met de Olympische goden vocht om de heerschappij over de Kosmos. We onderzoeken of Alcyoneus bijzondere eigenschappen bezit die direct verklaren waarom het uitvloeisel zo groot geworden is. Opvallend genoeg zendt Alcyoneus, ten opzichte van andere bekende uitvloeisels met straalstromen van zo'n één miljoen lichtjaar en langer, een mediale hoeveelheid radiostraling uit. Ook zijn, binnen deze referentiegroep, zowel de totale massa van de sterren in het bijbehorende sterrenstelsel als de massa van het zwarte gat dat de straalstromen geproduceerd heeft, typisch — zo niet aan de *lage* kant. Blijkbaar zijn een zeer hoge radiolichtkracht, stellaire massa of zwartgatmassa geen noodzakelijke voorwaarden om uitvloeisels zoals Alcyoneus te produceren. We schatten ten slotte de druk af in de plasmawolken, en vinden — dankzij de gevoeligheid van de LoTSS — de laagste druk ooit gemeten in een uitvloeisel. Deze druk is echter nog altijd hoger dan de typische druk in filamenten van het Kosmische Web. Aangezien Alcyoneus waarschijnlijk ook in een filament ligt, zijn de plasmawolken, en daarmee Alcyoneus als geheel, nog altijd bezig te groeien.

9.7 Een gigantisch uitvloeisel uit een spiraalsterrenstelsel

Na de spectaculaire vondst van Alcyoneus besluiten we een uitgebreidere LoTSS-zoektocht naar gigantische galactische uitvloeisels op poten te zetten. Één van de opvallendste uitvloeisels die we vinden, gepresenteerd in Hoofdstuk 5, is een uitvloeisel van ten minste 8 miljoen lichtjaar lang dat voortgebracht is door een centraal zwart gat in een spiraalsterrenstelsel. Spiraalsterrenstelsels, zoals de Melkweg, hebben meestal geen centrale zwarte gaten die gedurende lange tijd krachtige straalstromen genereren. Daardoor zijn, voor zover bekend, galactische uitvloeisels uit spiraalsterrenstelsels zeldzaam. Het uitvloeisel dat we hier presenteren, komt uit NGC 6185, en bevindt zich een half miljard lichtjaar van ons vandaan — dat wil zeggen, op een kosmische steenworpafstand. Het is het grootste bekende uitvloeisel uit een Melkwegachtig sterrenstelsel. De radiostraling afkomstig van de plasmawolken is nog maar nét in de LoTSS-foto's zichtbaar. Vergelijkbare uitvloeisels op grotere afstand zijn met de huidige data daarom niet, of bijna niet, te zien. De ontdekking van het uitvloeisel uit het nabijgelegen NGC 6185 suggeert dat gigantische galactische uitvloeisels uit een spiraalsterrenstelsels vaker voorkomen dan tot op heden gedacht. Het gebrek aan bekende exemplaren zou zomaar eens een observationeel selectie-effect kunnen zijn.

minste een miljoen keer de massa van de Zon.

Zoals bij Alcyoneus meten we de druk in de plasmawolken, en vinden een vergelijkbaar lage waarde. Omdat NGC 6185 relatief dichtbij staat, zijn er bayesiaanse Kosmische Webreconstructies beschikbaar waarmee de dichtheid van het filament waar het sterrenstelsel zich in bevindt, grofweg kan worden bepaald. Ook zijn er relatief gedetailleerde observaties mogelijk van ándere sterrenstelsels in de groep van sterrenstelsels waarin NGC 6185 het zwaarste is. Met schattingen van de druk in de plasmawolken van het uitvloeisel van NGC 6185 en van de dichtheid van het omliggende Kosmische Web, berekenen we de temperatuur van het intergalactisch medium rond NGC 6185. Het idee is, simpelweg, om gebruik te maken van de ideaalgaswet. Onze meting, $T = 11^{+12}_{-5} \cdot 10^6$ K, geldt als de eerste bepaling van de temperatuur van het intergalactisch medium middels een gigantisch galactisch uitvloeisel. Het voordeel van deze methode is dat er geen kostbare röntgenstraalobservaties nodig zijn. De methode kan in de toekomst preciezer gemaakt worden.

9.8 Een census van giganten

Geënthousiasmeerd door onze vondsten speuren we, zoals Hoofdstuk 6 beschrijft, naar meer onbekende gigantische galactische uitvloeisels. Dit doen we simpelweg door met het oog zorgvuldig de LoTSS-beelden te bekijken. Al gauw merken we dat er in principe zoveel giganten te vinden zijn dat we, gezien de omvang van ons onderzoeksteam, niet alle kandidaat-giganten binnen een redelijke termijn op gigantschap zouden kunnen toetsen. Om toch de grootste uitvloeisels in de LoTSS-beelden niet te missen, besluiten we ons in de eerste plaats te richten op kandidaat-giganten die aan de hemel vijf of meer boogminuten lijken te beslaan. Na een lange lijst van kandidaten kritisch af te werken, blijken we 2.060 nieuwe gigantische galactische uitvloeisels ontdekt te hebben (zie e.g. Fig. 9.7). Vóór onze census waren er in de wetenschappelijke literatuur circa 1.300 giganten bekend.

Omdat onze zoektocht een statistisch significante hoeveelheid hemel beslaat, en de LoTSS-waarnemingen volgens een vast procedé uitgevoerd en gekalibreerd zijn, is een betekenisvolle statistische analyse mogelijk. We modelleren hoe een geparametriseerde onderliggende populatie, samen met twee eveneens geparametriseerde selectie-effecten, tot een monster leidt. De selectie-effecten die we in ogenschouw nemen, zijn de eindige gevoeligheid van de LoTSS-foto's en de minimalehoekgrootte-eis van vijf boogminuten. Door het monster met onze theoretische resultaten te combineren, leiden we af dat de intrinsieke lengtes van gigantische galactische uitvloeisels consistent zijn met een paretoverdeling. Ook geven we een eerste inschatting van de intrinsieke populatiedichtheid van giganten in het Kosmische Web.

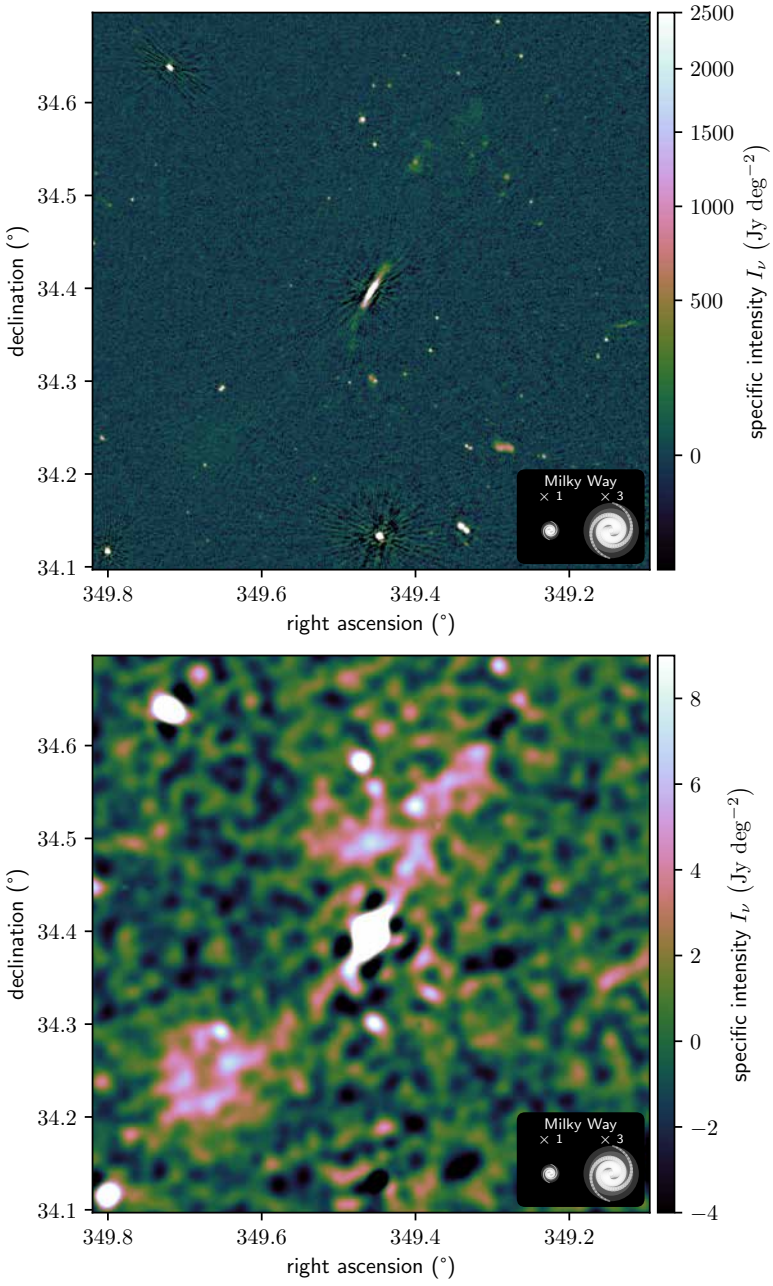


Figure 9.7: LoTSS DR2–uitsneden met resoluties van 6″ *(boven)* en 60″ *(onder)* van een pasontdekte gigant. Door compacte bronnen uit reeds gekalibreerde 150 MHz–data te verwijderen en lageresolutiebeeldvorming te bedrijven, leggen we anders speculatieve giganten bloot rond het nog onontgonnen oppervlaktehelderheidsniveau van ∼1 Jy deg^{-2}. Het geïdentificeerde bronsterrenstelsel bevindt zich in het midden van de plaatjes. Deze gigant heeft een geprojecteerde lengte $l_\mathrm{p} = 1{,}4 \pm 0{,}3$ Mpc en een hoeklengte $\phi = 32{,}3 \pm 0{,}2'$ — groter, dus, dan de volle Maan! Om een gevoel voor de omvang te krijgen, tonen we de stellaire Melkwegschijf (waarvan de diameter 50 kpc is) alsook een driemaal grotere versie. *Van: Hoofdstuk 6*

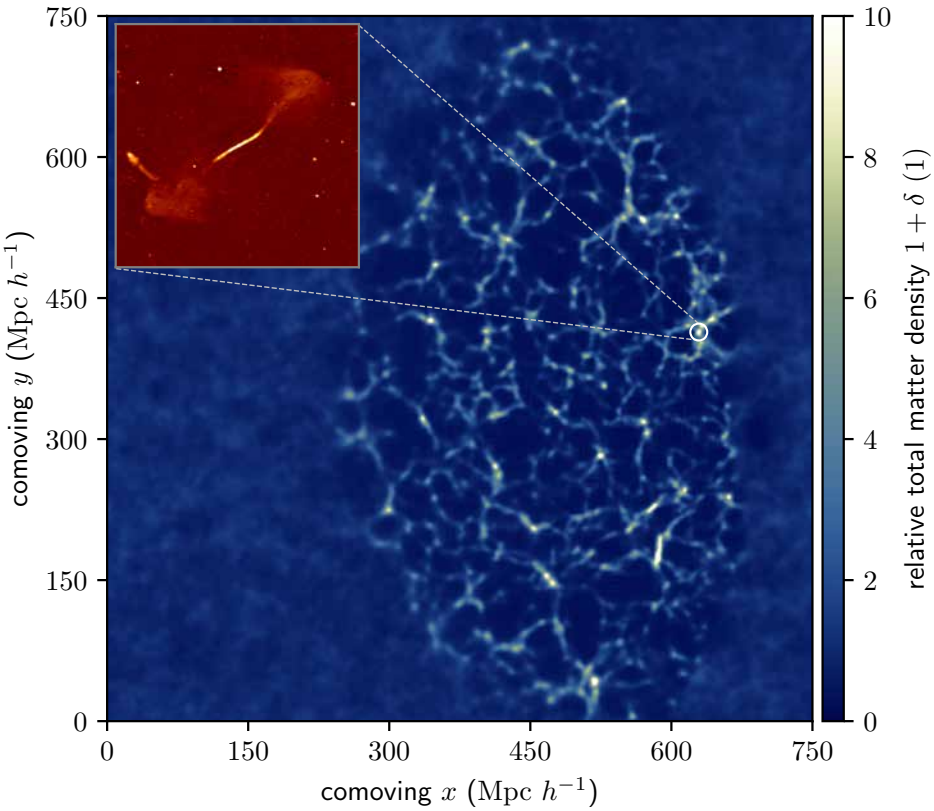


Figure 9.8: Dankzij een Kosmische Webreconstructiedoorbraak met montecarlotechnieken en herhaaldelijke simulaties van donkerematerie dynamica vanaf het vroege Heelal tot het heden, is het nu mogelijk om de dichtheid en gravitationele stabiliteit te meten van het Kosmische Web rond individuele galactische uitvloeisels (Hoofdstuk 7). Zo huist de gigant die we in Hoofdstuk 6 ontdekken en die in de vergroting weergegeven is, in een cluster van sterrenstelsels *(witte cirkel)*. Het hoofdpaneel toont een plakje van het driedimensionale a-posteriorigemiddelde van de BORG SDSS (Jasche et al., 2015). *Van: Hoofdstuk 7*

9.9 De omgevingen van gigantische galactische uitvloeisels in het Kosmische Web

In Hoofdstuk 7 demonstreren we dat moderne Kosmische Webreconstructies van het Lokale Universum ingezet kunnen worden om de dichtheid van het intergalactisch medium rondom individuele galactische uitvloeisels af te schatten. We lokaliseren in totaal 260 giganten (waarvan 80% ontdekkingen zijn uit Hoofdstuk 6) en 1.443 'algemene' galactische uitvloeisels, geselecteerd uit een groter monster zonder lengte als expliciete eis mee te nemen, in het Kosmische Web. Zie bijvoorbeeld Fig. 9.8.

We presenteren vervolgens de eerste kansverdelingen van de groteschaaldichtheid rond zowel giganten als algemene galactische uitvloeisels. Al enkele decennia bestaat

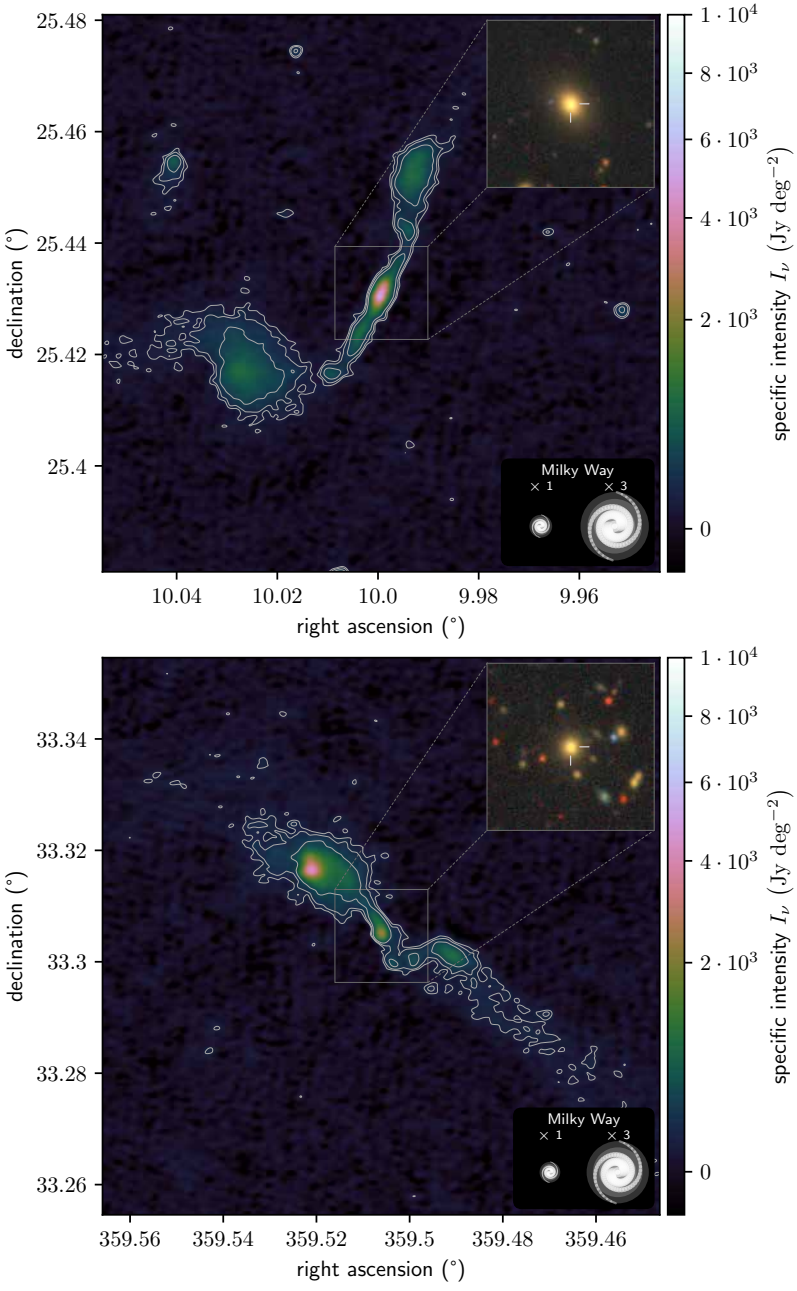


Figure 9.9: Voorbeelden van giganten die in de LoTSS DR2 gevonden zijn aan de hand van machinaal leren en geautomatiseerde, aannemelijkheidsquotiëntgebaseerde sterrenstelselassociatie. De vergrotingen tonen uitsneden van de DESI Legacy Imaging Surveys DR9 (g, r, z) (Dey et al., 2019) met afmetingen van $1' \times 1'$. *Van: Hoofdstuk 8*

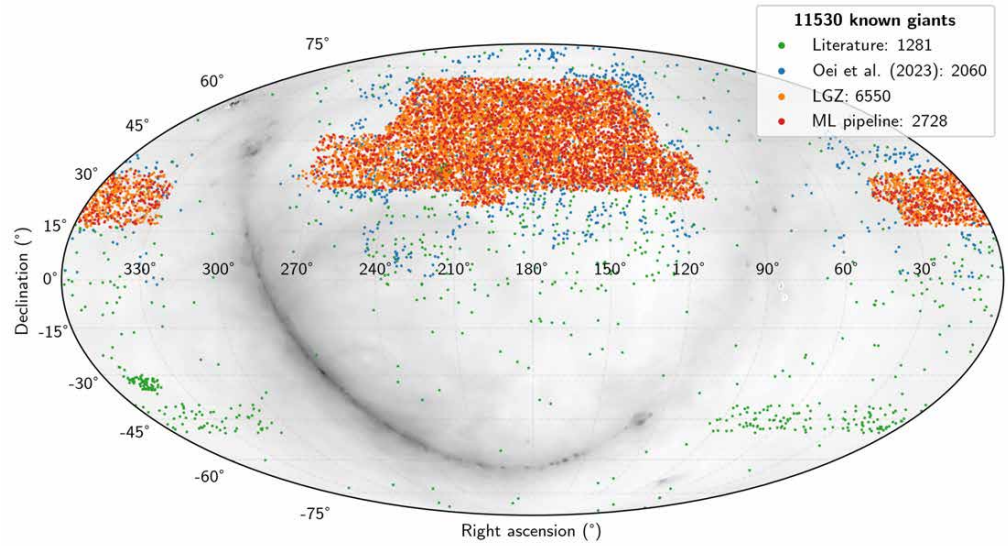


Figure 9.10: Overzicht van de hemel met de Melkweg en de coördinaten van bekende giganten ingetekend. Met in totaal 11.530 unieke bronnen presenteren we de omvangrijkste catalogus tot nu toe van grootschalige galactische terugkoppeling naar het Kosmische Web. Op het moment van schrijven zijn ∼90% van alle bekende giganten ontdekt in zoekcampagnes die we in Hoofdstukken 6 *(blauwe stippen)* en 8 *(oranje en rode stippen)* presenteren. *Van: Hoofdstuk 8*

het vermoeden dat giganten zoveel groter zijn dan andere galactische uitvloeisels doordat ze een bij uitstek ijle groteschaalomgeving zouden hebben. Onze resultaten laten echter zien dat de op dit moment bekende giganten in *dichtere* omgevingen leven dan algemene uitvloeisels. De huidig bekende populatie giganten zijn ook radiolichtkrachtiger dan algemene uitvloeisels, maar dat kan een selectie-effect zijn: omdat giganten veelal een groter stuk hemel beslaan dan andere uitvloeisels, moeten ze in totaal meer radiostraling afgeven om boven de ruis van de foto's uit te komen. We werpen de hypothese op dat de hogere dichtheden rondom giganten een reflectie zijn van een onderliggende, positieve relatie tussen Kosmische Webdichtheid en straalstroomvermogen die valide is voor álle galactische uitvloeisels. Om deze hypothese kracht bij te zetten, presenteren we de eerste kwantitatieve, observationele relatie tussen Kosmische Webdichtheid en radiolichtkracht voor galactische uitvloeisels. Deze radiolichtkracht is een indirecte maat voor het vermogen van de straalstromen die het uitvloeisel aandrijven. Onze bevindingen zijn consistent met een scenario waarin giganten gewone, in plaats van mechanistisch uitzonderlijke, leden van de uitvloeiselpopulatie zijn.

Hoofdstuk 6 beschrijft een handmatige, visuele zoektocht naar giganten in de LoTSS met een hoekgrootte van vijf boogminuten of meer. In Hoofdstuk 8 presenteren we een zoektocht naar giganten in de LoTSS met een hoekgrootte van *minder* dan vijf boogminuten, waarbij we de kracht van burgerwetenschap en machinaal leren combineren. Via het *Radio Galaxy Zoo*-project en convolutionele neurale netwerken bereiken we een census van giganten met meer dan 10^4 exemplaren. Figuur 9.9 toont twee giganten die aan de hand van de machinaallerenaanpak geïdentificeerd zijn, terwijl Fig. 9.10 onze uiteindelijke census weergeeft. Bovendien breiden we het geometrische uitvloeiselpopulatiemodel uit dat Hoofdstuk 6 heeft geïntroduceerd, en bepalen een kansverdeling over de parameters van het model met de nieuwe census en bayesiaanse statistiek. Dit resulteert in verbeterde inschattingen van de populatiedichtheid van giganten, $n_{\mathrm{GRG}} = 13 \pm 10\,(100\,\mathrm{Mpc})^{-3}$, en van de instantane vullingsgraad van de pluimen van giganten in clusters en filamenten van het Kosmische Web, $\mathcal{V}_{\mathrm{GRG-CW}} = 1{,}1 \pm 0{,}9 \cdot 10^{-5}$. Ons werk levert bewijs voor het idee dat giganten binnen de uitvloeiselpopulatie talrijk zijn. Dit suggereert dat giganten een standaarduitkomst van uitvloeiselevolutie belichamen. We redeneren dat de ingeschatte waarde van $\mathcal{V}_{\mathrm{GRG-CW}}$ voldoende hoog is voor een scenario waarin giganten significant hebben bijgedragen aan de magnetisatie van het Kosmische Web. Gigantische galactische uitvloeisels, die eerder als randverschijnsel beschouwd werden, lijken voorzichtig aan hoofdverdachten te worden in de zoektocht naar de bron van kosmische magnetogenese. ■

10
List of publications

The following list contains all first-author scientific publications, including shared† first-authorships, that I have completed during my time as a PhD student. The works are listed by publication date from high to low cosmological redshift. In the digital version, the coloured words link to the NASA ADS and the *Neurology* website.

'20 **1.** *A probabilistic approach to direction-dependent ionospheric calibration*
Martijn S. S. L. Oei†, Joshua G. Albert†, Reinout J. van Weeren, Huib T. Intema, Huub J. A. Röttgering
2020, *Astronomy & Astrophysics,* 633, 77, Published

'22 **2.** *The discovery of a radio galaxy of at least 5 Mpc*
Martijn S. S. L. Oei, Reinout J. van Weeren, Martin J. Hardcastle, Andrea Botteon, Tim W. Shimwell, Pratik Dabhade, Aivin R. D. J. G. I. B. Gast, Huub J. A. Röttgering, Marcus Brüggen, Cyril Tasse, Wendy L. Williams, Aleksandar Shulevski
2022, *Astronomy & Astrophysics,* 660, 2, Published

'22 **3.** *Filamentary Baryons and Where to Find Them: A forecast of synchrotron radiation from merger and accretion shocks in the local Cosmic Web*
Martijn S. S. L. Oei, Reinout J. van Weeren, Franco Vazza, Florent Leclercq, Akshatha Gopinath, Huub J. A. Röttgering
2022, *Astronomy & Astrophysics,* 662, 87, Published

'22 **4.** *Data-driven phenotyping of central disorders of hypersomnolence with unsupervised clustering*
Martijn S. S. L. Oei†, Jari K. Gool†, Zhongxing Zhang†, Stephanie Mathias, Yves Dauvilliers, Geert Mayer, Giuseppe Plazzi, Rafael del Rio-Villegas, Joan Santamaria Cano, Karel Šonka, Markku Partinen, Sebastiaan Overeem, Rosa Peraita-Adrados, Raphael Heinzer, Antonio Martins da Silva, Birgit Högl, Aleksandra Wierzbicka, Anna Heidbreder, Eva Feketeova, Mauro Manconi, Jitka Bušková, Francesca Canellas, Claudio L. Bassetti, Lucie Barateau, Fabio Pizza, Markus H. Schmidt, Rolf Fronczek, Ramin Khatami, Gert Jan Lammers
2022, *Neurology,* 98, 23, Published

'23 **5.** *An intergalactic medium temperature from a giant radio galaxy*
Martijn S. S. L. Oei, Reinout J. van Weeren, Martin J. Hardcastle, Franco Vazza, Tim W. Shimwell, Florent Leclercq, Marcus Brüggen, Huub J. A. Röttgering
2023, *Monthly Notices of the Royal Astronomical Society,* 518, 240, Published

'23 **6.** *Measuring the giant radio galaxy length distribution with the LoTSS*
Martijn S. S. L. Oei, Reinout J. van Weeren, Aivin R. D. J. G. I. B. Gast, Andrea Botteon, Martin J. Hardcastle, Pratik Dabhade, Tim W. Shimwell, Huub J. A. Röttgering, Alexander Drabent
2023, *Astronomy & Astrophysics,* 672, 163, Published

'23 **7.** *Do luminous giants populate special large-scale environments? Or: the radio luminosity–Cosmic Web density relation for radio galaxies*
Martijn S. S. L. Oei, Reinout J. van Weeren, Martin J. Hardcastle, Aivin R. D. J. G. I. B. Gast, Florent Leclercq, Huub J. A. Röttgering, Pratik Dabhade, Tim W. Shimwell, Andrea Botteon
2023, *Astronomy & Astrophysics,* Accepted

'23 **8.** *Constraining the giant radio galaxy population with machine learning–accelerated detection and Bayesian inference*
Martijn S. S. L. Oei†, Rafaël I. J. Mostert†, Bonny Barkus, Lara Alegre, Martin J. Hardcastle, Kenneth J. Duncan, Huub J. A. Röttgering, Reinout J. van Weeren, Maya Horton
2023, *Astronomy & Astrophysics,* Submitted

Above all, I have been a sentient being, a thinking animal, on this beautiful planet, and that in itself has been an enormous privilege and adventure.

Oliver W. Sacks, British neurologist and naturalist, in *The New York Times* essay *'My Own Life'* (2015)

11
Curriculum vitae

I was born in the Leiden University Medical Center (LUMC) on Saturday 2 October 1993 to my mother Antoinette Martine Maaike Kootte and my father Tjiauw Khing Oei, themselves both medics. My parents raised me in Oegstgeest, a town bordering Leiden and a place favoured to call home by many Leiden University academics.

In the autumn of 2005, I joined the Stedelijk Gymnasium Leiden — the high school of, among others, Rembrandt van Rijn, Herman Boerhaave, and Abraham Kuyper. I had great company, too. By introducing me to the world of computer programming and game design, my friend Jacob instilled in me the idea that, ultimately, hidden physical laws govern the world. On top of that, I had found an interest that was challenging on both a creative and a logical, mathematical level. With much joy and vigour, Jacob and I worked together on many ideas over our teenage years. Our works won national prizes. Thanks in part to popularisers of science, among them Stephen Hawking, I came to understand that physics addresses quantitatively many of the fundamental questions that so captured my imagination in my later high school years. It became clear, also, that physics and cosmology were inseparable.

In the autumn of 2011, I started the bachelor programmes Astronomy and Physics at Leiden University. With Jacob, and new friends Jos and Mel, I formed an invaluable bond through which we traversed the first academic years in Leiden. I wrote my bachelor thesis on simulating and building a prototype of a differential optical trans-

Figure 11.1: A visit to the Westerbork Synthesis Radio Telescope at Camp Westerbork, Drenthe, the Netherlands, in October 2016. This visit further fueled my interest in the field of radio astronomy.

fer function wavefront sensor. I was supervised by Professor Matthew Kenworthy of Leiden Observatory. Through this project, I know how exciting it can be to work on novel astronomical instrumentation. I obtained my bachelor degrees *cum laude*. I also completed the Honours College track Beta & Life Sciences.

In the autumn of 2014, I started the master's programme Research in Astronomy-Cosmology at Leiden University. I wrote two master theses: one on asteroseismology and one on radio astronomy. My first thesis concerned modelling asteroseismological activity on the surface of Beta Pictoris, a nearby, young star with at least two super-Jupiter planets. I worked on this project under the supervision of Professor Ignas Snellen of Leiden Observatory. I am still passionate about exoplanet research. My second thesis concerned, on the one hand, ionospheric calibration of visibility data from the Upgraded Giant Metrewave Radio Telescope (uGMRT), which by then was

still in active transformation from its GMRT origin. The goal of this project was to extend Huib Intema's SPAM pipeline (Intema et al., 2009b) for the GMRT to work for the uGMRT, preferably exploiting the latter's much wider frequency bands. Besides, the thesis concerned a statistical test of the alignment of radio galaxy jets with the filaments of the Cosmic Web in which they are often embedded. My supervisors were Huib Intema, Francesco de Gasperin, and Huub Röttgering. After completing my master's degree, I stayed at the institute over the summer of 2017 to continue working on radio astronomy. By this time, Reinout van Weeren returned to Leiden Observatory for a faculty position, and he started advising me on my research.

In the autumn of 2017, my academic *Bildung* continued overseas. I got admitted to Part III of the Mathematical Tripos at the University of Cambridge, a one-year intensive master's programme in mathematics with a history going back to the 18th century. I was part of Hughes Hall, but took up residence at Swirles Court of Girton College, which was close to the Centre for Mathematical Sciences (CMS). At the CMS, I took courses covering both theoretical physics and statistics. In particular Professor Kaisey Mandel's course *Astrostatistics*, which dealt with Gaussian processes, Markov chain Monte Carlo, and Bayesian inference (among other topics), has had a major influence on the choice of methodology used in this thesis. At the end of the year, in the summer of 2018, I decided to stay a few months longer in Cambridge to do research in biotechnology at the Wellcome Sanger Institute in Hixton, located well within the hilly Cambridgeshire countryside. At *Sanger*, under the supervision of Felicity Allen, I worked on data analysis for a series of CRISPR–Cas9 human gene editing experiments. The experiments were designed to map, and later predict, the mutations induced by the interplay of the Cas9 enzyme and cellular repair mechanisms as a function of the original target sequence.

In the autumn of 2018, after moving to Amsterdam, I started my PhD research on the intersection of radio astronomy and cosmology at Leiden Observatory. I was supervised by Assistant Professor Reinout van Weeren and Full Professor Huub Röttgering. The original research goal was to conduct a statistical experiment with LOFAR data in search of the synchrotron Cosmic Web. It were the inspiring exchanges with my friend and fellow PhD student Josh Albert and the serendipitious discovery of Alcyoneus that, in the end, made me divide my PhD time over studies of the ionosphere, the synchrotron Cosmic Web, and giant galactic outflows. In 2019, I was fortunate enough to visit the MWSKY-II conference in Pune, India, and Beijing Normal University in Beijing, China. Strikingly, when I travelled back home in December, the Chinese COVID-19 outbreak had already started. When the pandemic subsided, I held public talks for a primary school and for a club of astrophotography and astrophysics enthusiasts. In addition, my discovery of Alcyoneus (Chapter 4) unleashed a

larger than expected storm of media coverage, both nationally and internationally.

Over the years, I supervised the master theses of Akshatha Gopinath and Mel Voet, and helped teach the master's course *Radio Astronomy* in the academic years 2018–2019 and 2019–2020, and then the bachelor's course *Astronomical Observing Techniques* in the academic year 2020–2021. Meanwhile, my friend, housemate, and fellow PhD student Jari Gool and I started and completed a very enjoyable collaboration on clustering analysis applied to a large European dataset on patients with central disorders of hypersomnolence. The resulting article was published in *Neurology* and awarded a prize by the European Sleep Research Society at the Sleep Europe 2022 conference in Athens, Greece.

For the statistical analysis of giant galactic outflows presented in Chapter 6, I was selected as a finalist of the American Statistical Association's Astrostatistics Interest Group 2023 Student Paper Competition. In August 2023, I presented my work at the Joint Statistical Meetings in Toronto, Canada.

On 1 October 2023, the last day of my twenties, I started as the Prize Postdoctoral Scholar Research Associate in Observational Astronomy at the California Institute of Technology in Pasadena, United States of America. During my three-year postdoctoral fellowship at Caltech, I will continue my research on giant galactic outflows and shocks in the Cosmic Web. I very much look forward to this new life chapter in Greater Los Angeles — the *City of Stars*. ■

Figure 11.2: *Astronomer by Candlelight* by Gerard Dou (1613–1675), Rembrandt's first student. The transience of human life, symbolised by the hourglass and candle, contrasts with the astronomer's ambitious goal of understanding the largely unchanging, eternal heavens. Dou, born in Leiden, became the leader of the *fijnschilders*, a collective of Leiden painters known for their meticulous, naturalist style. This nocturnal astronomer's home is the Getty Center, J. Paul Getty Museum, Los Angeles.

I, a universe of atoms, an atom in the Universe.

Richard P. Feynman, American physicist, in *The Value of Science* (1955)

12 Acknowledgments

No major undertaking in life is ever the result of a single person's efforts, and PhDs are of course no exception. In both the recent and more distant past, I have been shaped, inspired, supported, and buoyed by many lovely people around me. Feeling at home in the world has been important, because the five years that separate the start of my time as a PhD student in September 2018 from my defence in December 2023 have been ones of great change. In the following, I want to briefly thank those directly involved in crafting this PhD thesis.

First of all, I want to thank you, Reinout, for five consistent years of inspiring and hugely helpful supervision. You care deeply about your students, and in guiding me through the scientific maturation process you, without a single exception, have put my interests first. Your advising style has always been clear, sincere, friendly, realistic, timely, and long-term minded. I also want to thank Huub for his personal support and supervision. As a leader, you are ambitious but realistic, effective, future-oriented, and refreshingly opinionated if needed. In our exchanges, I have always appreciated your wit and humour. I was fortunate to count on you in all the happy *and* the less happy moments — whether through a message, call, or video, you were always there when difficulty arose — director of Leiden Observatory or not. As the focus of the PhD shifted towards giant galactic outflows, I had the luck to start collaborating with Martin Hardcastle, one of the world's principal experts on giants. Martin, thank you for your generosity in answering the many questions that I have been ask-

ing you and for important comments and suggestions on the various articles that we have now published together. Your seemingly effortless proficiency in the reduction and interpretation of observations, analytical models, *and* numerical simulations is inspiring. Among my peers at Leiden Observatory, my friend Josh Albert has been particularly influential. Josh, you elevated my days at our joint office, by raising interesting discussions and by introducing me to demanding mathematics. You are not only very gifted, but also creative, kind, and thoughtful. I am proud of the article we wrote (Chapter 2). I am also glad for the friendship of Rafaël Mostert and Frits Sweijen, both of whom started their PhD trajectories at a similar time. Rafaël, during the last months of the PhD, I have much enjoyed our collaboration on our theses' last article (Chapter 8). This reminded me of how energising the creation of a joint scientific work can be — symbolically, linking the final experience of my PhD to my first. I would also like to thank my friend Aivin Gast for his important contribution to discovering thousands of previously unknown, angularly extended giants in the LoTSS (Chapter 6), and for providing extensive further support in writing the ensuing articles. I want to thank Jacob, Jos, and Mel, which whom I formed an invaluable quartet, dubbed the *Feyne Mannen*, during the first academic years in Leiden. Jacob, I want to thank you especially for being there whenever I need it, for your encouragement and heartfelt involvement in my PhD trajectory, and for your consistently wise advice. Jos, you do not only provide comforting companionship and much laughter, but you also constantly encourage me to keep appreciating the grandeur and mystery of the subject matter of mathematics, physics, and astronomy. Finally I adress Mel, my fellow astrophysicist. I want to explicitly thank you for your friendship, informed and interesting perspectives, and sincere interest in my PhD work. I want to thank my housemates Gerwin, Jari, Ties, Guus, Melle, and László, with whom I had the pleasure to spend various periods of my five-year residence at our beautiful penthouse *'the Titus'* in Amsterdam. You have made me feel at home in the fast-paced (or locked-down) city around us. Thank you, Jari, for keeping me company on countless nights while you worked on your PhD. You would often have interesting stories or funny remarks ready, which you would sprinkle through the peaceful quiet that marked these thoroughly cosy nights. Besides, I have also much enjoyed the scientific project we have forged together. Sharing research experiences and perspectives with eachother has had a positive impact on my wellbeing and scientific maturation. I have greatly enjoyed our days at the Titus too, Guus, for the same reason. I also consider invaluable the warm support and intellectual inspiration of Aurelie, Carli, Floris, Jelle, Jesse, Lara, Mattheus, Rogier, Sam, Tim, and Willem. Many others have supported me greatly. Finally, I thank my family for standing by me throughout the PhD process. I am especially endebted to my mum, for her unabated support.

Bibliography

1. Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., Allam, S. S., Allende Prieto, C., An, D., Anderson, K. S. J., Anderson, S. F., Annis, J., Bahcall, N. A., Bailer-Jones, C. A. L., Barentine, J. C., Bassett, B. A., Becker, A. C., Beers, T. C., Bell, E. F., Belokurov, V., Berlind, A. A., Berman, E. F., Bernardi, M., Bickerton, S. J., Bizyaev, D., Blakeslee, J. P., Blanton, M. R., Bochanski, J. J., Boroski, W. N., Brewington, H. J., Brinchmann, J., Brinkmann, J., Brunner, R. J., Budavári, T., Carey, L. N., Carliles, S., Carr, M. A., Castander, F. J., Cinabro, D., Connolly, A. J., Csabai, I., Cunha, C. E., Czarapata, P. C., Davenport, J. R. A., de Haas, E., Dilday, B., Doi, M., Eisenstein, D. J., Evans, M. L., Evans, N. W., Fan, X., Friedman, S. D., Frieman, J. A., Fukugita, M., Gänsicke, B. T., Gates, E., Gillespie, B., Gilmore, G., Gonzalez, B., Gonzalez, C. F., Grebel, E. K., Gunn, J. E., Györy, Z., Hall, P. B., Harding, P., Harris, F. H., Harvanek, M., Hawley, S. L., Hayes, J. J. E., Heckman, T. M., Hendry, J. S., Hennessy, G. S., Hindsley, R. B., Hoblitt, J., Hogan, C. J., Hogg, D. W., Holtzman, J. A., Hyde, J. B., Ichikawa, S.-i., Ichikawa, T., Im, M., Ivezić, Ž., Jester, S., Jiang, L., Johnson, J. A., Jorgensen, A. M., Jurić, M., Kent, S. M., Kessler, R., Kleinman, S. J., Knapp, G. R., Konishi, K., Kron, R. G., Krzesinski, J., Kuropatkin, N., Lampeitl, H., Lebedeva, S., Lee, M. G., Lee, Y. S., French Leger, R., Lépine, S., Li, N., Lima, M., Lin, H., Long, D. C., Loomis, C. P., Loveday, J., Lupton, R. H., Magnier, E., Malanushenko, O., Malanushenko, V., Mand elbaum, R., Margon, B., Marriner, J. P., Martínez-Delgado, D., Matsubara, T., McGehee, P. M., McKay, T. A., Meiksin, A., Morrison, H. L., Mullally, F., Munn, J. A., Murphy, T., Nash, T., Nebot, A., Neilsen, Eric H., J., Newberg, H. J., Newman, P. R., Nichol, R. C., Nicinski, T., Nieto-Santisteban, M., Nitta, A., Okamura, S., Oravetz, D. J., Ostriker, J. P., Owen, R., Padmanabhan, N., Pan, K., Park, C., Pauls, G., Peoples, John, J., Percival, W. J., Pier, J. R., Pope, A. C., Pourbaix, D., Price, P. A., Purger, N., Quinn, T., Raddick, M. J., Re Fiorentin, P., Richards, G. T., Richmond, M. W., Riess, A. G., Rix, H.-W., Rockosi, C. M., Sako, M., Schlegel, D. J., Schneider, D. P., Scholz, R.-D., Schreiber, M. R., Schwope, A. D., Seljak, U., Sesar, B., Sheldon, E., Shimasaku, K., Sibley, V. C., Simmons, A. E., Sivarani, T., Allyn Smith, J., Smith, M. C., Smolčić, V., Snedden, S. A., Stebbins, A., Steinmetz, M., Stoughton, C., Strauss, M. A., SubbaRao, M., Suto, Y., Szalay, A. S., Szapudi, I., Szkody, P., Tanaka, M., Tegmark, M., Teodoro, L. F. A., Thakar, A. R., Tremonti, C. A., Tucker, D. L., Uomoto, A., Vanden Berk, D. E., Vandenberg, J., Vidrih, S., Vogeley, M. S., Voges, W., Vogt, N. P., Wadadekar, Y., Watters, S., Weinberg, D. H., West, A. A., White, S. D. M., Wilhite, B. C., Wonders, A. C., Yanny, B., Yocum, D. R., York, D. G., Zehavi, I., Zibetti, S., & Zucker, D. B. (2009). The Seventh Data Release of the Sloan Digital Sky Survey. ApJS, 182(2), 543–558.
2. Acciari, V. A., Agudo, I., Aniello, T., Ansoldi, S., Antonelli, L. A., Arbet Engels, A., Artero, M., Asano, K., Baack, D., Babić, A., Baquero, A., Barres de Almeida, U., Barrio, J. A., Batković, I., Becerra González, J., Bednarek, W., Bernardini, E., Bernardos, M., Berti, A., Besenrieder, J., Bhattacharyya, W., Bigongiari, C., Biland, A., Blanch, O., Bökenkamp, H., Bonnoli, G., Bošnjak, Ž., Burelli, I., Busetto, G., Carosi, R., Ceribella, G., Cerruti, M., Chai, Y., Chilingarian, A., Cikota, S., Colombo, E., Contreras, J. L., Cortina, J., Covino, S., D'Amico, G., D'Elia, V., da Vela, P., Dazzi, F., de Angelis, A., de Lotto, B., Del Popolo, A., Delfino, M., Delgado, J., Delgado Mendez, C., Depaoli, D., di Pierro, F., di Venere, L., Do Souto Espiñeira, E., Dominis Prester, D., Donini, A., Dorner, D., Doro, M., Elsaesser, D., Fallah Ramazani, V., Fariña, L., Fattorini, A., Font, L., Fruck, C., Fukami, S., Fukazawa, Y., García López, R. J., Garczarczyk, M., Gasparyan, S., Gaug, M., Giglietto, N., Giordano, F., Gliwny, P., Godinović, N., Green, J. G., Green, D., Hadasch, D., Hahn, A., Hassan, T., Heckmann, L., Herrera, J., Hrupec, D., Hütten, M., Inada, T., Iotov, R., Ishio, K., Iwamura, Y., Jiménez Martínez, I., Jormanainen, J., Jouvin, L., Kerszberg, D., Kobayashi, Y., Kubo, H., Kushida, J., Lamastra, A., Lelas, D., Leone, F., Lindfors, E., Linhoff, L., Liodakis, I., Lombardi, S., Longo, F., López-Coto, R., López-Moya, M., López-Oramas, A., Loporchio, S., Lorini, A., Machado de Oliveira Fraga, B., Maggio, C., Majumdar, P., Makariev, M., Mallamaci, M., Maneva, G., Manganaro, M., Mannheim, K., Mariotti, M., Martínez, M., Mas Aguilar, A., Mazin, D., Menchiari, S., Mender, S., Mićanović, S., Miceli, D., Miener, T., Miranda, J. M., Mirzoyan, R., Molina, E., Mondal, H. A., Moralejo, A., Morcuende, D., Moreno, V., Moretti, E., Nakamori, T., Nanci, C., Nava, L., Neustroev, V., Nievas Rosillo, M., Nigro, C., Nilsson, K., Nishijima, K., Noda, K., Nozaki, S., Ohtani, Y., Oka, T., Otero-Santos, J., Paiano, S., Palatiello, M., Paneque, D., Paoletti, R., Paredes, J. M., Pavletić, L., Peñil, P., Persic, M., Pihet, M., Prada Moroni, P. G., Prandini, E., Priyadarshi, C., Puljak, I., Rhode, W., Ribó, M., Rico, J., Righi, C., Rugliancich, A., Sahakyan, N., Saito, T., Sakurai, S., Satalecka, K., Saturni, F. G., Schleicher, B., Schmidt, K., Schmuckermaier, F., Schubert, J. L., Schweizer, T., Sitarek, J., Šnidarić, I., Sobczynska, D., Spolon, A.,

Stamerra, A., Strišković, J., Strom, D., Strzys, M., Suda, Y., Surić, T., Takahashi, M., Takeishi, R., Tavecchio, F., Temnikov, P., Terzić, T., Teshima, M., Tosti, L., Truzzi, S., Tutone, A., Ubach, S., van Scherpenberg, J., Vanzo, G., Vazquez Acosta, M., Ventura, S., Verguilov, V., Viale, I., Vigorito, C. F., Vitale, V., Vovk, I., Will, M., Wunderlich, C., Yamamoto, T., Zarić, D., & MAGIC Collaboration (2023). A lower bound on intergalactic magnetic fields from time variability of 1ES 0229+200 from MAGIC and Fermi/LAT observations. A&A, 670, A145.

3. Ahn, C. P., Alexandroff, R., Allende Prieto, C., Anderson, S. F., Anderton, T., Andrews, B. H., Aubourg, É., Bailey, S., Balbinot, E., Barnes, R., Bautista, J., Beers, T. C., Beifiori, A., Berlind, A. A., Bhardwaj, V., Bizyaev, D., Blake, C. H., Blanton, M. R., Blomqvist, M., Bochanski, J. J., Bolton, A. S., Borde, A., Bovy, J., Brandt, W. N., Brinkmann, J., Brown, P. J., Brownstein, J. R., Bundy, K., Busca, N. G., Carithers, W., Carnero, A. R., Carr, M. A., Casetti-Dinescu, D. I., Chen, Y., Chiappini, C., Comparat, J., Connolly, N., Crepp, J. R., Cristiani, S., Croft, R. A. C., Cuesta, A. J., da Costa, L. N., Davenport, J. R. A., Dawson, K. S., de Putter, R., De Lee, N., Delubac, T., Dhital, S., Ealet, A., Ebelke, G. L., Edmondson, E. M., Eisenstein, D. J., Escoffier, S., Esposito, M., Evans, M. L., Fan, X., Femenía Castellá, B., Fernández Alvar, E., Ferreira, L. D., Filiz Ak, N., Finley, H., Fleming, S. W., Font-Ribera, A., Frinchaboy, P. M., García-Hernández, D. A., García Pérez, A. E., Ge, J., Génova-Santos, R., Gillespie, B. A., Girardi, L., González Hernández, J. I., Grebel, E. K., Gunn, J. E., Guo, H., Haggard, D., Hamilton, J.-C., Harris, D. W., Hawley, S. L., Hearty, F. R., Ho, S., Hogg, D. W., Holtzman, J. A., Honscheid, K., Huehnerhoff, J., Ivans, I. I., Ivezić, Ž., Jacobson, H. R., Jiang, L., Johansson, J., Johnson, J. A., Kauffmann, G., Kirkby, D., Kirkpatrick, J. A., Klaene, M. A., Knapp, G. R., Kneib, J.-P., Le Goff, J.-M., Leauthaud, A., Lee, K.-G., Lee, Y. S., Long, D. C., Loomis, C. P., Lucatello, S., Lundgren, B., Lupton, R. H., Ma, B., Ma, Z., MacDonald, N., Mack, C. E., Mahadevan, S., Maia, M. A. G., Majewski, S. R., Makler, M., Malanushenko, E., Malanushenko, V., Manchado, A., Mandelbaum, R., Manera, M., Maraston, C., Margala, D., Martell, S. L., McBride, C. K., McGreer, I. D., McMahon, R. G., Ménard, B., Meszaros, S., Miralda-Escudé, J., Montero-Dorta, A. D., Montesano, F., Morrison, H. L., Muna, D., Munn, J. A., Murayama, H., Myers, A. D., Neto, A. F., Nguyen, D. C., Nichol, R. C., Nidever, D. L., Noterdaeme, P., Nuza, S. E., Ogando, R. L. C., Olmstead, M. D., Oravetz, D. J., Owen, R., Padmanabhan, N., Palanque-Delabrouille, N., Pan, K., Parejko, J. K., Parihar, P., Pâris, I., Pattarakijwanich, P., Pepper, J., Percival, W. J., Pérez-Fournon, I., Pérez-Ràfols, I., Petitjean, P., Pforr, J., Pieri, M. M., Pinsonneault, M. H., Porto de Mello, G. F., Prada, F., Price-Whelan, A. M., Raddick, M. J., Rebolo, R., Rich, J., Richards, G. T., Robin, A. C., Rocha-Pinto, H. J., Rockosi, C. M., Roe, N. A., Ross, A. J., Ross, N. P., Rossi, G., Rubiño-Martin, J. A., Samushia, L., Sanchez Almeida, J., Sánchez, A. G., Santiago, B., Sayres, C., Schlegel, D. J., Schlesinger, K. J., Schmidt, S. J., Schneider, D. P., Schultheis, M., Schwope, A. D., Scóccola, C. G., Seljak, U., Sheldon, E., Shen, Y., Shu, Y., Simmerer, J., Simmons, A. E., Skibba, R. A., Skrutskie, M. F., Slosar, A., Sobreira, F., Sobeck, J. S., Stassun, K. G., Steele, O., Steinmetz, M., Strauss, M. A., Streblyanska, A., Suzuki, N., Swanson, M. E. C., Tal, T., Thakar, A. R., Thomas, D., Thompson, B. A., Tinker, J. L., Tojeiro, R., Tremonti, C. A., Vargas Magaña, M., Verde, L., Viel, M., Vikas, S. K., Vogt, N. P., Wake, D. A., Wang, J., Weaver, B. A., Weinberg, D. H., Weiner, B. J., West, A. A., White, M., Wilson, J. C., Wisniewski, J. P., Wood-Vasey, W. M., Yanny, B., Yèche, C., York, D. G., Zamora, O., Zasowski, G., Zehavi, I., Zhao, G.-B., Zheng, Z., Zhu, G., & Zinn, J. C. (2012). The Ninth Data Release of the Sloan Digital Sky Survey: First Spectroscopic Data from the SDSS-III Baryon Oscillation Spectroscopic Survey. ApJS, 203(2), 21.

4. Alam, S., Albareti, F. D., Prieto, C. A., Anders, F., Anderson, S. F., Anderton, T., Andrews, B. H., Armengaud, E., Aubourg, É., Bailey, S., Basu, S., Bautista, J. E., Beaton, R. L., Beers, T. C., Bender, C. F., Berlind, A. A., Beutler, F., Bhardwaj, V., Bird, J. C., Bizyaev, D., Blake, C. H., Blanton, M. R., Blomqvist, M., Bochanski, J. J., Bolton, A. S., Bovy, J., Bradley, A. S., Brandt, W. N., Brauer, D. E., Brinkmann, J., Brown, P. J., Brownstein, J. R., Burden, A., Burtin, E., Busca, N. G., Cai, Z., Capozzi, D., Rosell, A. C., Carr, M. A., Carrera, R., Chambers, K. C., Chaplin, W. J., Chen, Y.-C., Chiappini, C., Chojnowski, S. D., Chuang, C.-H., Clerc, N., Comparat, J., Covey, K., Croft, R. A. C., Cuesta, A. J., Cunha, K., da Costa, L. N., Rio, N. D., Davenport, J. R. A., Dawson, K. S., Lee, N. D., Delubac, T., Deshpande, R., Dhital, S., Dutra-Ferreira, L., Dwelly, T., Ealet, A., Ebelke, G. L., Edmondson, E. M., Eisenstein, D. J., Ellsworth, T., Elsworth, Y., Epstein, C. R., Eracleous, M., Escoffier, S., Esposito, M., Evans, M. L., Fan, X., Fernández-Alvar, E., Feuillet, D., Ak, N. F., Finley, H., Finoguenov, A., Flaherty, K., Fleming, S. W., Font-Ribera, A., Foster, J., Frinchaboy, P. M., Galbraith-Frew, J. G., García, R. A., García-Hernández, D. A., Pérez, A. E. G., Gaulme, P., Ge, J., Génova-Santos, R., Georgakakis, A., Ghezzi, L., Gillespie, B. A., Girardi, L., Goddard, D., Gontcho, S. G. A., Hernández, J. I. G., Grebel, E. K., Green, P. J., Grieb, J. N., Grieves, N., Gunn, J. E., Guo, H., Harding, P., Hasselquist, S., Hawley, S. L., Hayden, M., Hearty, F. R., Hekker, S., Ho, S., Hogg, D. W., Holley-Bockelmann, K., Holtzman, J. A., Honscheid, K., Huber, D., Huehnerhoff, J., Ivans, I. I., Jiang, L., Johnson, J. A., Kinemuchi, K., Kirkby, D., Kitaura, F., Klaene, M. A., Knapp, G. R., Kneib, J.-P., Koenig, X. P., Lam, C. R., Lan, T.-W., Lang, D., Laurent, P., Goff, J.-M. L., Leauthaud, A., Lee, K.-G., Lee, Y. S., Licquia, T. C., Liu, J., Long, D. C., López-Corredoira, M., Lorenzo-Oliveira, D., Lucatello, S., Lundgren, B., Lupton, R. H., III, C. E. M., Mahadevan, S., Maia, M. A. G., Majewski, S. R., Malanushenko, E., Malanushenko, V., Manchado, A., Manera, M., Mao, Q., Maraston, C., Marchwinski, R. C., Margala, D., Martell, S. L., Martig, M., Masters, K. L., Mathur, S., McBride, C. K., McGehee, P. M., McGreer, I. D., McMahon, R. G., Ménard, B., Menzel, M.-L., Merloni, A., Mészáros, S., Miller, A. A., Miralda-Escudé, J., Miyatake, H., Montero-Dorta, A. D., More, S., Morganson, E., Morice-Atkinson, X., Morrison, H. L., Mosser, B., Muna, D., Myers, A. D., Nandra, K., Newman, J. A., Neyrinck, M., Nguyen, D. C., Nichol, R. C., Nidever,

D. L., Noterdaeme, P., Nuza, S. E., O'Connell, J. E., O'Connell, R. W., O'Connell, R., Ogando, R. L. C., Olmstead, M. D., Oravetz, A. E., Oravetz, D. J., Osumi, K., Owen, R., Padgett, D. L., Padmanabhan, N., Paegert, M., Palanque-Delabrouille, N., Pan, K., Parejko, J. K., Pâris, I., Park, C., Pattarakijwanich, P., Pellejero-Ibanez, M., Pepper, J., Percival, W. J., Pérez-Fournon, I., Pe´rez-Ra`fols, I., Petitjean, P., Pieri, M. M., Pinsonneault, M. H., de Mello, G. F. P., Prada, F., Prakash, A., Price-Whelan, A. M., Protopapas, P., Raddick, M. J., Rahman, M., Reid, B. A., Rich, J., Rix, H.-W., Robin, A. C., Rockosi, C. M., Rodrigues, T. S., Rodríguez-Torres, S., Roe, N. A., Ross, A. J., Ross, N. P., Rossi, G., Ruan, J. J., Rubiño-Martín, J. A., Rykoff, E. S., Salazar-Albornoz, S., Salvato, M., Samushia, L., Sánchez, A. G., Santiago, B., Sayres, C., Schiavon, R. P., Schlegel, D. J., Schmidt, S. J., Schneider, D. P., Schultheis, M., Schwope, A. D., Scóccola, C. G., Scott, C., Sellgren, K., Seo, H.-J., Serenelli, A., Shane, N., Shen, Y., Shetrone, M., Shu, Y., Aguirre, V. S., Sivarani, T., Skrutskie, M. F., Slosar, A., Smith, V. V., Sobreira, F., Souto, D., Stassun, K. G., Steinmetz, M., Stello, D., Strauss, M. A., Streblyanska, A., Suzuki, N., Swanson, M. E. C., Tan, J. C., Tayar, J., Terrien, R. C., Thakar, A. R., Thomas, D., Thomas, N., Thompson, B. A., Tinker, J. L., Tojeiro, R., Troup, N. W., Vargas-Magaña, M., Vazquez, J. A., Verde, L., Viel, M., Vogt, N. P., Wake, D. A., Wang, J., Weaver, B. A., Weinberg, D. H., Weiner, B. J., White, M., Wilson, J. C., Wisniewski, J. P., Wood-Vasey, W. M., Ye`che, C., York, D. G., Zakamska, N. L., Zamora, O., Zasowski, G., Zehavi, I., Zhao, G.-B., Zheng, Z., (周旭), X. Z., (周志民), Z. Z., (邹虎), H. Z., & Zhu, G. (2015). The Eleventh and Twelfth Data Releases of the Sloan Digital Sky Survey: Final Data from SDSS-III. *ApJSS*, 219(1), 12.

5. Alegre, L., Sabater, J., Best, P., Mostert, R. I. J., Williams, W. L., Gürkan, G., Hardcastle, M. J., Kondapally, R., Shimwell, T. W., & Smith, D. J. B. (2022). A machine-learning classifier for LOFAR radio galaxy cross-matching techniques. MNRAS, 516(4), 4716–4738.

6. Altman, C. & Suchy, K. (2011). *Reciprocity, Spatial Mapping and Time Reversal in Electromagnetics*. Dordrecht: Springer.

7. Amirkhanyan, V. R. (2016). Radiation pattern of radio and optical components of extended radio sources. *Astrophysical Bulletin*, 71(4), 384–394.

8. Andernach, H., Feretti, L., Giovannini, G., Klein, U., Rossetti, E., & Schnaubelt, J. (1992). The large-scale radio structure of 3C 31 and 3C 449. A&AS, 93, 331–357.

9. Andernach, H., Jiménez-Andrade, E. F., & Willis, A. G. (2021). Discovery of 178 Giant Radio Galaxies in 1059 deg^2 of the Rapid ASKAP Continuum Survey at 888 MHz. *Galaxies*, 9(4), 99.

10. Anduaga, A. (2021). The formation of ionospheric physics - confluence of traditions and threads of continuity. *History of Geo- and Space Sciences*, 12(1), 57–75.

11. Angulo, R. E., Springel, V., White, S. D. M., Jenkins, A., Baugh, C. M., & Frenk, C. S. (2012). Scaling relations for galaxy clusters in the Millennium-XXL simulation. MNRAS, 426(3), 2046–2062.

12. Araya-Melo, P. A., Aragón-Calvo, M. A., Brüggen, M., & Hoeft, M. (2012). Radio emission in the cosmic web. MNRAS, 423(3), 2325–2341.

13. Arnaud, M. (2009). The β-model of the intracluster medium. Commentary on: Cavaliere A. and Fusco-Femiano R., 1976, A&A, 49, 137. A&A, 500(1), 103–104.

14. Arora, B. S., Morgan, J., Ord, S. M., Tingay, S. J., Bell, M., Callingham, J. R., Dwarakanath, K. S., For, B. Q., Hancock, P., Hindson, L., Hurley-Walker, N., Johnston-Hollitt, M., Kapińska, A. D., Lenc, E., McKinley, B., Offringa, A. R., Procopio, P., Staveley-Smith, L., Wayth, R. B., Wu, C., & Zheng, Q. (2016). Ionospheric Modelling using GPS to Calibrate the MWA. II: Regional Ionospheric Modelling using GPS and GLONASS to Estimate Ionospheric Gradients. PASA, 33, e031.

15. Attia, O., Teyssier, R., Katz, H., Kimm, T., Martin-Alvarez, S., Ocvirk, P., & Rosdahl, J. (2021). Cosmological magnetogenesis: the Biermann battery during the Epoch of Reionization. *Monthly notices of the Royal Astronomical Society*, 504(2), 2346–2359.

16. Axford, W. I., Leer, E., & Skadron, G. (1977). The Acceleration of Cosmic Rays by Shock Waves. In *International Cosmic Ray Conference*, volume 11 of *International Cosmic Ray Conference* (pp. 132).

17. Bagchi, J., Vivek, M., Vikram, V., Hota, A., Biju, K. G., Sirothia, S. K., Srianand, R., Gopal-Krishna, & Jacob, J. (2014). Megaparsec Relativistic Jets Launched from an Accreting Supermassive Black Hole in an Extreme Spiral Galaxy. ApJ, 788(2), 174.

18. Baring, M. G. (1997). Diffusive Shock Acceleration : the Fermi Mechanism. In Y. Giraud-Heraud & J. Tran Thanh van (Eds.), *Very High Energy Phenomena in the Universe; Moriond Workshop* (pp.97).

19. Barkus, B., Croston, J. H., Piotrowska, J., Mingo, B., Best, P. N., Hardcastle, M. J., Mostert, R. I. J., Röttgering, H. J. A., Sabater, J., Webster, B., & Williams, W. L. (2022). The application of ridgelines in extended radio source cross-identification. MNRAS, 509(1), 1–15.

20. Barrow, J. D., Tsagas, C. G., & Yamamoto, K. (2012). Origin of cosmic magnetic fields: Superadiabatically amplified modes in open Friedmann universes. Phys. Rev. D, 86(2), 023533.

21. Bassani, L., Ursini, F., Malizia, A., Bruni, G., Panessa, F., Masetti, N., Saviane, I., Monaco, L., Venturi, T., Dallacasa, D., Bazzano, A., & Ubertini, P. (2021). Soft gamma-ray selected giant radio galaxies: an update. MNRAS, 500(3), 3111–3122.

22. Beck, R. & Krause, M. (2005). Revised equipartition and minimum energy formula for magnetic field strength estimates from radio synchrotron observations. *Astronomische Nachrichten*, 326(6), 414–427.

23. Beckmann, R. S., Devriendt, J., Slyz, A., Peirani, S., Richardson, M. L. A., Dubois, Y., Pichon, C., Chisari, N. E., Kaviraj, S., Laigle, C., & Volonteri, M. (2017). Cosmic evolution of stellar quenching by AGN feedback: clues from the Horizon-AGN simulation. *Monthly Notices of the Royal Astronomical Society*, 472(1), 949–965.

24. Bédard, M. (2008). Optimal acceptance rates for Metropolis algorithms: Moving beyond 0.234. *Stochastic Processes and their Applications*, 118(12), 2198–2222.

25. Behroozi, P. S., Wechsler, R. H., & Conroy, C. (2013). The Average Star Formation Histories of Galaxies in Dark Matter Halos from z = 0-8. ApJ, 770(1), 57.

26. Bell, A. R. (1978a). The acceleration of cosmic rays in shock fronts - I. MNRAS, 182, 147–156.

27. Bell, A. R. (1978b). The acceleration of cosmic rays in shock fronts - II. MNRAS, 182, 443–455.

28. Benatar, D. (2006). *Better Never to Have Been: The Harm of Coming into Existence*. Oxford: Oxford University Press.

29. Best, P. N. & Heckman, T. M. (2012). On the fundamental dichotomy in the local radio-AGN population: accretion, evolution and host galaxy properties. MNRAS, 421(2), 1569–1582.

30. Best, P. N., Ker, L. M., Simpson, C., Rigby, E. E., & Sabater, J. (2014). The cosmic evolution of radio-AGN feedback to $z = 1$. MNRAS, 445(1), 955–969.

31. Bilitza, D. & Reinisch, B. W. (2008). International Reference Ionosphere 2007: Improvements and new parameters. *Advances in Space Research*, 42, 599–609.

32. Biretta, J. A., Zhou, F., & Owen, F. N. (1995). Detection of Proper Motions in the M87 Jet. ApJ, 447, 582.

33. Blandford, R. & Eichler, D. (1987). Particle acceleration at astrophysical shocks: A theory of cosmic ray origin. Phys. Rep., 154(1), 1–75.

34. Blandford, R., Meier, D., & Readhead, A. (2019). Relativistic Jets from Active Galactic Nuclei. ARA&A, 57, 467–509.

35. Blandford, R. D. & Ostriker, J. P. (1978). Particle acceleration by astrophysical shocks. ApJ, 221, L29–L32.

36. Blandford, R. D. & Rees, M. J. (1974). A "twin-exhaust" model for double radio sources. MNRAS, 169, 395–415.

37. Blandford, R. D. & Znajek, R. L. (1977). Electromagnetic extraction of energy from Kerr black holes. MNRAS, 179, 433–456.

38. Bloom, J. S., Frail, D. A., & Sari, R. (2001). The Prompt Energy Release of Gamma-Ray Bursts using a Cosmological k-Correction. AJ, 121(6), 2879–2888.

39. Boch, T. & Fernique, P. (2014). Aladin Lite: Embed your Sky in the Browser. In N. Manset & P. Forshay (Eds.), *Astronomical Data Analysis Software and Systems XXIII*, volume 485 of *Astronomical Society of the Pacific Conference Series* (pp. 277).

40. Bolton, J. G. (1948). Discrete Sources of Galactic Radio Frequency Noise. Nature, 162(4108), 141–142.

41. Bonnarel, F., Fernique, P., Bienaymé, O., Egret, D., Genova, F., Louys, M., Ochsenbein, F., Wenger, M., & Bartlett, J. G. (2000). The ALADIN interactive sky atlas. A reference tool for identification of astronomical sources. A&AS, 143, 33–40.

42. Bos, E. G. P., Kitaura, F.-S., & van de Weygaert, R. (2019). Bayesian cosmic density field inference from redshift space dark matter maps. MNRAS, 488(2), 2573–2604.

43. Botteon, A., Brunetti, G., van Weeren, R. J., Shimwell, T. W., Pizzo, R. F., Cassano, R., Iacobelli, M., Gastaldello, F., Bîrzan, L., Bonafede, A., Brüggen, M., Cuciti, V., Dallacasa, D., de Gasperin, F., Di Gennaro, G., Drabent, A., Hardcastle, M. J., Hoeft, M., Mandal, S., Röttgering, H. J. A., & Simionescu, A. (2020a). The Beautiful Mess in Abell 2255. ApJ, 897(1), 93.

44. Botteon, A., Shimwell, T. W., Bonafede, A., Dallacasa, D., Brunetti, G., Mandal, S., van Weeren, R. J., Brüggen, M., Cassano, R., de Gasperin, F., Hoang, D. N., Hoeft, M., Röttgering, H. J. A., Savini, F., White, G. J., Wilber, A., & Venturi, T. (2018). LOFAR discovery of a double radio halo system in Abell 1758 and radio/X-ray study of the cluster pair. MNRAS, 478(1), 885–898.

45. Botteon, A., van Weeren, R. J., Brunetti, G., de Gasperin, F., Intema, H. T., Osinga, E., Di Gennaro, G., Shimwell, T. W., Bonafede, A., Brüggen, M., Cassano, R., Cuciti, V., Dallacasa, D., Gastaldello, F., Mandal, S., Rossetti, M., & Röttgering, H. J. A. (2020b). A giant radio bridge connecting two galaxy clusters in Abell 1758. MNRAS, 499(1), L11–L15.

46. Botteon, A., van Weeren, R. J., Brunetti, G., Vazza, F., Shimwell, T. W., Brüggen, M., Röttgering, H. J. A., de Gasperin, F., Akamatsu, H., Bonafede, A., Cassano, R., Cuciti, V., Dallacasa, D., Gennaro, G. D., & Gastaldello, F. (2022). Magnetic

fields and relativistic electrons fill entire galaxy cluster. *Science Advances*, 8(44), eabq7623.

47. Bouchet, F. R., Colombi, S., Hivon, E., & Juszkiewicz, R. (1995). Perturbative Lagrangian approach to gravitational instability. A&A, 296, 575.

48. Bouman, K. L., Johnson, M. D., Zoran, D., Fish, V. L., Doeleman, S. S., & Freeman, W. T. (2016). Computational Imaging for VLBI Image Reconstruction. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

49. Boxelaar, J. M., van Weeren, R. J., & Botteon, A. (2021). A robust model for flux density calculations of radio halos in galaxy clusters: Halo-FDCA. *Astronomy and Computing*, 35, 100464.

50. Brandenburg, A. & Subramanian, K. (2005). Astrophysical magnetic fields and nonlinear dynamo theory. *Physics Reports*, 417(1), 1–209.

51. Bregman, J. N. & Irwin, J. A. (2007). The Search for Million Degree Gas through the N VII Hyperfine Line. ApJ, 666(1), 139–146.

52. Breiman, L. (1997). *Arcing the edge*. Technical report, Citeseer.

53. Brown, S., Vernstrom, T., Carretti, E., Dolag, K., Gaensler, B. M., Staveley-Smith, L., Bernardi, G., Haverkorn, M., Kesteven, M., & Poppi, S. (2017). Limiting magnetic fields in the cosmic web with diffuse radio emission. MNRAS, 468(4), 4246–4253.

54. Brüggen, M., Reiprich, T. H., Bulbul, E., Koribalski, B. S., Andernach, H., Rudnick, L., Hoang, D. N., Wilber, A. G., Duchesne, S. W., Veronica, A., Pacaud, F., Hopkins, A. M., Norris, R. P., Johnston-Hollitt, M., Brown, M. J. I., Bonafede, A., Brunetti, G., Collier, J. D., Sanders, J. S., Vardoulaki, E., Venturi, T., Kapinska, A. D., & Marvil, J. (2021). Radio observations of the merging galaxy cluster system Abell 3391-Abell 3395. A&A, 647, A3.

55. Brunetti, G., Setti, G., Feretti, L., & Giovannini, G. (2001). Particle reacceleration in the Coma cluster: radio properties and hard X-ray emission. MNRAS, 320(3), 365–378.

56. Brunetti, G. & Vazza, F. (2020). Second-order Fermi Reacceleration Mechanisms and Large-Scale Synchrotron Radio Emission in Intracluster Bridges. *Phys. Rev. Lett.*, 124, 051101.

57. Bryan, G. L., Norman, M. L., O'Shea, B. W., Abel, T., Wise, J. H., Turk, M. J., Reynolds, D. R., Collins, D. C., Wang, P., Skillman, S. W., Smith, B., Harkness, R. P., Bordner, J., Kim, J.-h., Kuhlen, M., Xu, H., Goldbaum, N., Hummels, C., Kritsuk, A. G., Tasker, E., Skory, S., Simpson, C. M., Hahn, O., Oishi, J. S., So, G. C., Zhao, F., Cen, R., Li, Y., & Enzo Collaboration (2014). ENZO: An Adaptive Mesh Refinement Code for Astrophysics. ApJS, 211(2), 19.

58. Burbidge, G. R. (1956). On Synchrotron Radiation from Messier 87. ApJ, 124, 416.

59. Byrd, R., Lu, P., Nocedal, J., & Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, 16(5), 1190–1208.

60. Cantwell, T. M., Bray, J. D., Croston, J. H., Scaife, A. M. M., Mulcahy, D. D., Best, P. N., Brüggen, M., Brunetti, G., Callingham, J. R., Clarke, A. O., Hardcastle, M. J., Harwood, J. J., Heald, G., Heesen, V., Iacobelli, M., Jamrozy, M., Morganti, R., Orrú, E., O'Sullivan, S. P., Riseley, C. J., Röttgering, H. J. A., Shulevski, A., Sridhar, S. S., Tasse, C., & Van Eck, C. L. (2020). Low-frequency observations of the giant radio galaxy NGC 6251. MNRAS, 495(1), 143–159.

61. Caprioli, D. & Haggerty, C. (2019). The Issue with Diffusive Shock Acceleration. In *36th International Cosmic Ray Conference (ICRC2019)*, volume 36 of *International Cosmic Ray Conference* (pp. 209).

62. Cargill, P. J. (2007). Fundamentals of plasma physics. *Plasma Physics and Controlled Fusion*, 49(2), 197.

63. Carretti, E., O'Sullivan, S. P., Vacca, V., Vazza, F., Gheller, C., Vernstrom, T., & Bonafede, A. (2023). Magnetic field evolution in cosmic filaments with LOFAR data. MNRAS, 518(2), 2273–2286.

64. Cavaliere, A. & Fusco-Femiano, R. (1976). X-rays from hot plasma in clusters of galaxies. A&A, 49, 137–144.

65. Cavaliere, A. & Fusco-Femiano, R. (1978). The Distribution of Hot Gas in Clusters of Galaxies. A&A, 70, 677.

66. Cavaliere, A. & Rephaeli, Y. (2011). *Astrophysics of Galaxy Clusters*, volume 172. IOS Press.

67. Cen, R. & Ostriker, J. P. (1999). Where Are the Baryons? ApJ, 514(1), 1–6.

68. Chambers, K. C., Magnier, E. A., Metcalfe, N., Flewelling, H. A., Huber, M. E., Waters, C. Z., Denneau, L., Draper, P. W., Farrow, D., Finkbeiner, D. P., Holmberg, C., Koppenhoefer, J., Price, P. A., Rest, A., Saglia, R. P., Schlafly, E. F., Smartt, S. J., Sweeney, W., Wainscoat, R. J., Burgett, W. S., Chastel, S., Grav, T., Heasley, J. N., Hodapp, K. W., Jedicke, R., Kaiser, N., Kudritzki, R. P., Luppino, G. A., Lupton, R. H., Monet, D. G., Morgan, J. S., Onaka, P. M., Shiao, B., Stubbs, C. W., Tonry, J. L., White, R., Bañados, E., Bell, E. F., Bender, R., Bernard, E. J., Boegner, M., Boffi, F., Botticella, M. T., Calamida, A., Casertano, S., Chen, W. P., Chen, X., Cole, S., Deacon, N., Frenk, C., Fitzsimmons, A., Gezari, S., Gibbs, V., Goessl, C., Goggia, T., Gourgue, R., Goldman, B., Grant, P., Grebel, E. K., Hambly, N. C., Hasinger, G., Heavens, A. F., Heckman, T. M., Henderson, R., Henning, T., Holman, M., Hopp, U., Ip, W. H., Isani, S., Jackson, M., Keyes, C. D., Koekemoer, A. M., Kotak, R., Le, D., Liska, D., Long, K. S., Lucey, J. R., Liu, M., Martin, N. F., Masci, G., McLean, B., Mindel, E.,

Misra, P., Morganson, E., Murphy, D. N. A., Obaika, A., Narayan, G., Nieto-Santisteban, M. A., Norberg, P., Peacock, J. A., Pier, E. A., Postman, M., Primak, N., Rae, C., Rai, A., Riess, A., Riffeser, A., Rix, H. W., Röser, S., Russel, R., Rutz, L., Schilbach, E., Schultz, A. S. B., Scolnic, D., Strolger, L., Szalay, A., Seitz, S., Small, E., Smith, K. W., Soderblom, D. R., Taylor, P., Thomson, R., Taylor, A. N., Thakar, A. R., Thiel, J., Thilker, D., Unger, D., Urata, Y., Valenti, J., Wagner, J., Walder, T., Walter, F., Watters, S. P., Werner, S., Wood-Vasey, W. M., & Wyse, R. (2016). The Pan-STARRS1 Surveys. *arXiv e-prints*, (pp. arXiv:1612.05560).

69. Chang, Y.-Y., van der Wel, A., da Cunha, E., & Rix, H.-W. (2015). Stellar Masses and Star Formation Rates for 1M Galaxies from SDSS+WISE. ApJS, 219(1), 8.

70. Christodoulou, D. (1970). Reversible and Irreversible Transformations in Black-Hole Physics. Phys. Rev. Lett., 25(22), 1596–1597.

71. Cohen, M. H. (1973). Introduction to very-long-baseline interferometry. *IEEE Proceedings*, 61, 1192–1197.

72. Cole, S., Percival, W. J., Peacock, J. A., Norberg, P., Baugh, C. M., Frenk, C. S., Baldry, I., Bland-Hawthorn, J., Bridges, T., Cannon, R., Colless, M., Collins, C., Couch, W., Cross, N. J. G., Dalton, G., Eke, V. R., de Propris, R., Driver, S. P., Efstathiou, G., Ellis, R. S., Glazebrook, K., Jackson, C., Jenkins, A., Lahav, O., Lewis, I., Lumsden, S., Maddox, S., Madgwick, D., Peterson, B. A., Sutherland, W., Taylor, K., & 2dFGRS Team, T. (2005). The 2dF Galaxy Redshift Survey: power-spectrum analysis of the final data set and cosmological implications. *Monthly Notices of the Royal Astronomical Society*, 362(2), 505–534.

73. Condon, J. J., Anderson, E., & Broderick, J. J. (1995). Radio Identifications of Extragalactic IRAS Sources. AJ, 109, 2318.

74. Condon, J. J., Cotton, W. D., Fomalont, E. B., Kellermann, K. I., Miller, N., Perley, R. A., Scott, D., Vernstrom, T., & Wall, J. V. (2012). Resolving the Radio Source Background: Deeper Understanding through Confusion. ApJ, 758(1), 23.

75. Condon, J. J., Cotton, W. D., Greisen, E. W., Yin, Q. F., Perley, R. A., Taylor, G. B., & Broderick, J. J. (1998). The NRAO VLA Sky Survey. AJ, 115(5), 1693–1716.

76. Connor, L., Bouman, K. L., Ravi, V., & Hallinan, G. (2022). Deep radio-interferometric imaging with POLISH: DSA-2000 and weak lensing. MNRAS, 514(2), 2614–2626.

77. Conselice, C. J., Wilkinson, A., Duncan, K., & Mortlock, A. (2016). The Evolution of Galaxy Number Density at $z < 8$ and Its Implications. ApJ, 830(2), 83.

78. Cooper, B. F. C., Price, R. M., & Cole, D. J. (1965). A study of the decimetric emission and polarization of Centaurus A. *Australian Journal of Physics*, 18, 589.

79. Crain, R. A., Schaye, J., Bower, R. G., Furlong, M., Schaller, M., Theuns, T., Dalla Vecchia, C., Frenk, C. S., McCarthy, I. G., Helly, J. C., Jenkins, A., Rosas-Guevara, Y. M., White, S. D. M., & Trayford, J. W. (2015). The EAGLE simulations of galaxy formation: calibration of subgrid physics and model variations. MNRAS, 450(2), 1937–1961.

80. Croston, J. H., Hardcastle, M. J., Mingo, B., Best, P. N., Sabater, J., Shimwell, T. M., Williams, W. L., Duncan, K. J., Röttgering, H. J. A., Brienza, M., Gürkan, G., Ineson, J., Miley, G. K., Morabito, L. M., O'Sullivan, S. P., & Prandoni, I. (2019). The environments of radio-loud AGN from the LOFAR Two-Metre Sky Survey (LoTSS). A&A, 622, A10.

81. Croston, J. H., Ineson, J., & Hardcastle, M. J. (2018). Particle content, radio-galaxy morphology, and jet power: all radio-loud AGN are not equal. MNRAS, 476(2), 1614–1623.

82. Croton, D. J., Springel, V., White, S. D. M., De Lucia, G., Frenk, C. S., Gao, L., Jenkins, A., Kauffmann, G., Navarro, J. F., & Yoshida, N. (2006). The many lives of active galactic nuclei: cooling flows, black holes and the luminosities and colours of galaxies. MNRAS, 365(1), 11–28.

83. Curtis, H. D. (1918). Descriptions of 762 Nebulae and Clusters Photographed with the Crossley Reflector. *Publications of Lick Observatory*, 13, 9–42.

84. Cutri, R. M. & et al. (2012). VizieR Online Data Catalog: WISE All-Sky Data Release (Cutri+ 2012). *VizieR Online Data Catalog*, (pp. II/311).

85. Cutri, R. M., Wright, E. L., Conrow, T., Fowler, J. W., Eisenhardt, P. R. M., Grillmair, C., Kirkpatrick, J. D., Masci, F., McCallon, H. L., Wheelock, S. L., Fajardo-Acosta, S., Yan, L., Benford, D., Harbut, M., Jarrett, T., Lake, S., Leisawitz, D., Ressler, M. E., Stanford, S. A., Tsai, C. W., Liu, F., Helou, G., Mainzer, A., Gettngs, D., Gonzalez, A., Hoffman, D., Marsh, K. A., Padgett, D., Skrutskie, M. F., Beck, R., Papin, M., & Wittman, M. (2021). VizieR Online Data Catalog: AllWISE Data Release (Cutri+ 2013). *VizieR Online Data Catalog*, (pp. II/328).

86. Dabhade, P., Gaikwad, M., Bagchi, J., Pandey-Pommier, M., Sankhyayan, S., & Raychaudhury, S. (2017). Discovery of giant radio galaxies from NVSS: radio and infrared properties. MNRAS, 469(3), 2886–2906.

87. Dabhade, P., Mahato, M., Bagchi, J., Saikia, D. J., Combes, F., Sankhyayan, S., Röttgering, H. J. A., Ho, L. C., Gaikwad, M., Raychaudhury, S., Vaidya, B., & Guiderdoni, B. (2020a). Search and analysis of giant radio galaxies with associated nuclei (SAGAN). I. New sample and multi-wavelength studies. A&A, 642, A153.

88. Dabhade, P., Röttgering, H. J. A., Bagchi, J., Shimwell, T. W., Hardcastle, M. J., Sankhyayan, S., Morganti, R., Jamrozy, M., Shulevski, A., & Duncan, K. J. (2020b). Giant radio galaxies in the LOFAR Two-metre Sky Survey. I. Radio and environmental properties. A&A, 635, A5.

89. Dabhade, P., Saikia, D. J., & Mahato, M. (2023). Decoding the giant extragalactic radio sources. *Journal of Astrophysics and Astronomy*, 44(1), 13.

90. Dálya, G., Galgóczi, G., Dobos, L., Frei, Z., Heng, I. S., Macas, R., Messenger, C., Raffai, P., & de Souza, R. S. (2018). GLADE: A galaxy catalogue for multimessenger searches in the advanced gravitational-wave detector era. MNRAS, 479(2), 2374–2381.

91. de Gasperin, F., Mevius, M., Rafferty, D., Intema, H., & Fallows, R. (2018). The effect of the ionosphere on ultra-low frequency radio-interferometric observations. A&A, 615.

92. de Graaff, A., Cai, Y.-C., Heymans, C., & Peacock, J. A. (2019). Probing the missing baryons with the Sunyaev–Zel'dovich effect from filaments. A&A, 624, A48.

93. de Jong, J. M. G. H. J., van Weeren, R. J., Botteon, A., Oonk, J. B. R., Brunetti, G., Shimwell, T. W., Cassano, R., Röttgering, H. J. A., & Tasse, C. (2022). Deep study of A399-401: Application of a wide-field facet calibration. A&A, 668, A107.

94. de Ruiter, H. R., Willis, A. G., & Arp, H. C. (1977). A Westerbork 1415 MHz survey of background radio sources. II. Optical identifications with deep IIIa-J plates. A&AS, 28, 211–293.

95. Delhaize, J., Heywood, I., Prescott, M., Jarvis, M. J., Delvecchio, I., Whittam, I. H., White, S. V., Hardcastle, M. J., Hale, C. L., Afonso, J., Ao, Y., Brienza, M., Brüggen, M., Collier, J. D., Daddi, E., Glowacki, M., Maddox, N., Morabito, L. K., Prandoni, I., Randriamanakoto, Z., Sekhar, S., An, F., Adams, N. J., Blyth, S., Bowler, R. A. A., Leeuw, L., Marchetti, L., Randriamampandry, S. M., Thorat, K., Seymour, N., Smirnov, O., Taylor, A. R., Tasse, C., & Vaccari, M. (2021). MIGHTEE: are giant radio galaxies more common than we thought? MNRAS, 501(3), 3833–3845.

96. Dewdney, P. E., Hall, P. J., Schilizzi, R. T., & Lazio, T. J. L. W. (2009). The Square Kilometre Array. *IEEE Proceedings*, 97(8), 1482–1496.

97. Dey, A., Schlegel, D. J., Lang, D., Blum, R., Burleigh, K., Fan, X., Findlay, J. R., Finkbeiner, D., Herrera, D., Juneau, S., Landriau, M., Levi, M., McGreer, I., Meisner, A., Myers, A. D., Moustakas, J., Nugent, P., Patej, A., Schlafly, E. F., Walker, A. R., Valdes, F., Weaver, B. A., Yèche, C., Zou, H., Zhou, X., Abareshi, B., Abbott, T. M. C., Abolfathi, B., Aguilera, C., Alam, S., Allen, L., Alvarez, A., Annis, J., Ansarinejad, B., Aubert, M., Beechert, J., Bell, E. F., BenZvi, S. Y., Beutler, F., Bielby, R. M., Bolton, A. S., Briceño, C., Buckley-Geer, E. J., Butler, K., Calamida, A., Carlberg, R. G., Carter, P., Casas, R., Castander, F. J., Choi, Y., Comparat, J., Cukanovaite, E., Delubac, T., DeVries, K., Dey, S., Dhungana, G., Dickinson, M., Ding, Z., Donaldson, J. B., Duan, Y., Duckworth, C. J., Eftekharzadeh, S., Eisenstein, D. J., Etourneau, T., Fagrelius, P. A., Farihi, J., Fitzpatrick, M., Font-Ribera, A., Fulmer, L., Gänsicke, B. T., Gaztanaga, E., George, K., Gerdes, D. W., Gontcho, S. G. A., Gorgoni, C., Green, G., Guy, J., Harmer, D., Hernandez, M., Honscheid, K., Huang, L. W., James, D. J., Jannuzi, B. T., Jiang, L., Joyce, R., Karcher, A., Karkar, S., Kehoe, R., Jean-Paul, K., Kueter-Young, A., Lan, T.-W., Lauer, T. R., Guillou, L. L., Suu, A. L. V., Lee, J. H., Lesser, M., Levasseur, L. P., Li, T. S., Mann, J. L., Marshall, R., Martínez-Vázquez, C. E., Martini, P., du Mas des Bourboux, H., McManus, S., Meier, T. G., Ménard, B., Metcalfe, N., Muñoz-Gutiérrez, A., Najita, J., Napier, K., Narayan, G., Newman, J. A., Nie, J., Nord, B., Norman, D. J., Olsen, K. A. G., Paat, A., Palanque-Delabrouille, N., Peng, X., Poppett, C. L., Poremba, M. R., Prakash, A., Rabinowitz, D., Raichoor, A., Rezaie, M., Robertson, A. N., Roe, N. A., Ross, A. J., Ross, N. P., Rudnick, G., Safonova, S., Saha, A., Sánchez, F. J., Savary, E., Schweiker, H., Scott, A., Seo, H.-J., Shan, H., Silva, D. R., Slepian, Z., Soto, C., Sprayberry, D., Staten, R., Stillman, C. M., Stupak, R. J., Summers, D. L., Tie, S. S., Tirado, H., Vargas-Magaña, M., Vivas, A. K., Wechsler, R. H., Williams, D., Yang, J., Yang, Q., Yapici, T., Zaritsky, D., Zenteno, A., Zhang, K., Zhang, T., Zhou, R., & Zhou, Z. (2019). Overview of the DESI legacy imaging surveys. *AJ*, 157(5), 168.

98. Di Gennaro, G., van Weeren, R. J., Hoeft, M., Kang, H., Ryu, D., Rudnick, L., Forman, W., Röttgering, H. J. A., Brüggen, M., Dawson, W. A., Golovich, N., Hoang, D. N., Intema, H. T., Jones, C., Kraft, R. P., Shimwell, T. W., & Stroe, A. (2018). Deep Very Large Array Observations of the Merging Cluster CIZA J2242.8+5301: Continuum and Spectral Imaging. ApJ, 865(1), 24.

99. Di Matteo, T., Springel, V., & Hernquist, L. (2005). Energy input from quasars regulates the growth and activity of black holes and their host galaxies. Nature, 433(7026), 604–607.

100. Dreher, J. W. & Feigelson, E. D. (1984). Rings and wiggles in Hercules A. Nature, 308, 43–45.

101. Dressel, L. L. & Condon, J. J. (1978). The Arecibo 2380 MHz survey of bright galaxies. ApJS, 36, 53–75.

102. Driver, S. (2021). The challenge of measuring and mapping the missing baryons. *Nature Astronomy*, 5, 852–854.

103. Driver, S. P., Norberg, P., Baldry, I. K., Bamford, S. P., Hopkins, A. M., Liske, J., Loveday, J., Peacock, J. A., Hill, D. T., Kelvin, L. S., Robotham, A. S. G., Cross, N. J. G., Parkinson, H. R., Prescott, M., Conselice, C. J., Dunne, L., Brough, S., Jones, H., Sharp, R. G., van Kampen, E., Oliver, S., Roseboom, I. G., Bland-Hawthorn, J., Croom, S. M., Ellis, S.,

Cameron, E., Cole, S., Frenk, C. S., Couch, W. J., Graham, A. W., Proctor, R., De Propris, R., Doyle, I. F., Edmondson, E. M., Nichol, R. C., Thomas, D., Eales, S. A., Jarvis, M. J., Kuijken, K., Lahav, O., Madore, B. F., Seibert, M., Meyer, M. J., Staveley-Smith, L., Phillipps, S., Popescu, C. C., Sansom, A. E., Sutherland, W. J., Tuffs, R. J., & Warren, S. J. (2009). GAMA: towards a physical understanding of galaxy formation. *Astronomy and Geophysics*, 50(5), 5.12–5.19.

104. Drury, L. O. (1983). Review article: An introduction to the theory of diffusive shock acceleration of energetic particles in tenuous plasmas. *Reports on Progress in Physics*, 46(8), 973–1027.

105. Duane, S., Kennedy, A. D., Pendleton, B. J., & Roweth, D. (1987). Hybrid Monte Carlo. *Physics Letters B*, 195(2), 216–222.

106. Dugan, Z., Gaibler, V., & Silk, J. (2017). Feedback by AGN Jets and Wide-angle Winds on a Galactic Scale. ApJ, 844(1), 37.

107. Duncan, K. J. (2022). All-purpose, all-sky photometric redshifts for the Legacy Imaging Surveys Data Release 8. MNRAS, 512(3), 3662–3683.

108. Eckert, D., Jauzac, M., Shan, H., Kneib, J.-P., Erben, T., Israel, H., Jullo, E., Klein, M., Massey, R., Richard, J., & Tchernin, C. (2015). Warm-hot baryons comprise 5–10 per cent of filaments in the cosmic web. Nature, 528(7580), 105–107.

109. Einstein, A. (1939). On a Stationary System with Spherical Symmetry Consisting of Many Gravitating Masses. *Annals of Mathematics*, 40, 922.

110. Eisenstein, D. J., Zehavi, I., Hogg, D. W., Scoccimarro, R., Blanton, M. R., Nichol, R. C., Scranton, R., Seo, H.-J., Tegmark, M., Zheng, Z., Anderson, S. F., Annis, J., Bahcall, N., Brinkmann, J., Burles, S., Castander, F. J., Connolly, A., Csabai, I., Doi, M., Fukugita, M., Frieman, J. A., Glazebrook, K., Gunn, J. E., Hendry, J. S., Hennessy, G., Ivezić, Z., Kent, S., Knapp, G. R., Lin, H., Loh, Y.-S., Lupton, R. H., Margon, B., McKay, T. A., Meiksin, A., Munn, J. A., Pope, A., Richmond, M. W., Schlegel, D., Schneider, D. P., Shimasaku, K., Stoughton, C., Strauss, M. A., SubbaRao, M., Szalay, A. S., Szapudi, I., Tucker, D. L., Yanny, B., & York, D. G. (2005). Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies. *The Astrophysical Journal*, 633(2), 560.

111. Ensslin, T. A., Biermann, P. L., Klein, U., & Kohle, S. (1998). Cluster radio relics as a tracer of shock waves of the large-scale structure formation. A&A, 332, 395–409.

112. Evans, I. N., Primini, F. A., Miller, J. B., Evans, J. D., Allen, C. E., Anderson, C. S., Becker, G., Budynkiewicz, J. A., Burke, D., Chen, J. C., Civano, F., D'Abrusco, R., Doe, S. M., Fabbiano, G., Martinez Galarza, J., Gibbs, D. G., I., Glotfelty, K. J., Graessle, D. E., Grier, J. D., J., Hain, R. M., Hall, D. M., Harbo, P. N., Houck, J. C., Lauer, J. L., Laurino, O., Lee, N. P., McCollough, M. L., McDowell, J. C., McLaughlin, W., Morgan, D. L., Mossman, A. E., Nguyen, D. T., Nichols, J. S., Nowak, M. A., Paxson, C., Perdikeas, M., Plummer, D. A., Rots, A. H., Siemiginowska, A. L., Sundheim, B. A., Thong, S., Tibbetts, M. S., Van Stone, D. W., Winkelman, S. L., & Zografou, P. (2020). The Chandra Source Catalog — A Billion X-ray Photons. In *American Astronomical Society Meeting Abstracts #235*, volume 235 of *American Astronomical Society Meeting Abstracts* (pp. 154.05).

113. Event Horizon Telescope Collaboration, Akiyama, K., Alberdi, A., Alef, W., Algaba, J. C., Anantua, R., Asada, K., Azulay, R., Bach, U., Baczko, A.-K., Ball, D., Baloković, M., Barrett, J., Bauböck, M., Benson, B. A., Bintley, D., Blackburn, L., Blundell, R., Bouman, K. L., Bower, G. C., Boyce, H., Bremer, M., Brinkerink, C. D., Brissenden, R., Britzen, S., Broderick, A. E., Broguiere, D., Bronzwaer, T., Bustamante, S., Byun, D.-Y., Carlstrom, J. E., Ceccobello, C., Chael, A., Chan, C.-k., Chatterjee, K., Chatterjee, S., Chen, M.-T., Chen, Y., Cheng, X., Cho, I., Christian, P., Conroy, N. S., Conway, J. E., Cordes, J. M., Crawford, T. M., Crew, G. B., Cruz-Osorio, A., Cui, Y., Davelaar, J., De Laurentis, M., Deane, R., Dempsey, J., Desvignes, G., Dexter, J., Dhruv, V., Doeleman, S. S., Dougal, S., Dzib, S. A., Eatough, R. P., Emami, R., Falcke, H., Farah, J., Fish, V. L., Fomalont, E., Ford, H. A., Fraga-Encinas, R., Freeman, W. T., Friberg, P., Fromm, C. M., Fuentes, A., Galison, P., Gammie, C. F., García, R., Gentaz, O., Georgiev, B., Goddi, C., Gold, R., Gómez-Ruiz, A. I., Gómez, J. L., Gu, M., Gurwell, M., Hada, K., Haggard, D., Haworth, K., Hecht, M. H., Hesper, R., Heumann, D., Ho, L. C., Ho, P., Honma, M., Huang, C.-W. L., Huang, L., Hughes, D. H., Ikeda, S., Impellizzeri, C. M. V., Inoue, M., Issaoun, S., James, D. J., Jannuzi, B. T., Janssen, M., Jeter, B., Jiang, W., Jiménez-Rosales, A., Johnson, M. D., Jorstad, S., Joshi, A. V., Jung, T., Karami, M., Karuppusamy, R., Kawashima, T., Keating, G. K., Kettenis, M., Kim, D.-J., Kim, J.-Y., Kim, J., Kim, J., Kino, M., Koay, J. Y., Kocherlakota, P., Kofuji, Y., Koch, P. M., Koyama, S., Kramer, C., Kramer, M., Krichbaum, T. P., Kuo, C.-Y., La Bella, N., Lauer, T. R., Lee, D., Lee, S.-S., Leung, P. K., Levis, A., Li, Z., Lico, R., Lindahl, G., Lindqvist, M., Lisakov, M., Liu, J., Liu, K., Liuzzo, E., Lo, W.-P., Lobanov, A. P., Loinard, L., Lonsdale, C. J., Lu, R.-S., Mao, J., Marchili, N., Markoff, S., Marrone, D. P., Marscher, A. P., Martí-Vidal, I., Matsushita, S., Matthews, L. D., Medeiros, L., Menten, K. M., Michalik, D., Mizuno, I., Mizuno, Y., Moran, J. M., Moriyama, K., Moscibrodzka, M., Müller, C., Mus, A., Musoke, G., Myserlis, I., Nadolski, A., Nagai, H., Nagar, N. M., Nakamura, M., Narayan, R., Narayanan, G., Natarajan, I., Nathanail, A., Fuentes, S. N., Neilsen, J., Neri, R., Ni, C., Noutsos, A., Nowak, M. A., Oh, J., Okino, H., Olivares, H., Ortiz-León, G. N., Oyama, T., Özel, F., Palumbo, D. C. M., Paraschos, G. F., Park, J., Parsons, H., Patel, N., Pen, U.-L., Pesce, D. W., Piétu, V., Plambeck, R., PopStefanija, A., Porth, O., Pötzl, F. M., Prather, B., Preciado-López, J. A., Psaltis, D., Pu, H.-Y., Ramakrishnan, V., Rao, R., Rawlings, M. G., Raymond, A. W., Rezzolla, L., Ricarte, A., Ripperda, B., Roelofs,

F., Rogers, A., Ros, E., Romero-Cañizales, C., Roshanineshat, A., Rottmann, H., Roy, A. L., Ruiz, I., Ruszczyk, C., Rygl, K. L. J., Sánchez, S., Sánchez-Argüelles, D., Sánchez-Portal, M., Sasada, M., Satapathy, K., Savolainen, T., Schloerb, F. P., Schonfeld, J., Schuster, K.-F., Shao, L., Shen, Z., Small, D., Sohn, B. W., SooHoo, J., Souccar, K., Sun, H., Tazaki, F., Tetarenko, A. J., Tiede, P., Tilanus, R. P. J., Titus, M., Torne, P., Traianou, E., Trent, T., Trippe, S., Turk, M., van Bemmel, I., van Langevelde, H. J., van Rossum, D. R., Vos, J., Wagner, J., Ward-Thompson, D., Wardle, J., Weintroub, J., Wex, N., Wharton, R., Wielgus, M., Wiik, K., Witzel, G., Wondrak, M. F., Wong, G. N., Wu, Q., Yamaguchi, P., Yoon, D., Young, A., Young, K., Younsi, Z., Yuan, F., Yuan, Y.-F., Zensus, J. A., Zhang, S., Zhao, G.-Y., Zhao, S.-S., Agurto, C., Allardi, A., Amestica, R., Araneda, J. P., Arriagada, O., Berghuis, J. L., Bertarini, A., Berthold, R., Blanchard, J., Brown, K., Cárdenas, M., Cantzler, M., Caro, P., Castillo-Domínguez, E., Chan, T. L., Chang, C.-C., Chang, D. O., Chang, S.-H., Chang, S.-C., Chen, C.-C., Chilson, R., Chuter, T. C., Ciechanowicz, M., Colin-Beltran, E., Coulson, I. M., Crowley, J., Degenaar, N., Dornbusch, S., Durán, C. A., Everett, W. B., Faber, A., Forster, K., Fuchs, M. M., Gale, D. M., Geertsema, G., González, E., Graham, D., Gueth, F., Halverson, N. W., Han, C.-C., Han, K.-C., Hasegawa, Y., Hernández-Rebollar, J. L., Herrera, C., Herrero-Illana, R., Heyminck, S., Hirota, A., Hoge, J., Hostler Schimpf, S. R., Howie, R. E., Huang, Y.-D., Jiang, H., Jinchi, H., John, D., Kimura, K., Klein, T., Kubo, D., Kuroda, J., Kwon, C., Lacasse, R., Laing, R., Leitch, E. M., Li, C.-T., Liu, C.-T., Liu, K.-Y., Lin, L. C. C., Lu, L.-M., Mac-Auliffe, F., Martin-Cocher, P., Matulonis, C., Maute, J. K., Messias, H., Meyer-Zhao, Z., Montaña, A., Montenegro-Montes, F., Montgomerie, W., Moreno Nolasco, M. E., Muders, D., Nishioka, H., Norton, T. J., Nystrom, G., Ogawa, H., Olivares, R., Oshiro, P., Pérez-Beaupuits, J. P., Parra, R., Phillips, N. M., Poirier, M., Pradel, N., Qiu, R., Raffin, P. A., Rahlin, A. S., Ramírez, J., Ressler, S., Reynolds, M., Rodríguez-Montoya, I., Saez-Madain, A. F., Santana, J., Shaw, P., Shirkey, L. E., Silva, K. M., Snow, W., Sousa, D., Sridharan, T. K., Stahm, W., Stark, A. A., Test, J., Torstensson, K., Venegas, P., Walther, C., Wei, T.-S., White, C., Wieching, G., Wijnands, R., Wouterloot, J. G. A., Yu, C.-Y., Yu (于威), W., & Zeballos, M. (2022). First Sagittarius A* Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole in the Center of the Milky Way. ApJ, 930(2), L12.

114. Event Horizon Telescope Collaboration, Akiyama, K., Alberdi, A., Alef, W., Asada, K., Azulay, R., Baczko, A.-K., Ball, D., Baloković, M., Barrett, J., Bintley, D., Blackburn, L., Boland, W., Bouman, K. L., Bower, G. C., Bremer, M., Brinkerink, C. D., Brissenden, R., Britzen, S., Broderick, A. E., Broguiere, D., Bronzwaer, T., Byun, D.-Y., Carlstrom, J. E., Chael, A., Chan, C.-k., Chatterjee, S., Chatterjee, K., Chen, M.-T., Chen, Y., Cho, I., Christian, P., Conway, J. E., Cordes, J. M., Crew, G. B., Cui, Y., Davelaar, J., De Laurentis, M., Deane, R., Dempsey, J., Desvignes, G., Dexter, J., Doeleman, S. S., Eatough, R. P., Falcke, H., Fish, V. L., Fomalont, E., Fraga-Encinas, R., Freeman, W. T., Friberg, P., Fromm, C. M., Gómez, J. L., Galison, P., Gammie, C. F., García, R., Gentaz, O., Georgiev, B., Goddi, C., Gold, R., Gu, M., Gurwell, M., Hada, K., Hecht, M. H., Hesper, R., Ho, L. C., Ho, P., Honma, M., Huang, C.-W. L., Huang, L., Hughes, D. H., Ikeda, S., Inoue, M., Issaoun, S., James, D. J., Jannuzi, B. T., Janssen, M., Jeter, B., Jiang, W., Johnson, M. D., Jorstad, S., Jung, T., Karami, M., Karuppusamy, R., Kawashima, T., Keating, G. K., Kettenis, M., Kim, J.-Y., Kim, J., Kim, J., Kino, M., Koay, J. Y., Koch, P. M., Koyama, S., Kramer, M., Kramer, C., Krichbaum, T. P., Kuo, C.-Y., Lauer, T. R., Lee, S.-S., Li, Y.-R., Li, Z., Lindqvist, M., Liu, K., Liuzzo, E., Lo, W.-P., Lobanov, A. P., Loinard, L., Lonsdale, C., Lu, R.-S., MacDonald, N. R., Mao, J., Markoff, S., Marrone, D. P., Marscher, A. P., Martí-Vidal, I., Matsushita, S., Matthews, L. D., Medeiros, L., Menten, K. M., Mizuno, Y., Mizuno, I., Moran, J. M., Moriyama, K., Moscibrodzka, M., Müller, C., Nagai, H., Nagar, N. M., Nakamura, M., Narayan, R., Narayanan, G., Natarajan, I., Neri, R., Ni, C., Noutsos, A., Okino, H., Olivares, H., Ortiz-León, G. N., Oyama, T., Özel, F., Palumbo, D. C. M., Patel, N., Pen, U.-L., Pesce, D. W., Piétu, V., Plambeck, R., PopStefanija, A., Porth, O., Prather, B., Preciado-López, J. A., Psaltis, D., Pu, H.-Y., Ramakrishnan, V., Rao, R., Rawlings, M. G., Raymond, A. W., Rezzolla, L., Ripperda, B., Roelofs, F., Rogers, A., Ros, E., Rose, M., Roshanineshat, A., Rottmann, H., Roy, A. L., Ruszczyk, C., Ryan, B. R., Rygl, K. L. J., Sánchez, S., Sánchez-Arguelles, D., Sasada, M., Savolainen, T., Schloerb, F. P., Schuster, K.-F., Shao, L., Shen, Z., Small, D., Sohn, B. W., SooHoo, J., Tazaki, F., Tiede, P., Tilanus, R. P. J., Titus, M., Toma, K., Torne, P., Trent, T., Trippe, S., Tsuda, S., van Bemmel, I., van Langevelde, H. J., van Rossum, D. R., Wagner, J., Wardle, J., Weintroub, J., Wex, N., Wharton, R., Wielgus, M., Wong, G. N., Wu, Q., Young, K., Young, A., Younsi, Z., Yuan, F., Yuan, Y.-F., Zensus, J. A., Zhao, G., Zhao, S.-S., Zhu, Z., Algaba, J.-C., Allardi, A., Amestica, R., Anczarski, J., Bach, U., Baganoff, F. K., Beaudoin, C., Benson, B. A., Berthold, R., Blanchard, J. M., Blundell, R., Bustamente, S., Cappallo, R., Castillo-Domínguez, E., Chang, C.-C., Chang, S.-H., Chang, S.-C., Chen, C.-C., Chilson, R., Chuter, T. C., Córdova Rosado, R., Coulson, I. M., Crawford, T. M., Crowley, J., David, J., Derome, M., Dexter, M., Dornbusch, S., Dudevoir, K. A., Dzib, S. A., Eckart, A., Eckert, C., Erickson, N. R., Everett, W. B., Faber, A., Farah, J. R., Fath, V., Folkers, T. W., Forbes, D. C., Freund, R., Gómez-Ruiz, A. I., Gale, D. M., Gao, F., Geertsema, G., Graham, D. A., Greer, C. H., Grosslein, R., Gueth, F., Haggard, D., Halverson, N. W., Han, C.-C., Han, K.-C., Hao, J., Hasegawa, Y., Henning, J. W., Hernández-Gómez, A., Herrero-Illana, R., Heyminck, S., Hirota, A., Hoge, J., Huang, Y.-D., Impellizzeri, C. M. V., Jiang, H., Kamble, A., Keisler, R., Kimura, K., Kono, Y., Kubo, D., Kuroda, J., Lacasse, R., Laing, R. A., Leitch, E. M., Li, C.-T., Lin, L. C. C., Liu, C.-T., Liu, K.-Y., Lu, L.-M., Marson, R. G., Martin-Cocher, P. L., Massingill, K. D., Matulonis, C., McColl, M. P., McWhirter, S. R., Messias, H., Meyer-Zhao, Z., Michalik, D., Montaña, A., Montgomerie, W., Mora-Klein, M., Muders, D., Nadolski, A., Navarro, S., Neilsen, J., Nguyen, C. H., Nishioka, H., Norton, T., Nowak, M. A., Nystrom, G., Ogawa, H., Oshiro, P., Oyama, T., Parsons, H., Paine, S. N., Peñalver, J., Phillips, N. M., Poirier, M., Pradel, N., Primiani, R. A.,

Raffin, P. A., Rahlin, A. S., Reiland, G., Risacher, C., Ruiz, I., Sáez-Madaín, A. F., Sassella, R., Schellart, P., Shaw, P., Silva, K. M., Shiokawa, H., Smith, D. R., Snow, W., Souccar, K., Sousa, D., Sridharan, T. K., Srinivasan, R., Stahm, W., Stark, A. A., Story, K., Timmer, S. T., Vertatschitsch, L., Walther, C., Wei, T.-S., Whitehorn, N., Whitney, A. R., Woody, D. P., Wouterloot, J. G. A., Wright, M., Yamaguchi, P., Yu, C.-Y., Zeballos, M., Zhang, S., & Ziurys, L. (2019a). First M87 Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole. ApJ, 875(1), L1.

115. Event Horizon Telescope Collaboration, Akiyama, K., Alberdi, A., Alef, W., Asada, K., Azulay, R., Baczko, A.-K., Ball, D., Baloković, M., Barrett, J., Bintley, D., Blackburn, L., Boland, W., Bouman, K. L., Bower, G. C., Bremer, M., Brinkerink, C. D., Brissenden, R., Britzen, S., Broderick, A. E., Broguiere, D., Bronzwaer, T., Byun, D.-Y., Carlstrom, J. E., Chael, A., Chan, C.-k., Chatterjee, S., Chatterjee, K., Chen, M.-T., Chen, Y., Cho, I., Christian, P., Conway, J. E., Cordes, J. M., Crew, G. B., Cui, Y., Davelaar, J., De Laurentis, M., Deane, R., Dempsey, J., Desvignes, G., Dexter, J., Doeleman, S. S., Eatough, R. P., Falcke, H., Fish, V. L., Fomalont, E., Fraga-Encinas, R., Friberg, P., Fromm, C. M., Gómez, J. L., Galison, P., Gammie, C. F., García, R., Gentaz, O., Georgiev, B., Goddi, C., Gold, R., Gu, M., Gurwell, M., Hada, K., Hecht, M. H., Hesper, R., Ho, L. C., Ho, P., Honma, M., Huang, C.-W. L., Huang, L., Hughes, D. H., Ikeda, S., Inoue, M., Issaoun, S., James, D. J., Jannuzi, B. T., Janssen, M., Jeter, B., Jiang, W., Johnson, M. D., Jorstad, S., Jung, T., Karami, M., Karuppusamy, R., Kawashima, T., Keating, G. K., Kettenis, M., Kim, J.-Y., Kim, J., Kim, J., Kino, M., Koay, J. Y., Koch, P. M., Koyama, S., Kramer, M., Kramer, C., Krichbaum, T. P., Kuo, C.-Y., Lauer, T. R., Lee, S.-S., Li, Y.-R., Li, Z., Lindqvist, M., Liu, K., Liuzzo, E., Lo, W.-P., Lobanov, A. P., Loinard, L., Lonsdale, C., Lu, R.-S., MacDonald, N. R., Mao, J., Markoff, S., Marrone, D. P., Marscher, A. P., Martí-Vidal, I., Matsushita, S., Matthews, L. D., Medeiros, L., Menten, K. M., Mizuno, Y., Mizuno, I., Moran, J. M., Moriyama, K., Moscibrodzka, M., Muller, C., Nagai, H., Nagar, N. M., Nakamura, M., Narayan, R., Narayanan, G., Natarajan, I., Neri, R., Ni, C., Noutsos, A., Okino, H., Olivares, H., Oyama, T., Özel, F., Palumbo, D. C. M., Patel, N., Pen, U.-L., Pesce, D. W., Piétu, V., Plambeck, R., PopStefanija, A., Porth, O., Prather, B., Preciado-López, J. A., Psaltis, D., Pu, H.-Y., Ramakrishnan, V., Rao, R., Rawlings, M. G., Raymond, A. W., Rezzolla, L., Ripperda, B., Roelofs, F., Rogers, A., Ros, E., Rose, M., Roshanineshat, A., Rottmann, H., Roy, A. L., Ruszczyk, C., Ryan, B. R., Rygl, K. L. J., Sánchez, S., Sánchez-Arguelles, D., Sasada, M., Savolainen, T., Schloerb, F. P., Schuster, K.-F., Shao, L., Shen, Z., Small, D., Sohn, B. W., SooHoo, J., Tazaki, F., Tiede, P., Tilanus, R. P. J., Titus, M., Toma, K., Torne, P., Trent, T., Trippe, S., Tsuda, S., van Bemmel, I., van Langevelde, H. J., van Rossum, D. R., Wagner, J., Wardle, J., Weintroub, J., Wex, N., Wharton, R., Wielgus, M., Wong, G. N., Wu, Q., Young, A., Young, K., Younsi, Z., Yuan, F., Yuan, Y.-F., Zensus, J. A., Zhao, G., Zhao, S.-S., Zhu, Z., Anczarski, J., Baganoff, F. K., Eckart, A., Farah, J. R., Haggard, D., Meyer-Zhao, Z., Michalik, D., Nadolski, A., Neilsen, J., Nishioka, H., Nowak, M. A., Pradel, N., Primiani, R. A., Souccar, K., Vertatschitsch, L., Yamaguchi, P., & Zhang, S. (2019b). First M87 Event Horizon Telescope Results. V. Physical Origin of the Asymmetric Ring. ApJ, 875(1), L5.

116. Event Horizon Telescope Collaboration, Akiyama, K., Algaba, J. C., Alberdi, A., Alef, W., Anantua, R., Asada, K., Azulay, R., Baczko, A.-K., Ball, D., Baloković, M., Barrett, J., Benson, B. A., Bintley, D., Blackburn, L., Blundell, R., Boland, W., Bouman, K. L., Bower, G. C., Boyce, H., Bremer, M., Brinkerink, C. D., Brissenden, R., Britzen, S., Broderick, A. E., Broguiere, D., Bronzwaer, T., Byun, D.-Y., Carlstrom, J. E., Chael, A., Chan, C.-k., Chatterjee, S., Chatterjee, K., Chen, M.-T., Chen, Y., Chesler, P. M., Cho, I., Christian, P., Conway, J. E., Cordes, J. M., Crawford, T. M., Crew, G. B., Cruz-Osorio, A., Cui, Y., Davelaar, J., De Laurentis, M., Deane, R., Dempsey, J., Desvignes, G., Dexter, J., Doeleman, S. S., Eatough, R. P., Falcke, H., Farah, J., Fish, V. L., Fomalont, E., Ford, H. A., Fraga-Encinas, R., Freeman, W. T., Friberg, P., Fromm, C. M., Fuentes, A., Galison, P., Gammie, C. F., García, R., Gentaz, O., Georgiev, B., Goddi, C., Gold, R., Gómez, J. L., Gómez-Ruiz, A. I., Gu, M., Gurwell, M., Hada, K., Haggard, D., Hecht, M. H., Hesper, R., Ho, L. C., Ho, P., Honma, M., Huang, C.-W. L., Huang, L., Hughes, D. H., Ikeda, S., Inoue, M., Issaoun, S., James, D. J., Jannuzi, B. T., Janssen, M., Jeter, B., Jiang, W., Jimenez-Rosales, A., Johnson, M. D., Jorstad, S., Jung, T., Karami, M., Karuppusamy, R., Kawashima, T., Keating, G. K., Kettenis, M., Kim, D.-J., Kim, J.-Y., Kim, J., Kim, J., Kino, M., Koay, J. Y., Kofuji, Y., Koch, P. M., Koyama, S., Kramer, M., Kramer, C., Krichbaum, T. P., Kuo, C.-Y., Lauer, T. R., Lee, S.-S., Levis, A., Li, Y.-R., Li, Z., Lindqvist, M., Lico, R., Lindahl, G., Liu, J., Liu, K., Liuzzo, E., Lo, W.-P., Lobanov, A. P., Loinard, L., Lonsdale, C., Lu, R.-S., MacDonald, N. R., Mao, J., Marchili, N., Markoff, S., Marrone, D. P., Marscher, A. P., Martí-Vidal, I., Matsushita, S., Matthews, L. D., Medeiros, L., Menten, K. M., Mizuno, I., Mizuno, Y., Moran, J. M., Moriyama, K., Moscibrodzka, M., Müller, C., Musoke, G., Mejías, A. M., Michalik, D., Nadolski, A., Nagai, H., Nagar, N. M., Nakamura, M., Narayan, R., Narayanan, G., Natarajan, I., Nathanail, A., Neilsen, J., Neri, R., Ni, C., Noutsos, A., Nowak, M. A., Okino, H., Olivares, H., Ortiz-León, G. N., Oyama, T., Özel, F., Palumbo, D. C. M., Park, J., Patel, N., Pen, U.-L., Pesce, D. W., Piétu, V., Plambeck, R., PopStefanija, A., Porth, O., Pötzl, F. M., Prather, B., Preciado-López, J. A., Psaltis, D., Pu, H.-Y., Ramakrishnan, V., Rao, R., Rawlings, M. G., Raymond, A. W., Rezzolla, L., Ricarte, A., Ripperda, B., Roelofs, F., Rogers, A., Ros, E., Rose, M., Roshanineshat, A., Rottmann, H., Roy, A. L., Ruszczyk, C., Rygl, K. L. J., Sánchez, S., Sánchez-Arguelles, D., Sasada, M., Savolainen, T., Schloerb, F. P., Schuster, K.-F., Shao, L., Shen, Z., Small, D., Sohn, B. W., SooHoo, J., Sun, H., Tazaki, F., Tetarenko, A. J., Tiede, P., Tilanus, R. P. J., Titus, M., Toma, K., Torne, P., Trent, T., Traianou, E., Trippe, S., van Bemmel, I., van Langevelde, H. J., van Rossum, D. R., Wagner, J., Ward-Thompson, D., Wardle, J., Weintroub, J., Wex, N., Wharton, R., Wielgus, M., Wong, G. N., Wu, Q., Yoon, D., Young, A., Young, K.,

Younsi, Z., Yuan, F., Yuan, Y.-F., Zensus, J. A., Zhao, G.-Y., & Zhao, S.-S. (2021a). First M87 Event Horizon Telescope Results. VII. Polarization of the Ring. ApJ, 910(1), L12.

117. Event Horizon Telescope Collaboration, Akiyama, K., Algaba, J. C., Alberdi, A., Alef, W., Anantua, R., Asada, K., Azulay, R., Baczko, A.-K., Ball, D., Baloković, M., Barrett, J., Benson, B. A., Bintley, D., Blackburn, L., Blundell, R., Boland, W., Bouman, K. L., Bower, G. C., Boyce, H., Bremer, M., Brinkerink, C. D., Brissenden, R., Britzen, S., Broderick, A. E., Broguiere, D., Bronzwaer, T., Byun, D.-Y., Carlstrom, J. E., Chael, A., Chan, C.-k., Chatterjee, S., Chatterjee, K., Chen, M.-T., Chen, Y., Chesler, P. M., Cho, I., Christian, P., Conway, J. E., Cordes, J. M., Crawford, T. M., Crew, G. B., Cruz-Osorio, A., Cui, Y., Davelaar, J., De Laurentis, M., Deane, R., Dempsey, J., Desvignes, G., Dexter, J., Doeleman, S. S., Eatough, R. P., Falcke, H., Farah, J., Fish, V. L., Fomalont, E., Ford, H. A., Fraga-Encinas, R., Friberg, P., Fromm, C. M., Fuentes, A., Galison, P., Gammie, C. F., García, R., Gelles, Z., Gentaz, O., Georgiev, B., Goddi, C., Gold, R., Gómez, J. L., Gómez-Ruiz, A. I., Gu, M., Gurwell, M., Hada, K., Haggard, D., Hecht, M. H., Hesper, R., Himwich, E., Ho, L. C., Ho, P., Honma, M., Huang, C.-W. L., Huang, L., Hughes, D. H., Ikeda, S., Inoue, M., Issaoun, S., James, D. J., Jannuzi, B. T., Janssen, M., Jeter, B., Jiang, W., Jimenez-Rosales, A., Johnson, M. D., Jorstad, S., Jung, T., Karami, M., Karuppusamy, R., Kawashima, T., Keating, G. K., Kettenis, M., Kim, D.-J., Kim, J.-Y., Kim, J., Kim, J., Kino, M., Koay, J. Y., Kofuji, Y., Koch, P. M., Koyama, S., Kramer, M., Kramer, C., Krichbaum, T. P., Kuo, C.-Y., Lauer, T. R., Lee, S.-S., Levis, A., Li, Y.-R., Li, Z., Lindqvist, M., Lico, R., Lindahl, G., Liu, J., Liu, K., Liuzzo, E., Lo, W.-P., Lobanov, A. P., Loinard, L., Lonsdale, C., Lu, R.-S., MacDonald, N. R., Mao, J., Marchili, N., Markoff, S., Marrone, D. P., Marscher, A. P., Martí-Vidal, I., Matsushita, S., Matthews, L. D., Medeiros, L., Menten, K. M., Mizuno, I., Mizuno, Y., Moran, J. M., Moriyama, K., Moscibrodzka, M., Müller, C., Musoke, G., Mus Mejías, A., Michalik, D., Nadolski, A., Nagai, H., Nagar, N. M., Nakamura, M., Narayan, R., Narayanan, G., Natarajan, I., Nathanail, A., Neilsen, J., Neri, R., Ni, C., Noutsos, A., Nowak, M. A., Okino, H., Olivares, H., Ortiz-León, G. N., Oyama, T., Özel, F., Palumbo, D. C. M., Park, J., Patel, N., Pen, U.-L., Pesce, D. W., Piétu, V., Plambeck, R., PopStefanija, A., Porth, O., Pötzl, F. M., Prather, B., Preciado-López, J. A., Psaltis, D., Pu, H.-Y., Ramakrishnan, V., Rao, R., Rawlings, M. G., Raymond, A. W., Rezzolla, L., Ricarte, A., Ripperda, B., Roelofs, F., Rogers, A., Ros, E., Rose, M., Roshanineshat, A., Rottmann, H., Roy, A. L., Ruszczyk, C., Rygl, K. L. J., Sánchez, S., Sánchez-Arguelles, D., Sasada, M., Savolainen, T., Schloerb, F. P., Schuster, K.-F., Shao, L., Shen, Z., Small, D., Sohn, B. W., SooHoo, J., Sun, H., Tazaki, F., Tetarenko, A. J., Tiede, P., Tilanus, R. P. J., Titus, M., Toma, K., Torne, P., Trent, T., Traianou, E., Trippe, S., van Bemmel, I., van Langevelde, H. J., van Rossum, D. R., Wagner, J., Ward-Thompson, D., Wardle, J., Weintroub, J., Wex, N., Wharton, R., Wielgus, M., Wong, G. N., Wu, Q., Yoon, D., Young, A., Young, K., Younsi, Z., Yuan, F., Yuan, Y.-F., Zensus, J. A., Zhao, G.-Y., & Zhao, S.-S. (2021b). First M87 Event Horizon Telescope Results. VIII. Magnetic Field Structure near The Event Horizon. ApJ, 910(1), L13.

118. Fabian, A. C. (2012). Observational Evidence of Active Galactic Nuclei Feedback. ARA&A, 50, 455–489.

119. Fabian, A. C., Nulsen, P. E. J., & Canizares, C. R. (1984). Cooling flows in clusters of galaxies. Nature, 310(5980), 733–740.

120. Falco, E. E., Kurtz, M. J., Geller, M. J., Huchra, J. P., Peters, J., Berlind, P., Mink, D. J., Tokarz, S. P., & Elwell, B. (1999). The Updated Zwicky Catalog (UZC). PASP, 111(758), 438–452.

121. Feretti, L., Boehringer, H., Giovannini, G., & Neumann, D. (1997). The radio and X-ray properties of Abell 2255. A&A, 317, 432–440.

122. Ford, A. L., Keenan, B. D., & Medvedev, M. V. (2018). Electron-positron cascade in magnetospheres of spinning black holes. *Physical review. D*, 98(6), 063016.

123. Forero-Romero, J. E., Hoffman, Y., Gottlöber, S., Klypin, A., & Yepes, G. (2009). A dynamical classification of the cosmic web. MNRAS, 396(3), 1815–1824.

124. Frank, P., Jasche, J., & Enßlin, T. A. (2016). SOMBI: Bayesian identification of parameter relations in unstructured cosmological data. A&A, 595, A75.

125. Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, (pp. 1189–1232).

126. Gaia Collaboration, Brown, A. G. A., Vallenari, A., Prusti, T., de Bruijne, J. H. J., Babusiaux, C., Biermann, M., Creevey, O. L., Evans, D. W., Eyer, L., Hutton, A., Jansen, F., Jordi, C., Klioner, S. A., Lammers, U., Lindegren, L., Luri, X., Mignard, F., Panem, C., Pourbaix, D., Randich, S., Sartoretti, P., Soubiran, C., Walton, N. A., Arenou, F., Bailer-Jones, C. A. L., Bastian, U., Cropper, M., Drimmel, R., Katz, D., Lattanzi, M. G., van Leeuwen, F., Bakker, J., Cacciari, C., Castañeda, J., De Angeli, F., Ducourant, C., Fabricius, C., Fouesneau, M., Frémat, Y., Guerra, R., Guerrier, A., Guiraud, J., Jean-Antoine Piccolo, A., Masana, E., Messineo, R., Mowlavi, N., Nicolas, C., Nienartowicz, K., Pailler, F., Panuzzo, P., Riclet, F., Roux, W., Seabroke, G. M., Sordo, R., Tanga, P., Thévenin, F., Gracia-Abril, G., Portell, J., Teyssier, D., Altmann, M., Andrae, R., Bellas-Velidis, I., Benson, K., Berthier, J., Blomme, R., Brugaletta, E., Burgess, P. W., Busso, G., Carry, B., Cellino, A., Cheek, N., Clementini, G., Damerdji, Y., Davidson, M., Delchambre, L., Dell'Oro, A., Fernández-Hernández, J., Galluccio, L., García-Lario, P., Garcia-Reinaldos, M., González-Núñez, J., Gosset, E., Haigron, R., Halbwachs, J. L., Hambly, N. C., Harrison, D. L., Hatzidimitriou, D., Heiter, U., Hernández, J., Hestroffer, D., Hodgkin, S. T., Holl, B.,

Janßen, K., Jevardat de Fombelle, G., Jordan, S., Krone-Martins, A., Lanzafame, A. C., Löffler, W., Lorca, A., Manteiga, M., Marchal, O., Marrese, P. M., Moitinho, A., Mora, A., Muinonen, K., Osborne, P., Pancino, E., Pauwels, T., Petit, J. M., Recio-Blanco, A., Richards, P. J., Riello, M., Rimoldini, L., Robin, A. C., Roegiers, T., Rybizki, J., Sarro, L. M., Siopis, C., Smith, M., Sozzetti, A., Ulla, A., Utrilla, E., van Leeuwen, M., van Reeven, W., Abbas, U., Abreu Aramburu, A., Accart, S., Aerts, C., Aguado, J. J., Ajaj, M., Altavilla, G., Álvarez, M. A., Álvarez Cid-Fuentes, J., Alves, J., Anderson, R. I., Anglada Varela, E., Antoja, T., Audard, M., Baines, D., Baker, S. G., Balaguer-Núñez, L., Balbinot, E., Balog, Z., Barache, C., Barbato, D., Barros, M., Barstow, M. A., Bartolomé, S., Bassilana, J. L., Bauchet, N., Baudesson-Stella, A., Becciani, U., Bellazzini, M., Bernet, M., Bertone, S., Bianchi, L., Blanco-Cuaresma, S., Boch, T., Bombrun, A., Bossini, D., Bouquillon, S., Bragaglia, A., Bramante, L., Breedt, E., Bressan, A., Brouillet, N., Bucciarelli, B., Burlacu, A., Busonero, D., Butkevich, A. G., Buzzi, R., Caffau, E., Cancelliere, R., Cánovas, H., Cantat-Gaudin, T., Carballo, R., Carlucci, T., Carnerero, M. I., Carrasco, J. M., Casamiquela, L., Castellani, M., Castro-Ginard, A., Castro Sampol, P., Chaoul, L., Charlot, P., Chemin, L., Chiavassa, A., Cioni, M. R. L., Comoretto, G., Cooper, W. J., Cornez, T., Cowell, S., Crifo, F., Crosta, M., Crowley, C., Dafonte, C., Dapergolas, A., David, M., David, P., de Laverny, P., De Luise, F., De March, R., De Ridder, J., de Souza, R., de Teodoro, P., de Torres, A., del Peloso, E. F., del Pozo, E., Delbo, M., Delgado, A., Delgado, H. E., Delisle, J. B., Di Matteo, P., Diakite, S., Diener, C., Distefano, E., Dolding, C., Eappachen, D., Edvardsson, B., Enke, H., Esquej, P., Fabre, C., Fabrizio, M., Faigler, S., Fedorets, G., Fernique, P., Fienga, A., Figueras, F., Fouron, C., Fragkoudi, F., Fraile, E., Franke, F., Gai, M., Garabato, D., Garcia-Gutierrez, A., García-Torres, M., Garofalo, A., Gavras, P., Gerlach, E., Geyer, R., Giacobbe, P., Gilmore, G., Girona, S., Giuffrida, G., Gomel, R., Gomez, A., Gonzalez-Santamaria, I., González-Vidal, J. J., Granvik, M., Gutiérrez-Sánchez, R., Guy, L. P., Hauser, M., Haywood, M., Helmi, A., Hidalgo, S. L., Hilger, T., Hładczuk, N., Hobbs, D., Holland, G., Huckle, H. E., Jasniewicz, G., Jonker, P. G., Juaristi Campillo, J., Julbe, F., Karbevska, L., Kervella, P., Khanna, S., Kochoska, A., Kontizas, M., Kordopatis, G., Korn, A. J., Kostrzewa-Rutkowska, Z., Kruszyńska, K., Lambert, S., Lanza, A. F., Lasne, Y., Le Campion, J. F., Le Fustec, Y., Lebreton, Y., Lebzelter, T., Leccia, S., Leclerc, N., Lecoeur-Taibi, I., Liao, S., Licata, E., Lindstrøm, E. P., Lister, T. A., Livanou, E., Lobel, A., Madrero Pardo, P., Managau, S., Mann, R. G., Marchant, J. M., Marconi, M., Marcos Santos, M. M. S., Marinoni, S., Marocco, F., Marshall, D. J., Martin Polo, L., Martín-Fleitas, J. M., Masip, A., Massari, D., Mastrobuono-Battisti, A., Mazeh, T., McMillan, P. J., Messina, S., Michalik, D., Millar, N. R., Mints, A., Molina, D., Molinaro, R., Molnár, L., Montegriffo, P., Mor, R., Morbidelli, R., Morel, T., Morris, D., Mulone, A. F., Munoz, D., Muraveva, T., Murphy, C. P., Musella, I., Noval, L., Ordénovic, C., Orrù, G., Osinde, J., Pagani, C., Pagano, I., Palaversa, L., Palicio, P. A., Panahi, A., Pawlak, M., Peñalosa Esteller, X., Penttilä, A., Piersimoni, A. M., Pineau, F. X., Plachy, E., Plum, G., Poggio, E., Poretti, E., Poujoulet, E., Prša, A., Pulone, L., Racero, E., Ragaini, S., Rainer, M., Raiteri, C. M., Rambaux, N., Ramos, P., Ramos-Lerate, M., Re Fiorentin, P., Regibo, S., Reylé, C., Ripepi, V., Riva, A., Rixon, G., Robichon, N., Robin, C., Roelens, M., Rohrbasser, L., Romero-Gómez, M., Rowell, N., Royer, F., Rybicki, K. A., Sadowski, G., Sagristà Sellés, A., Sahlmann, J., Salgado, J., Salguero, E., Samaras, N., Sanchez Gimenez, V., Sanna, N., Santoveña, R., Sarasso, M., Schultheis, M., Sciacca, E., Segol, M., Segovia, J. C., Ségransan, D., Semeux, D., Shahaf, S., Siddiqui, H. I., Siebert, A., Siltala, L., Slezak, E., Smart, R. L., Solano, E., Solitro, F., Souami, D., Souchay, J., Spagna, A., Spoto, F., Steele, I. A., Steidelmüller, H., Stephenson, C. A., Süveges, M., Szabados, L., Szegedi-Elek, E., Taris, F., Tauran, G., Taylor, M. B., Teixeira, R., Thuillot, W., Tonello, N., Torra, F., Torra, J., Turon, C., Unger, N., Vaillant, M., van Dillen, E., Vanel, O., Vecchiato, A., Viala, Y., Vicente, D., Voutsinas, S., Weiler, M., Wevers, T., Wyrzykowski, Ł., Yoldas, A., Yvard, P., Zhao, H., Zorec, J., Zucker, S., Zurbach, C., & Zwitter, T. (2021). Gaia Early Data Release 3. Summary of the contents and survey properties. A&A, 649, A1.

127. Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., Brown, A. G. A., Vallenari, A., Babusiaux, C., Bailer-Jones, C. A. L., Bastian, U., Biermann, M., Evans, D. W., Eyer, L., Jansen, F., Jordi, C., Klioner, S. A., Lammers, U., Lindegren, L., Luri, X., Mignard, F., Milligan, D. J., Panem, C., Poinsignon, V., Pourbaix, D., Randich, S., Sarri, G., Sartoretti, P., Siddiqui, H. I., Soubiran, C., Valette, V., van Leeuwen, F., Walton, N. A., Aerts, C., Arenou, F., Cropper, M., Drimmel, R., Høg, E., Katz, D., Lattanzi, M. G., O'Mullane, W., Grebel, E. K., Holland, A. D., Huc, C., Passot, X., Bramante, L., Cacciari, C., Castañeda, J., Chaoul, L., Cheek, N., De Angeli, F., Fabricius, C., Guerra, R., Hernández, J., Jean-Antoine-Piccolo, A., Masana, E., Messineo, R., Mowlavi, N., Nienartowicz, K., Ordóñez-Blanco, D., Panuzzo, P., Portell, J., Richards, P. J., Riello, M., Seabroke, G. M., Tanga, P., Thévenin, F., Torra, J., Els, S. G., Gracia-Abril, G., Comoretto, G., Garcia-Reinaldos, M., Lock, T., Mercier, E., Altmann, M., Andrae, R., Astraatmadja, T. L., Bellas-Velidis, I., Benson, K., Berthier, J., Blomme, R., Busso, G., Carry, B., Cellino, A., Clementini, G., Cowell, S., Creevey, O., Cuypers, J., Davidson, M., De Ridder, J., de Torres, A., Delchambre, L., Dell'Oro, A., Ducourant, C., Frémat, Y., García-Torres, M., Gosset, E., Halbwachs, J. L., Hambly, N. C., Harrison, D. L., Hauser, M., Hestroffer, D., Hodgkin, S. T., Huckle, H. E., Hutton, A., Jasniewicz, G., Jordan, S., Kontizas, M., Korn, A. J., Lanzafame, A. C., Manteiga, M., Moitinho, A., Muinonen, K., Osinde, J., Pancino, E., Pauwels, T., Petit, J. M., Recio-Blanco, A., Robin, A. C., Sarro, L. M., Siopis, C., Smith, M., Smith, K. W., Sozzetti, A., Thuillot, W., van Reeven, W., Viala, Y., Abbas, U., Abreu Aramburu, A., Accart, S., Aguado, J. J., Allan, P. M., Allasia, W., Altavilla, G., Álvarez, M. A., Alves, J., Anderson, R. I., Andrei, A. H., Anglada Varela, E., Antiche, E., Antoja, T., Antón, S., Arcay, B., Atzei, A., Ayache, L., Bach, N., Baker, S. G., Balaguer-Núñez, L., Barache, C., Barata, C., Barbier, A., Barblan, F., Baroni, M., Barrado y Navascués, D., Barros, M., Barstow, M. A., Becciani, U., Bellazzini, M., Bellei, G., Bello García, A.,

Belokurov, V., Bendjoya, P., Berihuete, A., Bianchi, L., Bienaymé, O., Billebaud, F., Blagorodnova, N., Blanco-Cuaresma, S., Boch, T., Bombrun, A., Borrachero, R., Bouquillon, S., Bourda, G., Bouy, H., Bragaglia, A., Breddels, M. A., Brouillet, N., Brüsemeister, T., Bucciarelli, B., Budnik, F., Burgess, P., Burgon, R., Burlacu, A., Busonero, D., Buzzi, R., Caffau, E., Cambras, J., Campbell, H., Cancelliere, R., Cantat-Gaudin, T., Carlucci, T., Carrasco, J. M., Castellani, M., Charlot, P., Charnas, J., Charvet, P., Chassat, F., Chiavassa, A., Clotet, M., Cocozza, G., Collins, R. S., Collins, P., Costigan, G., Crifo, F., Cross, N. J. G., Crosta, M., Crowley, C., Dafonte, C., Damerdji, Y., Dapergolas, A., David, P., David, M., De Cat, P., de Felice, F., de Laverny, P., De Luise, F., De March, R., de Martino, D., de Souza, R., Debosscher, J., del Pozo, E., Delbo, M., Delgado, A., Delgado, H. E., di Marco, F., Di Matteo, P., Diakite, S., Distefano, E., Dolding, C., Dos Anjos, S., Drazinos, P., Durán, J., Dzigan, Y., Ecale, E., Edvardsson, B., Enke, H., Erdmann, M., Escolar, D., Espina, M., Evans, N. W., Eynard Bontemps, G., Fabre, C., Fabrizio, M., Faigler, S., Falcão, A. J., Farràs Casas, M., Faye, F., Federici, L., Fedorets, G., Fernández-Hernández, J., Fernique, P., Fienga, A., Figueras, F., Filippi, F., Findeisen, K., Fonti, A., Fouesneau, M., Fraile, E., Fraser, M., Fuchs, J., Furnell, R., Gai, M., Galleti, S., Galluccio, L., Garabato, D., García-Sedano, F., Garé, P., Garofalo, A., Garralda, N., Gavras, P., Gerssen, J., Geyer, R., Gilmore, G., Girona, S., Giuffrida, G., Gomes, M., González-Marcos, A., González-Núñez, J., González-Vidal, J. J., Granvik, M., Guerrier, A., Guillout, P., Guiraud, J., Gúrpide, A., Gutiérrez-Sánchez, R., Guy, L. P., Haigron, R., Hatzidimitriou, D., Haywood, M., Heiter, U., Helmi, A., Hobbs, D., Hofmann, W., Holl, B., Holland, G., Hunt, J. A. S., Hypki, A., Icardi, V., Irwin, M., Jevardat de Fombelle, G., Jofré, P., Jonker, P. G., Jorissen, A., Julbe, F., Karampelas, A., Kochoska, A., Kohley, R., Kolenberg, K., Kontizas, E., Koposov, S. E., Kordopatis, G., Koubsky, P., Kowalczyk, A., Krone-Martins, A., Kudryashova, M., Kull, I., Bachchan, R. K., Lacoste-Seris, F., Lanza, A. F., Lavigne, J. B., Le Poncin-Lafitte, C., Lebreton, Y., Lebzelter, T., Leccia, S., Leclerc, N., Lecoeur-Taibi, I., Lemaitre, V., Lenhardt, H., Leroux, F., Liao, S., Licata, E., Lindstrøm, H. E. P., Lister, T. A., Livanou, E., Lobel, A., Löffler, W., López, M., Lopez-Lozano, A., Lorenz, D., Loureiro, T., MacDonald, I., Magalhães Fernandes, T., Managau, S., Mann, R. G., Mantelet, G., Marchal, O., Marchant, J. M., Marconi, M., Marie, J., Marinoni, S., Marrese, P. M., Marschalkó, G., Marshall, D. J., Martín-Fleitas, J. M., Martino, M., Mary, N., Matijevič, G., Mazeh, T., McMillan, P. J., Messina, S., Mestre, A., Michalik, D., Millar, N. R., Miranda, B. M. H., Molina, D., Molinaro, R., Molinaro, M., Molnár, L., Moniez, M., Montegriffo, P., Monteiro, D., Mor, R., Mora, A., Morbidelli, R., Morel, T., Morgenthaler, S., Morley, T., Morris, D., Mulone, A. F., Muraveva, T., Musella, I., Narbonne, J., Nelemans, G., Nicastro, L., Noval, L., Ordénovic, C., Ordieres-Meré, J., Osborne, P., Pagani, C., Pagano, I., Pailler, F., Palacin, H., Palaversa, L., Parsons, P., Paulsen, T., Pecoraro, M., Pedrosa, R., Pentikäinen, H., Pereira, J., Pichon, B., Piersimoni, A. M., Pineau, F. X., Plachy, E., Plum, G., Poujoulet, E., Prša, A., Pulone, L., Ragaini, S., Rago, S., Rambaux, N., Ramos-Lerate, M., Ranalli, P., Rauw, G., Read, A., Regibo, S., Renk, F., Reylé, C., Ribeiro, R. A., Rimoldini, L., Ripepi, V., Riva, A., Rixon, G., Roelens, M., Romero-Gómez, M., Rowell, N., Royer, F., Rudolph, A., Ruiz-Dern, L., Sadowski, G., Sagristà Sellés, T., Sahlmann, J., Salgado, J., Salguero, E., Sarasso, M., Savietto, H., Schnorhk, A., Schultheis, M., Sciacca, E., Segol, M., Segovia, J. C., Segransan, D., Serpell, E., Shih, I. C., Smareglia, R., Smart, R. L., Smith, C., Solano, E., Solitro, F., Sordo, R., Soria Nieto, S., Souchay, J., Spagna, A., Spoto, F., Stampa, U., Steele, I. A., Steidelmüller, H., Stephenson, C. A., Stoev, H., Suess, F. F., Süveges, M., Surdej, J., Szabados, L., Szegedi-Elek, E., Tapiador, D., Taris, F., Tauran, G., Taylor, M. B., Teixeira, R., Terrett, D., Tingley, B., Trager, S. C., Turon, C., Ulla, A., Utrilla, E., Valentini, G., van Elteren, A., Van Hemelryck, E., van Leeuwen, M., Varadi, M., Vecchiato, A., Veljanoski, J., Via, T., Vicente, D., Vogt, S., Voss, H., Votruba, V., Voutsinas, S., Walmsley, G., Weiler, M., Weingrill, K., Werner, D., Wevers, T., Whitehead, G., Wyrzykowski, Ł., Yoldas, A., Žerjal, M., Zucker, S., Zurbach, C., Zwitter, T., Alecu, A., Allen, M., Allende Prieto, C., Amorim, A., Anglada-Escudé, G., Arsenijevic, V., Azaz, S., Balm, P., Beck, M., Bernstein, H. H., Bigot, L., Bijaoui, A., Blasco, C., Bonfigli, M., Bono, G., Boudreault, S., Bressan, A., Brown, S., Brunet, P. M., Bunclark, P., Buonanno, R., Butkevich, A. G., Carret, C., Carrion, C., Chemin, L., Chéreau, F., Corcione, L., Darmigny, E., de Boer, K. S., de Teodoro, P., de Zeeuw, P. T., Delle Luche, C., Domingues, C. D., Dubath, P., Fodor, F., Frézouls, B., Fries, A., Fustes, D., Fyfe, D., Gallardo, E., Gallegos, J., Gardiol, D., Gebran, M., Gomboc, A., Gómez, A., Grux, E., Gueguen, A., Heyrovsky, A., Hoar, J., Iannicola, G., Isasi Parache, Y., Janotto, A. M., Joliet, E., Jonckheere, A., Keil, R., Kim, D. W., Klagyivik, P., Klar, J., Knude, J., Kochukhov, O., Kolka, I., Kos, J., Kutka, A., Lainey, V., LeBouquin, D., Liu, C., Loreggia, D., Makarov, V. V., Marseille, M. G., Martayan, C., Martinez-Rubi, O., Massart, B., Meynadier, F., Mignot, S., Munari, U., Nguyen, A. T., Nordlander, T., Ocvirk, P., O'Flaherty, K. S., Olias Sanz, A., Ortiz, P., Osorio, J., Oszkiewicz, D., Ouzounis, A., Palmer, M., Park, P., Pasquato, E., Peltzer, C., Peralta, J., Pétuaud, F., Pieniluoma, T., Pigozzi, E., Poels, J., Prat, G., Prod'homme, T., Raison, F., Rebordao, J. M., Risquez, D., Rocca-Volmerange, B., Rosen, S., Ruiz-Fuertes, M. I., Russo, F., Sembay, S., Serraller Vizcaino, I., Short, A., Siebert, A., Silva, H., Sinachopoulos, D., Slezak, E., Soffel, M., Sosnowska, D., Straižys, V., ter Linden, M., Terrell, D., Theil, S., Tiede, C., Troisi, L., Tsalmantza, P., Tur, D., Vaccari, M., Vachier, F., Valles, P., Van Hamme, W., Veltz, L., Virtanen, J., Wallut, J. M., Wichmann, R., Wilkinson, M. I., Ziaeepour, H., & Zschocke, S. (2016). The Gaia mission. A&A, 595, A1.

128. Galvin, T. J., Huynh, M. T., Norris, R. P., Wang, X. R., Hopkins, E., Polsterer, K., Ralph, N. O., O'Brien, A. N., & Heald, G. H. (2020). Cataloguing the radio-sky with unsupervised machine learning: a new approach for the SKA era. MNRAS, 497(3), 2730–2758.

129. Garaldi, E., Pakmor, R., & Springel, V. (2021). Magnetogenesis around the first galaxies: the impact of different field seeding

processes on galaxy formation. MNRAS, 502(4), 5726–5744.

130. Gershtein, S. S. & Zel'dovich, Y. B. (1966). Rest Mass of Muonic Neutrino and Cosmology. *Soviet Journal of Experimental and Theoretical Physics Letters*, 4, 120–122.

131. Gheller, C. & Vazza, F. (2019). A survey of the thermal and non-thermal properties of cosmic filaments. MNRAS, 486(1), 981–1002.

132. Gheller, C. & Vazza, F. (2020). Multiwavelength cross-correlation analysis of the simulated cosmic web. MNRAS, 494(4), 5603–5618.

133. Gheller, C., Vazza, F., & Bonafede, A. (2018). Deep learning based detection of cosmological diffuse radio sources. MNRAS, 480(3), 3749–3761.

134. Gheller, C., Vazza, F., Brüggen, M., Alpaslan, M., Holwerda, B. W., Hopkins, A. M., & Liske, J. (2016). Evolution of cosmic filaments and of their galaxy population from MHD cosmological simulations. MNRAS, 462(1), 448–463.

135. Gheller, C., Vazza, F., Favre, J., & Brüggen, M. (2015). Properties of cosmological filaments extracted from Eulerian simulations. MNRAS, 453(2), 1164–1185.

136. Ghosh, A., Prasad, J., Bharadwaj, S., Ali, S. S., & Chengalur, J. N. (2012). Characterizing foreground for redshifted 21 cm radiation: 150 MHz Giant Metrewave Radio Telescope observations. MNRAS, 426(4), 3295–3314.

137. Gillmor, C. S. (1982). Wilhelm Altar, Edward Appleton, and the Magneto-Ionic Theory. *Proceedings of the American Philosophical Society*, 126(5), 395–440.

138. Girshick, R. (2015). Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision* (pp. 1440–1448).

139. Gordon, Y. A., Boyce, M. M., O'Dea, C. P., Rudnick, L., Andernach, H., Vantyghem, A. N., Baum, S. A., Bui, J.-P., Dionyssiou, M., Safi-Harb, S., & Sander, I. (2021). A Quick Look at the 3 GHz Radio Sky. I. Source Statistics from the Very Large Array Sky Survey. ApJS, 255(2), 30.

140. Govoni, F., Orrù, E., Bonafede, A., Iacobelli, M., Paladino, R., Vazza, F., Murgia, M., Vacca, V., Giovannini, G., Feretti, L., Loi, F., Bernardi, G., Ferrari, C., Pizzo, R. F., Gheller, C., Manti, S., Brüggen, M., Brunetti, G., Cassano, R., de Gasperin, F., Enßlin, T. A., Hoeft, M., Horellou, C., Junklewitz, H., Röttgering, H. J. A., Scaife, A. M. M., Shimwell, T. W., van Weeren, R. J., & Wise, M. (2019). A radio ridge connecting two galaxy clusters in a filament of the cosmic web. *Science*, 364(6444), 981–984.

141. Gregori, G., Ravasio, A., Murphy, C. D., Schaar, K., Baird, A., Bell, A. R., Benuzzi-Mounaix, A., Bingham, R., Constantin, C., Drake, R. P., Edwards, M., Everson, E. T., Gregory, C. D., Kuramitsu, Y., Lau, W., Mithen, J., Niemann, C., Park, H. S., Remington, B. A., Reville, B., Robinson, A. P. L., Ryutov, D. D., Sakawa, Y., Yang, S., Woolsey, N. C., Koenig, M., & Miniati, F. (2012). Generation of scaled protogalactic seed magnetic fields in laser-produced shock waves. Nature, 481(7382), 480–483.

142. Guidetti, D., Laing, R. A., Bridle, A. H., Parma, P., & Gregorini, L. (2011). Ordered magnetic fields around radio galaxies: evidence for interaction with the environment. MNRAS, 413(4), 2525–2544.

143. Gürkan, G., Hardcastle, M. J., & Jarvis, M. J. (2014). The Wide-field Infrared Survey Explorer properties of complete samples of radio-loud active galactic nucleus. MNRAS, 438(2), 1149–1161.

144. Gürkan, G., Prandoni, I., O'Brien, A., Raja, W., Marchetti, L., Vaccari, M., Driver, S., Taylor, E., Franzen, T., Brown, M. J. I., Shabala, S., Andernach, H., Hopkins, A. M., Norris, R. P., Leahy, D., Bilicki, M., Farajollahi, H., Galvin, T., Heald, G., Koribalski, B. S., An, T., & Warhurst, K. (2022). Deep ASKAP EMU Survey of the GAMA23 field: properties of radio sources. MNRAS, 512(4), 6104–6121.

145. Hahn, O., Porciani, C., Carollo, C. M., & Dekel, A. (2007). Properties of dark matter haloes in clusters, filaments, sheets and voids. MNRAS, 375(2), 489–499.

146. Hales, S. E. G., Baldwin, J. E., & Warner, P. J. (1988). The 6C survey of radio sources - II. The zone $30 < \delta < 51$, 08h30m < α <17h30m. MNRAS, 234, 919–936.

147. Hales, S. E. G., Masson, C. R., Warner, P. J., & Baldwin, J. E. (1990). The 6C survey of radio sources - III. The zone 48<Dec<68, 05h25<RA<18h17. MNRAS, 246, 256–262.

148. Hallinan, G., Ravi, V., Weinreb, S., Kocz, J., Huang, Y., Woody, D. P., Lamb, J., D'Addario, L., Catha, M., Law, C., Kulkarni, S. R., Phinney, E. S., Eastwood, M. W., Bouman, K., McLaughlin, M., Ransom, S., Siemens, X., Cordes, J., Lynch, R., Kaplan, D., Brazier, A., Bhatnagar, S., Myers, S., Walter, F., & Gaensler, B. (2019). The DSA-2000 — A Radio Survey Camera. In *Bulletin of the American Astronomical Society*, volume 51 (pp. 255).

149. Hamaker, J. P., Bregman, J. D., & Sault, R. J. (1996). Understanding radio polarimetry. I. Mathematical foundations. A&AS, 117, 137–147.

150. Hardcastle, M. J. (2018). A simulation-based analytic model of radio galaxies. MNRAS, 475(2), 2768–2786.

151. Hardcastle, M. J., Alexander, P., Pooley, G. G., & Riley, J. M. (1998a). FRII radio galaxies with $z < 0.3$ - I. Properties of jets, cores and hotspots. MNRAS, 296(2), 445–462.

152. Hardcastle, M. J. & Croston, J. H. (2020). Radio galaxies and feedback from AGN jets. New A Rev., 88, 101539.

153. Hardcastle, M. J., Horton, M. A., Williams, W. L., Duncan, K. J., Alegre, L., Barkus, B., Croston, J. H., Dickinson, H., Osinga, E., Röttgering, H. J. A., Sabater, J., Shimwell, T. W., Smith, D. J. B., Best, P. N., Botteon, A., Brüggen, M., Drabent, A., de Gasperin, F., Gürkan, G., Hajduk, M., Hale, C. L., Hoeft, M., Jamrozy, M., Kunert-Bajraszewska, M., Kondapally, R., Magliocchetti, M., Mahatma, V. H., Mostert, R. I. J., O'Sullivan, S. P., Pajdosz-Śmierciak, U., Petley, J., Pierce, J. C. S., Prandoni, I., Schwarz, D. J., Shulewski, A., Siewert, T. M., Stott, J. P., Tang, H., Vaccari, M., Zheng, X., Bailey, T., Desbled, S., Goyal, A., Gonano, V., Hanset, M., Kurtz, W., Lim, S. M., Mielle, L., Molloy, C. S., Roth, R., Terentev, I. A., & Torres, M. (2023). The LOFAR Two-Metre Sky Survey. VI. Optical identifications for the second data release. A&A, 678, A151.

154. Hardcastle, M. J. & Krause, M. G. H. (2013). Numerical modelling of the lobes of radio galaxies in cluster environments. MNRAS, 430(1), 174–196.

155. Hardcastle, M. J., Shimwell, T. W., Tasse, C., Best, P. N., Drabent, A., Jarvis, M. J., Prandoni, I., Röttgering, H. J. A., Sabater, J., & Schwarz, D. J. (2021). The contribution of discrete sources to the sky temperature at 144 MHz. A&A, 648, A10.

156. Hardcastle, M. J., Williams, W. L., Best, P. N., Croston, J. H., Duncan, K. J., Röttgering, H. J. A., Sabater, J., Shimwell, T. W., Tasse, C., Callingham, J. R., Cochrane, R. K., de Gasperin, F., Gürkan, G., Jarvis, M. J., Mahatma, V., Miley, G. K., Mingo, B., Mooney, S., Morabito, L. K., O'Sullivan, S. P., Prandoni, I., Shulevski, A., & Smith, D. J. B. (2019). Radio-loud AGN in the first LoTSS data release. The lifetimes and environmental impact of jet-driven sources. A&A, 622, A12.

157. Hardcastle, M. J., Worrall, D. M., & Birkinshaw, M. (1998b). Dynamics of the radio galaxy 3C449. MNRAS, 296(4), 1098–1104.

158. Harrison, I., Camera, S., Zuntz, J., & Brown, M. L. (2016). SKA weak lensing - I. Cosmological forecasts and the power of radio-optical cross-correlations. MNRAS, 463, 3674–3685.

159. Hart, R. E., Bamford, S. P., Willett, K. W., Masters, K. L., Cardamone, C., Lintott, C. J., Mackay, R. J., Nichol, R. C., Rosslowe, C. K., Simmons, B. D., & Smethurst, R. J. (2016). Galaxy Zoo: comparing the demographics of spiral arm number and a new method for correcting redshift bias. MNRAS, 461(4), 3663–3682.

160. Heckman, T. M. & Best, P. N. (2014). The Coevolution of Galaxies and Supermassive Black Holes: Insights from Surveys of the Contemporary Universe. ARA&A, 52, 589–660.

161. Heesen, V., Croston, J. H., Morganti, R., Hardcastle, M. J., Stewart, A. J., Best, P. N., Broderick, J. W., Brüggen, M., Brunetti, G., ChyŻy, K. T., Harwood, J. J., Haverkorn, M., Hess, K. M., Intema, H. T., Jamrozy, M., Kunert-Bajraszewska, M., McKean, J. P., Orrú, E., Röttgering, H. J. A., Shimwell, T. W., Shulevski, A., White, G. J., Wilcots, E. M., & Williams, W. L. (2018). LOFAR reveals the giant: a low-frequency radio continuum study of the outflow in the nearby FR I radio galaxy 3C 31. MNRAS, 474(4), 5049–5067.

162. Hendriks, J. N., Jidling, C., Wills, A., & Schön, T. B. (2018). Evaluating the squared-exponential covariance function in Gaussian processes with integral observations. *arXiv e-prints*, (pp. arXiv:1812.07319).

163. Hodgson, T., Vazza, F., Johnston-Hollitt, M., & McKinley, B. (2021). Figaro simulation: Filaments & galactic radio simulation. *Publications of the Astronomical Society of Australia*, 38, e047.

164. Hoeft, M. & Brüggen, M. (2007). Radio signature of cosmological structure formation shocks. MNRAS, 375(1), 77–91.

165. Hoffman, Y., Metuki, O., Yepes, G., Gottlöber, S., Forero-Romero, J. E., Libeskind, N. I., & Knebe, A. (2012). A kinematic classification of the cosmic web. MNRAS, 425(3), 2049–2057.

166. Högbom, J. A. (1974). Aperture Synthesis with a Non-Regular Distribution of Interferometer Baselines. A&AS, 15, 417.

167. Hurley-Walker, N., Callingham, J. R., Hancock, P. J., Franzen, T. M. O., Hindson, L., Kapińska, A. D., Morgan, J., Offringa, A. R., Wayth, R. B., Wu, C., Zheng, Q., Murphy, T., Bell, M. E., Dwarakanath, K. S., For, B., Gaensler, B. M., Johnston-Hollitt, M., Lenc, E., Procopio, P., Staveley-Smith, L., Ekers, R., Bowman, J. D., Briggs, F., Cappallo, R. J., Deshpande, A. A., Greenhill, L., Hazelton, B. J., Kaplan, D. L., Lonsdale, C. J., McWhirter, S. R., Mitchell, D. A., Morales, M. F., Morgan, E., Oberoi, D., Ord, S. M., Prabu, T., Shankar, N. U., Srivani, K. S., Subrahmanyan, R., Tingay, S. J., Webster, R. L., Williams, A., & Williams, C. L. (2017). GaLactic and Extragalactic All-sky Murchison Widefield Array (GLEAM) survey - I. A low-frequency extragalactic catalogue. MNRAS, 464, 1146–1167.

168. Ineson, J., Croston, J. H., Hardcastle, M. J., & Mingo, B. (2017). A representative survey of the dynamics and energetics of FR II radio galaxies. MNRAS, 467(2), 1586–1607.

169. Intema, H. T., van der Tol, S., Cotton, W. D., Cohen, A. S., van Bemmel, I. M., & Röttgering, H. J. A. (2009a). Ionospheric calibration of low frequency radio interferometric observations using the peeling scheme. I. Method description and first results. A&A, 501, 1185–1205.

170. Intema, H. T., van der Tol, S., Cotton, W. D., Cohen, A. S., van Bemmel, I. M., & Röttgering, H. J. A. (2009b). Ionospheric calibration of low frequency radio interferometric observations using the peeling scheme. I. Method description and first results. A&A, 501(3), 1185–1205.

171. Ishwara-Chandra, C. H. & Saikia, D. J. (1999). Giant radio sources. MNRAS, 309(1), 100–112.

172. Ishwara-Chandra, C. H., Taylor, A. R., Green, D. A., Stil, J. M., Vaccari, M., & Ocran, E. F. (2020). A wide-area GMRT 610-MHz survey of ELAIS N1 field. MNRAS, 497(4), 5383–5394.

173. Jamrozy, M., Konar, C., Machalski, J., & Saikia, D. J. (2008). A multifrequency study of giant radio sources - II. Spectral ageing analysis of the lobes of selected sources. MNRAS, 385(3), 1286–1296.

174. Jamrozy, M. & Machalski, J. (2002). Energy Density and Radiation Losses in Giant Radio Galaxies. In R. F. Green, E. Y. Khachikian, & D. B. Sanders (Eds.), *IAU Colloq. 184: AGN Surveys*, volume 284 of *Astronomical Society of the Pacific Conference Series* (pp. 295).

175. Jamrozy, M., Machalski, J., Mack, K. H., & Klein, U. (2005). Ageing analysis of the giant radio galaxy J1343+3758. A&A, 433(2), 467–477.

176. Jansen, R. A., Franx, M., Fabricant, D., & Caldwell, N. (2000). Surface Photometry of Nearby Field Galaxies: The Data. ApJS, 126(2), 271–329.

177. Jasche, J. & Kitaura, F. S. (2010a). Fast Hamiltonian sampling for large-scale structure inference. MNRAS, 407(1), 29–42.

178. Jasche, J. & Kitaura, F. S. (2010b). Fast Hamiltonian sampling for large-scale structure inference. MNRAS, 407(1), 29–42.

179. Jasche, J., Kitaura, F. S., Wandelt, B. D., & Enßlin, T. A. (2010a). Bayesian power-spectrum inference for large-scale structure data. MNRAS, 406(1), 60–85.

180. Jasche, J., Kitaura, F. S., Wandelt, B. D., & Enßlin, T. A. (2010b). Bayesian power-spectrum inference for large-scale structure data. MNRAS, 406(1), 60–85.

181. Jasche, J. & Lavaux, G. (2019). Physical Bayesian modelling of the non-linear matter distribution: New insights into the nearby universe. A&A, 625, A64.

182. Jasche, J., Leclercq, F., & Wandelt, B. D. (2015). Past and present cosmic structure in the SDSS DR7 main sample. J. Cosmology Astropart. Phys., 2015(1), 036.

183. Jasche, J. & Wandelt, B. D. (2012). Bayesian inference from photometric redshift surveys. MNRAS, 425(2), 1042–1056.

184. Jasche, J. & Wandelt, B. D. (2013). Bayesian physical reconstruction of initial conditions from large-scale structure surveys. MNRAS, 432(2), 894–913.

185. Jedamzik, K. & Pogosian, L. (2020). Relieving the Hubble Tension with Primordial Magnetic Fields. Phys. Rev. Lett., 125(18), 181302.

186. Jeffreys, H. (1925). On Certain Approximate Solutions of Linear Differential Equations of the Second Order*. *Proceedings of the London Mathematical Society*, s2-23(1), 428–436.

187. Jidling, C., Hendriks, J., Wahlström, N., Gregg, A., Schön, T. B., Wensrich, C., & Wills, A. (2018). Probabilistic modelling and reconstruction of strain. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 436, 141 – 155.

188. Jones, D. H., Saunders, W., Colless, M., Read, M. A., Parker, Q. A., Watson, F. G., Campbell, L. A., Burkey, D., Mauch, T., Moore, L., Hartley, M., Cass, P., James, D., Russell, K., Fiegert, K., Dawe, J., Huchra, J., Jarrett, T., Lahav, O., Lucey, J., Mamon, G. A., Proust, D., Sadler, E. M., & Wakamatsu, K.-i. (2004). The 6dF Galaxy Survey: samples, observational techniques and the first data release. MNRAS, 355(3), 747–763.

189. Jones, F. C. & Ellison, D. C. (1991). The plasma physics of shock acceleration. Space Sci. Rev., 58(1), 259–346.

190. Jones, R. C. (1941). New calculus for the treatment of optical systems. I. Description and discussion of the calculus. *Journal of the Optical Society of America (1917-1983)*, 31, 488.

191. Jordan, C. H., Murray, S., Trott, C. M., Wayth, R. B., Mitchell, D. A., Rahimi, M., Pindor, B., Procopio, P., & Morgan, J. (2017). Characterization of the ionosphere above the Murchison Radio Observatory using the Murchison Widefield Array. MNRAS, 471(4), 3974–3987.

192. Kaiser, C. R. & Alexander, P. (1997). A self-similar model for extragalactic radio sources. MNRAS, 286(1), 215–222.

193. Kale, R. (2020). Low frequency observations of radio relics and halos. In K. Asada, E. de Gouveia Dal Pino, M. Giroletti, H. Nagai, & R. Nemmen (Eds.), *IAU Symposium*, volume 342 of *IAU Symposium* (pp. 37–43).

194. Kandus, A., Kunze, K. E., & Tsagas, C. G. (2011). Primordial magnetogenesis. Phys. Rep., 505(1), 1–58.

195. Kannappan, S. J., Guie, J. M., & Baker, A. J. (2009). E/So Galaxies on the Blue Color-Stellar Mass Sequence at $z = 0$: Fading Mergers or Future Spirals? *The Astronomical Journal*, 138(2), 579–597.

196. Kannappan, S. J., Stark, D. V., Eckert, K. D., Moffett, A. J., Wei, L. H., Pisano, D. J., Baker, A. J., Vogel, S. N., Fabricant, D. G., Laine, S., Norris, M. A., Jogee, S., Lepore, N., Hough, L. E., & Weinberg-Wolf, J. (2013). Connecting Transitions in Galaxy Properties to Refueling. ApJ, 777(1), 42.

197. Kazemi, S., Yatawatta, S., Zaroubi, S., Lampropoulos, P., de Bruyn, A. G., Koopmans, L. V. E., & Noordam, J. (2011). Radio interferometric calibration using the SAGE algorithm. MNRAS, 414, 1656–1666.

198. Keller, J., Bellman, R., & Society, A. M. (1964). *Stochastic Equations and Wave Propagation in Random Media*. Proceedings of symposia in applied mathematics. American Mathematical Society.

199. Keshet, U., Waxman, E., & Loeb, A. (2004). Imprint of Intergalactic Shocks on the Radio Sky. ApJ, 617(1), 281–302.

200. King, A. (2016). How big can a black hole grow? MNRAS, 456(1), L109–L112.

201. King, A. & Pounds, K. (2015). Powerful outflows and feedback from active galactic nuclei. *Annual Review of Astronomy and Astrophysics*, 53(1), 115–154.

202. Kirk, J. G. & Schneider, P. (1987). On the Acceleration of Charged Particles at Relativistic Shock Fronts. ApJ, 315, 425.

203. Kitaura, F.-S., Ata, M., Rodríguez-Torres, S. A., Hernández-Sánchez, M., Balaguera-Antolínez, A., & Yepes, G. (2021). COSMIC BIRTH: efficient Bayesian inference of the evolving cosmic web from galaxy surveys. MNRAS, 502(3), 3456–3475.

204. Kitaura, F. S. & Enßlin, T. A. (2008). Bayesian reconstruction of the cosmological large-scale structure: methodology, inverse algorithms and numerical optimization. MNRAS, 389(2), 497–544.

205. Kivelson, M. G. & Russell, C. T. (1995). *Introduction to Space Physics*.

206. Kolmogorov, A. N. (1956). *Foundations of the theory of probability*. Chelsea Pub Co, 2 english edition.

207. Kolmogorov, A. N. (1991). The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. *Proceedings of the Royal Society of London Series A*, 434, 9–13.

208. Koopmans, L. V. E. (2010). Ionospheric Power-spectrum Tomography in Radio Interferometry. ApJ, 718, 963–971.

209. Kormendy, J. & Ho, L. C. (2013). Coevolution (Or Not) of Supermassive Black Holes and Host Galaxies. ARA&A, 51(1), 511–653.

210. Krause, M., Alexander, P., Riley, J., & Hopton, D. (2012). A new connection between the jet opening angle and the large-scale morphology of extragalactic radio sources. MNRAS, 427(4), 3196–3208.

211. Kroupa, P. (2001). On the variation of the initial mass function. MNRAS, 322(2), 231–246.

212. Krumholz, M. R. & Federrath, C. (2019). The role of magnetic fields in setting the star formation rate and the initial mass function. *Frontiers in Astronomy and Space Sciences*, 6.

213. Krymskii, G. F. (1977). A regular mechanism for the acceleration of charged particles on the front of a shock wave. *Akademiia Nauk SSSR Doklady*, 234, 1306–1308.

214. Kuminski, E. & Shamir, L. (2016). A Computer-generated Visual Morphology Catalog of ~3,000,000 SDSS Galaxies. ApJS, 223(2), 20.

215. Kuźmicz, A. & Jamrozy, M. (2021). Giant Radio Quasars: Sample and Basic Properties. ApJS, 253(1), 25.

216. Kuźmicz, A., Jamrozy, M., Bronarska, K., Janda-Boczar, K., & Saikia, D. J. (2018). An Updated Catalog of Giant Radio Sources. ApJS, 238(1), 9.

217. La Porta, L., Burigana, C., Reich, W., & Reich, P. (2008). The impact of Galactic synchrotron emission on CMB anisotropy measurements. I. Angular power spectrum analysis of total intensity all-sky surveys. A&A, 479(3), 641–654.

218. Lacy, M., Baum, S. A., Chandler, C. J., Chatterjee, S., Clarke, T. E., Deustua, S., English, J., Farnes, J., Gaensler, B. M., Gugliucci, N., Hallinan, G., Kent, B. R., Kimball, A., Law, C. J., Lazio, T. J. W., Marvil, J., Mao, S. A., Medlin, D., Mooley, K., Murphy, E. J., Myers, S., Osten, R., Richards, G. T., Rosolowsky, E., Rudnick, L., Schinzel, F., Sivakoff, G. R., Sjouwerman, L. O., Taylor, R., White, R. L., Wrobel, J., Andernach, H., Beasley, A. J., Berger, E., Bhatnager, S., Birkinshaw, M., Bower, G. C., Brandt, W. N., Brown, S., Burke-Spolaor, S., Butler, B. J., Comerford, J., Demorest, P. B., Fu, H., Giacintucci, S., Golap, K., Güth, T., Hales, C. A., Hiriart, R., Hodge, J., Horesh, A., Ivezić, Ž., Jarvis, M. J., Kamble, A., Kassim, N., Liu, X., Loinard, L., Lyons, D. K., Masters, J., Mezcua, M., Moellenbrock, G. A., Mroczkowski, T., Nyland, K., O'Dea, C. P., O'Sullivan, S. P., Peters, W. M., Radford, K., Rao, U., Robnett, J., Salcido, J., Shen, Y., Sobotka, A., Witz, S., Vaccari, M., van Weeren, R. J., Vargas, A., Williams, P. K. G., & Yoon, I. (2020). The Karl G. Jansky Very Large Array Sky Survey (VLASS). Science Case and Survey Design. PASP, 132(1009), 035001.

219. Lacy, M., Rawlings, S., Saunders, R., & Warner, P. J. (1993). 8C 0821+695 : a giant radio galaxy at Z = 0.538. MNRAS, 264, 721–728.

220. Laing, R. A. & Bridle, A. H. (2014). Systematic properties of decelerating relativistic jets in low-luminosity radio galaxies. MNRAS, 437(4), 3405–3441.

221. Lan, T.-W. & Prochaska, J. X. (2021). On the environments of giant radio galaxies. MNRAS, 502(4), 5104–5114.

222. Lara, L., Mack, K. H., Lacy, M., Klein, U., Cotton, W. D., Feretti, L., Giovannini, G., & Murgia, M. (2000). The giant radio galaxy 8C 0821+695 and its environment. A&A, 356, 63–72.

223. Lavaux, G. & Jasche, J. (2016). Unmasking the masked Universe: the 2M++ catalogue through Bayesian eyes. MNRAS, 455(3), 3169–3179.

224. Lavaux, G., Jasche, J., & Leclercq, F. (2019). Systematic-free inference of the cosmic matter density field from SDSS3-BOSS data. *arXiv e-prints*, (pp. arXiv:1909.06396).

225. Lavaux, G. & Wandelt, B. D. (2010). Precision cosmology with voids: definition, methods, dynamics. MNRAS, 403(3), 1392–1408.

226. Leclercq, F., Jasche, J., Lavaux, G., Wand elt, B., & Percival, W. (2017). The phase-space structure of nearby dark matter as constrained by the SDSS. J. Cosmology Astropart. Phys., 2017(6), 049.

227. Leclercq, F., Jasche, J., & Wandelt, B. (2015). Bayesian analysis of the dynamic cosmic web in the SDSS galaxy survey. J. Cosmology Astropart. Phys., 2015(6), 015.

228. Leclercq, F., Lavaux, G., Jasche, J., & Wandelt, B. (2016). Comparing cosmic web classifiers using information theory. J. Cosmology Astropart. Phys., 2016(8), 027.

229. Li, F., Zhang, H., Liu, S., Zhang, L., Ni, L. M., Shum, H.-Y., et al. (2022). Mask dino: Towards a unified transformer-based framework for object detection and segmentation. *arXiv preprint arXiv:2206.02777*.

230. Li, Y.-L., Konhauser, K. O., & Zhai, M. (2017). The formation of magnetite in the early Archean oceans. *Earth and Planetary Science Letters*, 466, 103–114.

231. Linde, A. (2008). *Inflationary Cosmology*, volume 738, (pp.1). Springer-Verlag Berlin Heidelberg.

232. Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., & Guo, B. (2021). Swin transformer: Hierarchical vision transformer using shifted windows. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 10012–10022).

233. Liu, Z., Mao, H., Wu, C.-Y., Feichtenhofer, C., Darrell, T., & Xie, S. (2022). A convnet for the 2020s. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 11976–11986).

234. Locatelli, N. T., Rajpurohit, K., Vazza, F., Gastaldello, F., Dallacasa, D., Bonafede, A., Rossetti, M., Stuardi, C., Bonassieux, E., Brunetti, G., Brüggen, M., & Shimwell, T. (2020). Discovering the most elusive radio relic in the sky: diffuse shock acceleration caught in the act? MNRAS, 496(1), L48–L53.

235. Lochner, M. & Bassett, B. A. (2021). ASTRONOMALY: Personalised active anomaly detection in astronomical data. *Astronomy and Computing*, 36, 100481.

236. Lockman, F. J., Jahoda, K., & McCammon, D. (1986). The Structure of Galactic H i in Directions of Low Total Column Density. ApJ, 302, 432.

237. Loi, S. T., Murphy, T., Cairns, I. H., Menk, F. W., Waters, C. L., Erickson, P. J., Trott, C. M., Hurley-Walker, N., Morgan, J., Lenc, E., Offringa, A. R., Bell, M. E., Ekers, R. D., Gaensler, B. M., Lonsdale, C. J., Feng, L., Hancock, P. J., Kaplan, D. L., Bernardi, G., Bowman, J. D., Briggs, F., Cappallo, R. J., Deshpande, A. A., Greenhill, L. J., Hazelton, B. J., Johnston-Hollitt, M., McWhirter, S. R., Mitchell, D. A., Morales, M. F., Morgan, E., Oberoi, D., Ord, S. M., Prabu, T., Shankar, N. U., Srivani, K. S., Subrahmanyan, R., Tingay, S. J., Wayth, R. B., Webster, R. L., Williams, A., & Williams, C. L. (2015). Real-time imaging of density ducts between the plasmasphere and ionosphere. Geophys. Res. Lett., 42, 3707–3714.

238. Lovisari, L., Ettori, S., Gaspari, M., & Giles, P. A. (2021). Scaling properties of galaxy groups. *Universe*, 7(5).

239. Lovisari, L., Reiprich, T. H., & Schellenberger, G. (2015). Scaling properties of a complete X-ray selected galaxy group sample. A&A, 573, A118.

240. Lu, R.-S., Asada, K., Krichbaum, T. P., Park, J., Tazaki, F., Pu, H.-Y., Nakamura, M., Lobanov, A., Hada, K., Akiyama, K., Kim, J.-Y., Marti-Vidal, I., Gómez, J. L., Kawashima, T., Yuan, F., Ros, E., Alef, W., Britzen, S., Bremer, M., Broderick, A. E., Doi, A., Giovannini, G., Giroletti, M., Ho, P. T. P., Honma, M., Hughes, D. H., Inoue, M., Jiang, W., Kino, M., Koyama, S., Lindqvist, M., Liu, J., Marscher, A. P., Matsushita, S., Nagai, H., Rottmann, H., Savolainen, T., Schuster, K.-F., Shen, Z.-Q., de Vicente, P., Walker, R. C., Yang, H., Zensus, J. A., Algaba, J. C., Allardi, A., Bach, U., Berthold, R., Bintley, D., Byun, D.-Y., Casadio, C., Chang, S.-H., Chang, C.-C., Chang, S.-C., Chen, C.-C., Chen, M.-T., Chilson, R., Chuter, T. C., Conway, J., Crew, G. B., Dempsey, J. T., Dornbusch, S., Faber, A., Friberg, P., García, J. G., Garrido, M. G., Han, C.-C., Han, K.-C., Hasegawa, Y., Herrero-Illana, R., Huang, Y.-D., Huang, C.-W. L., Impellizzeri, V., Jiang, H., Jinchi, H., Jung, T., Kallunki, J., Kirves, P., Kimura, K., Koay, J. Y., Koch, P. M., Kramer, C., Kraus, A., Kubo, D., Kuo, C.-Y., Li, C.-T., Lin, L. C.-C., Liu, C.-T., Liu, K.-Y., Lo, W.-P., Lu, L.-M., MacDonald, N., Martin-Cocher, P., Messias, H., Meyer-Zhao, Z., Minter, A., Nair, D. G., Nishioka, H., Norton, T. J., Nystrom, G., Ogawa, H., Oshiro, P., Patel, N. A., Pen, U.-L., Pidopryhora, Y., Pradel, N., Raffin, P. A., Rao, R., Ruiz, I., Sanchez, S., Shaw, P., Snow, W., Sridharan, T. K.,

Srinivasan, R., Tercero, B., Torne, P., Traianou, E., Wagner, J., Walther, C., Wei, T.-S., Yang, J., & Yu, C.-Y. (2023). A ring-like accretion structure in M87 connecting its black hole and jet. Nature, 616(7958), 686–690.

241. Maccagni, F. M., Murgia, M., Serra, P., Govoni, F., Morokuma-Matsui, K., Kleiner, D., Buchner, S., Józsa, G. I. G., Kamphuis, P., Makhathini, S., Molnár, D. C., Prokhorov, D. A., Ramaila, A., Ramatsoku, M., Thorat, K., & Smirnov, O. (2020). The flickering nuclear activity of Fornax A. A&A, 634, A9.

242. Machalski, J. (2011). A multifrequency study of giant radio sources - IV. Length/luminosity asymmetries and dynamical evolution of the lobes of selected sources. MNRAS, 413(4), 2429–2438.

243. Machalski, J. & Jamrozy, M. (2000). J1343+3758: the third largest FRII-type radio galaxy in the Universe. A&A, 363, L17–L20.

244. Machalski, J. & Jamrozy, M. (2006). The new sample of giant radio sources. III. Statistical trends and correlations. A&A, 454(1), 95–102.

245. Machalski, J., Jamrozy, M., & Zola, S. (2001). The new sample of Giant radio sources. I. Radio imaging, optical identification and spectroscopy of selected candidates. A&A, 371, 445–469.

246. Machalski, J., Jamrozy, M., Zola, S., & Koziel, D. (2006). The new sample of giant radio sources. II. Update of optical counterparts, further spectroscopy of identified faint host galaxies, high-frequency radio maps, and polarisation properties of the sources. A&A, 454(1), 85–94.

247. Machalski, J., Koziel-Wierzbowska, D., & Jamrozy, M. (2007). Giant Radio Galaxies as a Probe of the Cosmological Evolution of the IGM. I.Preliminary Deep Detections and Low-Resolution Spectroscopy with SALT. *AcA*, 57, 227–248.

248. Machalski, J., Kozieł-Wierzbowska, D., Jamrozy, M., & Saikia, D. J. (2008). J1420-0545: The Radio Galaxy Larger than 3C 236. ApJ, 679(1), 149–155.

249. Mack, K. H., Klein, U., O'Dea, C. P., Willis, A. G., & Saripalli, L. (1998). Spectral indices, particle ages, and the ambient medium of giant radio galaxies. A&A, 329, 431–442.

250. Macquart, J. P., Prochaska, J. X., McQuinn, M., Bannister, K. W., Bhandari, S., Day, C. K., Deller, A. T., Ekers, R. D., James, C. W., Marnoch, L., Osłowski, S., Phillips, C., Ryder, S. D., Scott, D. R., Shannon, R. M., & Tejos, N. (2020). A census of baryons in the Universe from localized fast radio bursts. Nature, 581(7809), 391–395.

251. Mahajan, S., Drinkwater, M. J., Driver, S., Hopkins, A. M., Graham, A. W., Brough, S., Brown, M. J. I., Holwerda, B. W., Owers, M. S., & Pimbblet, K. A. (2018). Galaxy And Mass Assembly (GAMA): blue spheroids within 87 Mpc. MNRAS, 475(1), 788–799.

252. Mahatma, V. H., Hardcastle, M. J., Williams, W. L., Best, P. N., Croston, J. H., Duncan, K., Mingo, B., Morganti, R., Brienza, M., Cochrane, R. K., Gürkan, G., Harwood, J. J., Jarvis, M. J., Jamrozy, M., Jurlin, N., Morabito, L. K., Röttgering, H. J. A., Sabater, J., Shimwell, T. W., Smith, D. J. B., Shulevski, A., & Tasse, C. (2019). LoTSS DR1: Double-double radio galaxies in the HETDEX field. A&A, 622, A13.

253. Mahato, M., Dabhade, P., Saikia, D. J., Combes, F., Bagchi, J., Ho, L. C., & Raychaudhury, S. (2022). Search and analysis of giant radio galaxies with associated nuclei (SAGAN). III. New insights into giant radio quasars. A&A, 660, A59.

254. Mahony, E. K., Morganti, R., Prandoni, I., van Bemmel, I. M., Shimwell, T. W., Brienza, M., Best, P. N., Brüggen, M., Calistro Rivera, G., de Gasperin, F., Hardcastle, M. J., Harwood, J. J., Heald, G., Jarvis, M. J., Mandal, S., Miley, G. K., Retana-Montenegro, E., Röttgering, H. J. A., Sabater, J., Tasse, C., van Velzen, S., van Weeren, R. J., Williams, W. L., & White, G. J. (2016). The Lockman Hole project: LOFAR observations and spectral index properties of low-frequency radio sources. MNRAS, 463(3), 2997–3020.

255. Makarov, V. V. & Efroimsky, M. (2014). Tidal Dissipation in a Homogeneous Spherical Body. II. Three Examples: Mercury, IO, and Kepler-10 b. ApJ, 795(1), 7.

256. Malanushenko, A., Cheung, M. C. M., DeForest, C. E., Klimchuk, J. A., & Rempel, M. (2022). The Coronal Veil. ApJ, 927(1), 1.

257. Malarecki, J. M., Jones, D. H., Saripalli, L., Staveley-Smith, L., & Subrahmanyan, R. (2015). Giant radio galaxies - II. Tracers of large-scale structure. MNRAS, 449(1), 955–986.

258. Malarecki, J. M., Staveley-Smith, L., Saripalli, L., Subrahmanyan, R., Jones, D. H., Duffy, A. R., & Rioja, M. (2013). Giant radio galaxies - I. Intergalactic barometers. MNRAS, 432(1), 200–224.

259. Malevergne, Y., Pisarenko, V., & Sornette, D. (2011). Testing the Pareto against the lognormal distributions with the uniformly most powerful unbiased test applied to the distribution of cities. Phys. Rev. E, 83(3), 036111.

260. Malkov, M. A. & Drury, L. O. (2001). Nonlinear theory of diffusive acceleration of particles by shock waves. *Reports on Progress in Physics*, 64(4), 429–481.

261. Mamon, G. A., Trevisan, M., Thuan, T. X., Gallazzi, A., & Davé, R. (2020). The frequency of very young galaxies in the

local Universe - II. The view from SDSS spectra. MNRAS, 492(2), 1791–1811.

262. Mandal, S., Intema, H. T., van Weeren, R. J., Shimwell, T. W., Botteon, A., Brunetti, G., de Gasperin, F., Brüggen, M., Di Gennaro, G., Kraft, R., Röttgering, H. J. A., Hardcastle, M., & Tasse, C. (2020). Revived fossil plasma sources in galaxy clusters. A&A, 634, A4.

263. Marinacci, F., Vogelsberger, M., Pakmor, R., Torrey, P., Springel, V., Hernquist, L., Nelson, D., Weinberger, R., Pillepich, A., Naiman, J., & Genel, S. (2018). First results from the IllustrisTNG simulations: radio haloes and magnetic fields. MNRAS, 480(4), 5113–5139.

264. Masini, A., Celotti, A., Grandi, P., Moravec, E., & Williams, W. L. (2021). A new distant giant radio galaxy in the Boötes field serendipitously detected by Chandra. A&A, 650, A51.

265. Mason, A., Morrison, P., & Sadun, A. C. (1988). The radio rings of Hercules A. Nature, 333(6174), 640–642.

266. Massaro, F., Giroletti, M., D'Abrusco, R., Masetti, N., Paggi, A., Cowperthwaite, P. S., Tosti, G., & Funk, S. (2014). The Low-frequency Radio Catalog of Flat-spectrum Sources. ApJS, 213(1), 3.

267. Matthews, A. G. d. G., van der Wilk, M., Nickson, T., Fujii, K., Boukouvalas, A., León-Villagrá, P., Ghahramani, Z., & Hensman, J. (2017). GPflow: A Gaussian process library using TensorFlow. *Journal of Machine Learning Research*, 18(40), 1–6.

268. McKinley, B., Tingay, S. J., Gaspari, M., Kraft, R. P., Matherne, C., Offringa, A. R., McDonald, M., Calzadilla, M. S., Veilleux, S., Shabala, S. S., Gwyn, S. D. J., Bland-Hawthorn, J., Crnojević, D., Gaensler, B. M., & Johnston-Hollitt, M. (2022). Multi-scale feedback and feeding in the closest radio galaxy Centaurus A. *Nature Astronomy*, 6, 109–120.

269. McNamara, B. R. & Nulsen, P. E. J. (2012). Mechanical feedback from active galactic nuclei in galaxies, groups and clusters. *New Journal of Physics*, 14(5), 055023.

270. Mevius, M., van der Tol, S., Pandey, V. N., Vedantham, H. K., Brentjens, M. A., de Bruyn, A. G., Abdalla, F. B., Asad, K. M. B., Bregman, J. D., Brouw, W. N., Bus, S., Chapman, E., Ciardi, B., Fernandez, E. R., Ghosh, A., Harker, G., Iliev, I. T., Jelić, V., Kazemi, S., Koopmans, L. V. E., Noordam, J. E., Offringa, A. R., Patil, A. H., van Weeren, R. J., Wijnholds, S., Yatawatta, S., & Zaroubi, S. (2016). Probing ionospheric structures using the LOFAR radio telescope. *Radio Science*, 51, 927–941.

271. Miley, G. & De Breuck, C. (2008). Distant radio galaxies and their environments. A&A Rev., 15(2), 67–144.

272. Miley, G. G. K. (2010). LOFAR - Origins and Hopes. In J. van Leeuwen (Ed.), *ISKAF2010 Science Meeting* (pp.40).

273. Mingo, B., Croston, J. H., Hardcastle, M. J., Best, P. N., Duncan, K. J., Morganti, R., Rottgering, H. J. A., Sabater, J., Shimwell, T. W., Williams, W. L., Brienza, M., Gurkan, G., Mahatma, V. H., Morabito, L. K., Prandoni, I., Bondi, M., Ineson, J., & Mooney, S. (2019). Revisiting the Fanaroff-Riley dichotomy and radio-galaxy morphology with the LOFAR Two-Metre Sky Survey (LoTSS). MNRAS, 488(2), 2701–2721.

274. Miniati, F., Jones, T. W., Kang, H., & Ryu, D. (2001). Cosmic-Ray Electrons in Groups and Clusters of Galaxies: Primary and Secondary Populations from a Numerical Cosmological Simulation. ApJ, 562(1), 233–253.

275. Mohan, N. & Rafferty, D. (2015). PyBDSF: Python Blob Detection and Source Finder. Astrophysics Source Code Library.

276. Morabito, L. K., Jackson, N. J., Mooney, S., Sweijen, F., Badole, S., Kukreti, P., Venkattu, D., Groeneveld, C., Kappes, A., Bonnassieux, E., Drabent, A., Iacobelli, M., Croston, J. H., Best, P. N., Bondi, M., Callingham, J. R., Conway, J. E., Deller, A. T., Hardcastle, M. J., McKean, J. P., Miley, G. K., Moldon, J., Röttgering, H. J. A., Tasse, C., Shimwell, T. W., van Weeren, R. J., Anderson, J. M., Asgekar, A., Avruch, I. M., van Bemmel, I. M., Bentum, M. J., Bonafede, A., Brouw, W. N., Butcher, H. R., Ciardi, B., Corstanje, A., Coolen, A., Damstra, S., de Gasperin, F., Duscha, S., Eislöffel, J., Engels, D., Falcke, H., Garrett, M. A., Griessmeier, J., Gunst, A. W., van Haarlem, M. P., Hoeft, M., van der Horst, A. J., Jütte, E., Kadler, M., Koopmans, L. V. E., Krankowski, A., Mann, G., Nelles, A., Oonk, J. B. R., Orru, E., Paas, H., Pandey, V. N., Pizzo, R. F., Pandey-Pommier, M., Reich, W., Rothkaehl, H., Ruiter, M., Schwarz, D. J., Shulevski, A., Soida, M., Tagger, M., Vocks, C., Wijers, R. A. M. J., Wijnholds, S. J., Wucknitz, O., Zarka, P., & Zucca, P. (2022). Sub-arcsecond imaging with the International LOFAR Telescope. I. Foundational calibration strategy and pipeline. A&A, 658, A1.

277. Mostert, R. & Oei, M. (2023). Constraining the giant radio galaxy population with machine learning–accelerated detection and Bayesian inference. *in prep.*

278. Mostert, R. I. J., Duncan, K. J., Alegre, L., Röttgering, H. J. A., Williams, W. L., Best, P. N., Hardcastle, M. J., & Morganti, R. (2022). Radio source-component association for the LOFAR Two-metre Sky Survey with region-based convolutional neural networks. A&A, 668, A28.

279. Mostert, R. I. J., Duncan, K. J., Röttgering, H. J. A., Polsterer, K. L., Best, P. N., Brienza, M., Brüggen, M., Hardcastle, M. J., Jurlin, N., Mingo, B., Morganti, R., Shimwell, T., Smith, D., & Williams, W. L. (2021). Unveiling the rarest morphologies of the LOFAR Two-metre Sky Survey radio source population with self-organised maps. A&A, 645, A89.

280. Mostert, R. I. J., Morganti, R., Brienza, M., Duncan, K. J., Oei, M. S. S. L., Röttgering, H. J. A., Alegre, L., Hardcastle,

M. J., & Jurlin, N. (2023). Finding AGN remnant candidates based on radio morphology with machine learning. A&A, 674, A208.

281. Mullis, C. R., Henry, J. P., Gioia, I. M., Böhringer, H., Briel, U. G., Voges, W., & Huchra, J. P. (2001). The North Ecliptic Pole Supercluster. ApJ, 553(2), L115–L118.

282. Murthy, S. K., Kasif, S., Salzberg, S., & Beigel, R. (1993). Oc1: A randomized algorithm for building oblique decision trees. In *Proceedings of AAAI*, volume 93 (pp. 322–327).: Citeseer.

283. Myers, S. T. & Spangler, S. R. (1985). Synchrotron aging in the lobes of luminous radio galaxies. ApJ, 291, 52–62.

284. Naiman, J. P., Pillepich, A., Springel, V., Ramirez-Ruiz, E., Torrey, P., Vogelsberger, M., Pakmor, R., Nelson, D., Marinacci, F., Hernquist, L., Weinberger, R., & Genel, S. (2018). First results from the IllustrisTNG simulations: a tale of two elements - chemical evolution of magnesium and europium. MNRAS, 477(1), 1206–1224.

285. Nakamura, M., Asada, K., Hada, K., Pu, H.-Y., Noble, S., Tseng, C., Toma, K., Kino, M., Nagai, H., Takahashi, K., Algaba, J.-C., Orienti, M., Akiyama, K., Doi, A., Giovannini, G., Giroletti, M., Honma, M., Koyama, S., Lico, R., Niinuma, K., & Tazaki, F. (2018). Parabolic Jets from the Spinning Black Hole in M87. ApJ, 868(2), 146.

286. Narayan, R., Igumenshchev, I. V., & Abramowicz, M. A. (2003). Magnetically Arrested Disk: an Energetically Efficient Accretion Flow. PASJ, 55, L69–L72.

287. Navarro, J. F., Frenk, C. S., & White, S. D. M. (1996). The Structure of Cold Dark Matter Halos. ApJ, 462, 563.

288. Nelson, D., Pillepich, A., Springel, V., Weinberger, R., Hernquist, L., Pakmor, R., Genel, S., Torrey, P., Vogelsberger, M., Kauffmann, G., Marinacci, F., & Naiman, J. (2018). First results from the IllustrisTNG simulations: the galaxy colour bimodality. MNRAS, 475(1), 624–647.

289. Neronov, A. & Vovk, I. (2010). Evidence for Strong Extragalactic Magnetic Fields from Fermi Observations of TeV Blazars. *Science*, 328(5974), 73.

290. Nicastro, F., Kaastra, J., Krongold, Y., Borgani, S., Branchini, E., Cen, R., Dadina, M., Danforth, C. W., Elvis, M., Fiore, F., Gupta, A., Mathur, S., Mayya, D., Paerels, F., Piro, L., Rosa-Gonzalez, D., Schaye, J., Shull, J. M., Torres-Zafra, J., Wijers, N., & Zappacosta, L. (2018). Observations of the missing baryons in the warm–hot intergalactic medium. Nature, 558(7710), 406–409.

291. Nobels, F. S. J., Schaye, J., Schaller, M., Bahé, Y. M., & Chaikin, E. (2022). The interplay between AGN feedback and precipitation of the intracluster medium in simulations of galaxy groups and clusters. MNRAS, 515(4), 4838–4859.

292. Nolting, C., Jones, T. W., O'Neill, B. J., & Mendygral, P. J. (2019a). Interactions between Radio Galaxies and Cluster Shocks. I. Jet Axes Aligned with Shock Normals. ApJ, 876(2), 154.

293. Nolting, C., Jones, T. W., O'Neill, B. J., & Mendygral, P. J. (2019b). Simulated Interactions between Radio Galaxies and Cluster Shocks. II. Jet Axes Orthogonal to Shock Normals. ApJ, 885(1), 80.

294. Norris, R. P., Hopkins, A. M., Afonso, J., Brown, S., Condon, J. J., Dunne, L., Feain, I., Hollow, R., Jarvis, M., Johnston-Hollitt, M., Lenc, E., Middelberg, E., Padovani, P., Prandoni, I., Rudnick, L., Seymour, N., Umana, G., Andernach, H., Alexander, D. M., Appleton, P. N., Bacon, D., Banfield, J., Becker, W., Brown, M. J. I., Ciliegi, P., Jackson, C., Eales, S., Edge, A. C., Gaensler, B. M., Giovannini, G., Hales, C. A., Hancock, P., Huynh, M. T., Ibar, E., Ivison, R. J., Kennicutt, R., Kimball, A. E., Koekemoer, A. M., Koribalski, B. S., López-Sánchez, Á. R., Mao, M. Y., Murphy, T., Messias, H., Pimbblet, K. A., Raccanelli, A., Randall, K. E., Reiprich, T. H., Roseboom, I. G., Röttgering, H., Saikia, D. J., Sharp, R. G., Slee, O. B., Smail, I., Thompson, M. A., Urquhart, J. S., Wall, J. V., & Zhao, G. B. (2011). EMU: Evolutionary Map of the Universe. PASA, 28(3), 215–248.

295. Novikov, I. D. & Thorne, K. S. (1973). Astrophysics of black holes. In *Black Holes (Les Astres Occlus)* (pp. 343–450).

296. Ochsenbein, F., Bauer, P., & Marcout, J. (2000). The VizieR database of astronomical catalogues. A&AS, 143, 23–32.

297. O'Dea, C. P., Baum, S. A., & Stanghellini, C. (1991). What Are the Gigahertz Peaked-Spectrum Radio Sources? ApJ, 380, 66.

298. O'Dea, C. P. & Saikia, D. J. (2021). Compact steep-spectrum and peaked-spectrum radio sources. A&A Rev., 29(1), 3.

299. Oei, M. S. S. L., van Weeren, R. J., de Gasperin, F., Botteon, A., Hardcastle, M. J., Shimwell, T. W., & Röttgering, H. J. A. (in prep.a). The Milky Way in Stokes I at metre wavelengths. A&A.

300. Oei, M. S. S. L., van Weeren, R. J., Gast, A. R. D. J. G. I. B., Botteon, A., Hardcastle, M. J., Dabhade, P., Shimwell, T. W., Röttgering, H. J. A., & Drabent, A. (2023a). Measuring the giant radio galaxy length distribution with the LoTSS. A&A, 672, A163.

301. Oei, M. S. S. L., van Weeren, R. J., Hardcastle, M. J., Botteon, A., Shimwell, T. W., Dabhade, P., Gast, A. R. D. J. G. I. B., Röttgering, H. J. A., Brüggen, M., Tasse, C., Williams, W. L., & Shulevski, A. (2022a). The discovery of a radio galaxy of at least 5 Mpc. A&A, 660, A2.

302. Oei, M. S. S. L., van Weeren, R. J., Hardcastle, M. J., Gast, A. R. D. J. G. I. B., Leclercq, F., Röttgering, H. J. A., Dabhade, P., Shimwell, T. W., & Botteon, A. (in prep.b). Do luminous giants populate special large-scale environments? Or: the radio luminosity–Cosmic Web density relation for radio galaxies. A&A.

303. Oei, M. S. S. L., van Weeren, R. J., Hardcastle, M. J., Vazza, F., Shimwell, T. W., Leclercq, F., Brüggen, M., & Röttgering, H. J. A. (2023b). An intergalactic medium temperature from a giant radio galaxy. MNRAS, 518(1), 240–256.

304. Oei, M. S. S. L., van Weeren, R. J., Vazza, F., Leclercq, F., Gopinath, A., & Röttgering, H. J. A. (2022b). Filamentary baryons and where to find them. A forecast of synchrotron radiation from merger and accretion shocks in the local Cosmic Web. A&A, 662, A87.

305. Offringa, A. R., McKinley, B., Hurley-Walker, N., Briggs, F. H., Wayth, R. B., Kaplan, D. L., Bell, M. E., Feng, L., Neben, A. R., Hughes, J. D., Rhee, J., Murphy, T., Bhat, N. D. R., Bernardi, G., Bowman, J. D., Cappallo, R. J., Corey, B. E., Deshpande, A. A., Emrich, D., Ewall-Wice, A., Gaensler, B. M., Goeke, R., Greenhill, L. J., Hazelton, B. J., Hindson, L., Johnston-Hollitt, M., Jacobs, D. C., Kasper, J. C., Kratzenberg, E., Lenc, E., Lonsdale, C. J., Lynch, M. J., McWhirter, S. R., Mitchell, D. A., Morales, M. F., Morgan, E., Kudryavtseva, N., Oberoi, D., Ord, S. M., Pindor, B., Procopio, P., Prabu, T., Riding, J., Roshi, D. A., Shankar, N. U., Srivani, K. S., Subrahmanyan, R., Tingay, S. J., Waterson, M., Webster, R. L., Whitney, A. R., Williams, A., & Williams, C. L. (2014). WSCLEAN: an implementation of a fast, generic wide-field imager for radio astronomy. MNRAS, 444(1), 606–619.

306. Offringa, A. R. & Smirnov, O. (2017). An optimized algorithm for multiscale wideband deconvolution of radio astronomical images. MNRAS, 471(1), 301–316.

307. O'Sullivan, E., Vrtilek, J. M., David, L. P., Giacintucci, S., Zezas, A., Ponman, T. J., Mamon, G. A., Nulsen, P., & Raychaudhury, S. (2014). Deep Chandra Observations of HCG 16. II. The Development of the Intra-group Medium in a Spiral-rich Group. ApJ, 793(2), 74.

308. O'Sullivan, S. P., Brüggen, M., Vazza, F., Carretti, E., Locatelli, N. T., Stuardi, C., Vacca, V., Vernstrom, T., Heald, G., Horellou, C., Shimwell, T. W., Hardcastle, M. J., Tasse, C., & Röttgering, H. (2020). New constraints on the magnetization of the cosmic web using LOFAR Faraday rotation observations. MNRAS, 495(3), 2607–2619.

309. O'Sullivan, S. P., Feain, I. J., McClure-Griffiths, N. M., Ekers, R. D., Carretti, E., Robishaw, T., Mao, S. A., Gaensler, B. M., Bland-Hawthorn, J., & Stawarz, Ł. (2013). Thermal Plasma in the Giant Lobes of the Radio Galaxy Centaurus A. ApJ, 764(2), 162.

310. O'Sullivan, S. P., Machalski, J., Van Eck, C. L., Heald, G., Brüggen, M., Fynbo, J. P. U., Heintz, K. E., Lara-Lopez, M. A., Vacca, V., Hardcastle, M. J., Shimwell, T. W., Tasse, C., Vazza, F., Andernach, H., Birkinshaw, M., Haverkorn, M., Horellou, C., Williams, W. L., Harwood, J. J., Brunetti, G., Anderson, J. M., Mao, S. A., Nikiel-Wroczyński, B., Takahashi, K., Carretti, E., Vernstrom, T., van Weeren, R. J., Orrú, E., Morabito, L. K., & Callingham, J. R. (2019). The intergalactic magnetic field probed by a giant radio galaxy. A&A, 622, A16.

311. Owen, F. N., Eilek, J. A., & Kassim, N. E. (2000). M87 at 90 Centimeters: A Different Picture. ApJ, 543(2), 611–619.

312. Oxford English Dictionary (2000). Oxford English Dictionary.

313. Pacholczyk, A. G. (1970). *Radio astrophysics. Nonthermal processes in galactic and extragalactic sources*.

314. Parfrey, K., Philippov, A., & Cerutti, B. (2019). First-Principles Plasma Simulations of Black-Hole Jet Launching. Phys. Rev. Lett., 122(3), 035101.

315. Parma, P., de Ruiter, H. R., Mack, K. H., van Breugel, W., Dey, A., Fanti, R., & Klein, U. (1996). 1358+305: a giant radio galaxy at z=0.206. A&A, 311, 49–56.

316. Pasini, T., Finoguenov, A., Brüggen, M., Gaspari, M., de Gasperin, F., & Gozaliasl, G. (2021). Radio galaxies in galaxy groups: kinematics, scaling relations, and AGN feedback. MNRAS, 505(2), 2628–2637.

317. Patil, A. H., Yatawatta, S., Koopmans, L. V. E., de Bruyn, A. G., Brentjens, M. A., Zaroubi, S., Asad, K. M. B., Hatef, M., Jelić, V., Mevius, M., Offringa, A. R., Pandey, V. N., Vedantham, H., Abdalla, F. B., Brouw, W. N., Chapman, E., Ciardi, B., Gehlot, B. K., Ghosh, A., Harker, G., Iliev, I. T., Kakiichi, K., Majumdar, S., Mellema, G., Silva, M. B., Schaye, J., Vrbanec, D., & Wijnholds, S. J. (2017). Upper Limits on the 21 cm Epoch of Reionization Power Spectrum from One Night with LOFAR. ApJ, 838, 65.

318. Peacock, J. A. (1999). *Cosmological Physics*. Cambridge University Press.

319. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.

320. Petrosian, V. (2001). On the Nonthermal Emission and Acceleration of Electrons in Coma and Other Clusters of Galaxies. ApJ, 557(2), 560–572.

321. Pierpaoli, E., Borgani, S., Scott, D., & White, M. (2003). On determining the cluster abundance normalization. MNRAS,

342(1), 163–175.

322. Pillepich, A., Nelson, D., Hernquist, L., Springel, V., Pakmor, R., Torrey, P., Weinberger, R., Genel, S., Naiman, J. P., Marinacci, F., & Vogelsberger, M. (2018). First results from the IllustrisTNG simulations: the stellar mass content of groups and clusters of galaxies. MNRAS, 475(1), 648–675.

323. Pizzo, R. F., de Bruyn, A. G., Feretti, L., & Govoni, F. (2008). Detection of diffuse radio emission at large distance from the center of the galaxy cluster A 2255. A&A, 481(3), L91–L94.

324. Planck Collaboration, Ade, P. A. R., Aghanim, N., Arnaud, M., Ashdown, M., Aumont, J., Baccigalupi, C., Banday, A. J., Barreiro, R. B., Barrena, R., Bartlett, J. G., Bartolo, N., Battaner, E., Battye, R., Benabed, K., Benoît, A., Benoit-Lévy, A., Bernard, J. P., Bersanelli, M., Bielewicz, P., Bikmaev, I., Böhringer, H., Bonaldi, A., Bonavera, L., Bond, J. R., Borrill, J., Bouchet, F. R., Bucher, M., Burenin, R., Burigana, C., Butler, R. C., Calabrese, E., Cardoso, J. F., Carvalho, P., Catalano, A., Challinor, A., Chamballu, A., Chary, R. R., Chiang, H. C., Chon, G., Christensen, P. R., Clements, D. L., Colombi, S., Colombo, L. P. L., Combet, C., Comis, B., Couchot, F., Coulais, A., Crill, B. P., Curto, A., Cuttaia, F., Dahle, H., Danese, L., Davies, R. D., Davis, R. J., de Bernardis, P., de Rosa, A., de Zotti, G., Delabrouille, J., Désert, F. X., Dickinson, C., Diego, J. M., Dolag, K., Dole, H., Donzelli, S., Doré, O., Douspis, M., Ducout, A., Dupac, X., Efstathiou, G., Eisenhardt, P. R. M., Elsner, F., Enßlin, T. A., Eriksen, H. K., Falgarone, E., Fergusson, J., Feroz, F., Ferragamo, A., Finelli, F., Forni, O., Frailis, M., Fraisse, A. A., Franceschi, E., Frejsel, A., Galeotta, S., Galli, S., Ganga, K., Génova-Santos, R. T., Giard, M., Giraud-Héraud, Y., Gjerløw, E., González-Nuevo, J., Górski, K. M., Grainge, K. J. B., Gratton, S., Gregorio, A., Gruppuso, A., Gudmundsson, J. E., Hansen, F. K., Hanson, D., Harrison, D. L., Hempel, A., Henrot-Versillé, S., Hernández-Monteagudo, C., Herranz, D., Hildebrandt, S. R., Hivon, E., Hobson, M., Holmes, W. A., Hornstrup, A., Hovest, W., Huffenberger, K. M., Hurier, G., Jaffe, A. H., Jaffe, T. R., Jin, T., Jones, W. C., Juvela, M., Keihänen, E., Keskitalo, R., Khamitov, I., Kisner, T. S., Kneissl, R., Knoche, J., Kunz, M., Kurki-Suonio, H., Lagache, G., Lamarre, J. M., Lasenby, A., Lattanzi, M., Lawrence, C. R., Leonardi, R., Lesgourgues, J., Levrier, F., Liguori, M., Lilje, P. B., Linden-Vørnle, M., López-Caniego, M., Lubin, P. M., Macías-Pérez, J. F., Maggio, G., Maino, D., Mak, D. S. Y., Mandolesi, N., Mangilli, A., Martin, P. G., Martínez-González, E., Masi, S., Matarrese, S., Mazzotta, P., McGehee, P., Mei, S., Melchiorri, A., Melin, J. B., Mendes, L., Mennella, A., Migliaccio, M., Mitra, S., Miville-Deschênes, M. A., Moneti, A., Montier, L., Morgante, G., Mortlock, D., Moss, A., Munshi, D., Murphy, J. A., Naselsky, P., Nastasi, A., Nati, F., Natoli, P., Netterfield, C. B., Nørgaard-Nielsen, H. U., Noviello, F., Novikov, D., Novikov, I., Olamaie, M., Oxborrow, C. A., Paci, F., Pagano, L., Pajot, F., Paoletti, D., Pasian, F., Patanchon, G., Pearson, T. J., Perdereau, O., Perotto, L., Perrott, Y. C., Perrotta, F., Pettorino, V., Piacentini, F., Piat, M., Pierpaoli, E., Pietrobon, D., Plaszczynski, S., Pointecouteau, E., Polenta, G., Pratt, G. W., Prézeau, G., Prunet, S., Puget, J. L., Rachen, J. P., Reach, W. T., Rebolo, R., Reinecke, M., Remazeilles, M., Renault, C., Renzi, A., Ristorcelli, I., Rocha, G., Rosset, C., Rossetti, M., Roudier, G., Rozo, E., Rubiño-Martín, J. A., Rumsey, C., Rusholme, B., Rykoff, E. S., Sandri, M., Santos, D., Saunders, R. D. E., Savelainen, M., Savini, G., Schammel, M. P., Scott, D., Seiffert, M. D., Shellard, E. P. S., Shimwell, T. W., Spencer, L. D., Stanford, S. A., Stern, D., Stolyarov, V., Stompor, R., Streblyanska, A., Sudiwala, R., Sunyaev, R., Sutton, D., Suur-Uski, A. S., Sygnet, J. F., Tauber, J. A., Terenzi, L., Toffolatti, L., Tomasi, M., Tramonte, D., Tristram, M., Tucci, M., Tuovinen, J., Umana, G., Valenziano, L., Valiviita, J., Van Tent, B., Vielva, P., Villa, F., Wade, L. A., Wandelt, B. D., Wehus, I. K., White, S. D. M., Wright, E. L., Yvon, D., Zacchei, A., & Zonca, A. (2016). Planck 2015 results. XXVII. The second Planck catalogue of Sunyaev-Zeldovich sources. A&A, 594, A27.

325. Planck Collaboration, Aghanim, N., Akrami, Y., Ashdown, M., Aumont, J., Baccigalupi, C., Ballardini, M., Banday, A. J., Barreiro, R. B., Bartolo, N., Basak, S., Battye, R., Benabed, K., Bernard, J. P., Bersanelli, M., Bielewicz, P., Bock, J. J., Bond, J. R., Borrill, J., Bouchet, F. R., Boulanger, F., Bucher, M., Burigana, C., Butler, R. C., Calabrese, E., Cardoso, J. F., Carron, J., Challinor, A., Chiang, H. C., Chluba, J., Colombo, L. P. L., Combet, C., Contreras, D., Crill, B. P., Cuttaia, F., de Bernardis, P., de Zotti, G., Delabrouille, J., Delouis, J. M., Di Valentino, E., Diego, J. M., Doré, O., Douspis, M., Ducout, A., Dupac, X., Dusini, S., Efstathiou, G., Elsner, F., Enßlin, T. A., Eriksen, H. K., Fantaye, Y., Farhang, M., Fergusson, J., Fernandez-Cobos, R., Finelli, F., Forastieri, F., Frailis, M., Fraisse, A. A., Franceschi, E., Frolov, A., Galeotta, S., Galli, S., Ganga, K., Génova-Santos, R. T., Gerbino, M., Ghosh, T., González-Nuevo, J., Górski, K. M., Gratton, S., Gruppuso, A., Gudmundsson, J. E., Hamann, J., Handley, W., Hansen, F. K., Herranz, D., Hildebrandt, S. R., Hivon, E., Huang, Z., Jaffe, A. H., Jones, W. C., Karakci, A., Keihänen, E., Keskitalo, R., Kiiveri, K., Kim, J., Kisner, T. S., Knox, L., Krachmalnicoff, N., Kunz, M., Kurki-Suonio, H., Lagache, G., Lamarre, J. M., Lasenby, A., Lattanzi, M., Lawrence, C. R., Le Jeune, M., Lemos, P., Lesgourgues, J., Levrier, F., Lewis, A., Liguori, M., Lilje, P. B., Lilley, M., Lindholm, V., López-Caniego, M., Lubin, P. M., Ma, Y. Z., Macías-Pérez, J. F., Maggio, G., Maino, D., Mandolesi, N., Mangilli, A., Marcos-Caballero, A., Maris, M., Martin, P. G., Martinelli, M., Martínez-González, E., Matarrese, S., Mauri, N., McEwen, J. D., Meinhold, P. R., Melchiorri, A., Mennella, A., Migliaccio, M., Millea, M., Mitra, S., Miville-Deschênes, M. A., Molinari, D., Montier, L., Morgante, G., Moss, A., Natoli, P., Nørgaard-Nielsen, H. U., Pagano, L., Paoletti, D., Partridge, B., Patanchon, G., Peiris, H. V., Perrotta, F., Pettorino, V., Piacentini, F., Polastri, L., Polenta, G., Puget, J. L., Rachen, J. P., Reinecke, M., Remazeilles, M., Renzi, A., Rocha, G., Rosset, C., Roudier, G., Rubiño-Martín, J. A., Ruiz-Granados, B., Salvati, L., Sandri, M., Savelainen, M., Scott, D., Shellard, E. P. S., Sirignano, C., Sirri, G., Spencer, L. D., Sunyaev, R., Suur-Uski, A. S., Tauber, J. A., Tavagnacco,

D., Tenti, M., Toffolatti, L., Tomasi, M., Trombetti, T., Valenziano, L., Valiviita, J., Van Tent, B., Vibert, L., Vielva, P., Villa, F., Vittorio, N., Wandelt, B. D., Wehus, I. K., White, M., White, S. D. M., Zacchei, A., & Zonca, A. (2020). Planck 2018 results. VI. Cosmological parameters. A&A, 641, A6.

326. Porqueres, N., Jasche, J., Enßlin, T. A., & Lavaux, G. (2018). Imprints of the large-scale structure on AGN formation and evolution. A&A, 612, A31.

327. Pracy, M. B., Ching, J. H. Y., Sadler, E. M., Croom, S. M., Baldry, I. K., Bland-Hawthorn, J., Brough, S., Brown, M. J. I., Couch, W. J., Davis, T. M., Drinkwater, M. J., Hopkins, A. M., Jarvis, M. J., Jelliffe, B., Jurek, R. J., Loveday, J., Pimbblet, K. A., Prescott, M., Wisnioski, E., & Woods, D. (2016). GAMA/WiggleZ: the 1.4 GHz radio luminosity functions of high- and low-excitation radio galaxies and their redshift evolution to $z = 0.75$. MNRAS, 460(1), 2–17.

328. Proctor, D. D. (2016). A Selection of Giant Radio Sources from NVSS. ApJS, 224(2), 18.

329. Ramatsoku, M., Murgia, M., Vacca, V., Serra, P., Makhathini, S., Govoni, F., Smirnov, O., Andati, L. A. L., de Blok, E., Józsa, G. I. G., Kamphuis, P., Kleiner, D., Maccagni, F. M., Molnár, D. C., Ramaila, A. J. T., Thorat, K., & White, S. V. (2020). Collimated synchrotron threads linking the radio lobes of ESO 137-006. A&A, 636, L1.

330. Rasmussen, C. E. & Williams, C. K. I. (2006). *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press.

331. Rees, M. J. (1984). Black Hole Models for Active Galactic Nuclei. ARA&A, 22, 471–506.

332. Rengelink, R. B., Tang, Y., de Bruyn, A. G., Miley, G. K., Bremer, M. N., Roettgering, H. J. A., & Bremer, M. A. R. (1997). The Westerbork Northern Sky Survey (WENSS), I. A 570 square degree Mini-Survey around the North Ecliptic Pole. A&AS, 124, 259–280.

333. Ribeiro, A. L. B., de Carvalho, R. R., Trevisan, M., Capelato, H. V., La Barbera, F., Lopes, P. A. A., & Schilling, A. C. (2013). SPIDER - IX. Classifying galaxy groups according to their velocity distribution. MNRAS, 434(1), 784–795.

334. Rice, J. (2006). *Mathematical Statistics and Data Analysis*. Advanced series. Cengage Learning.

335. Richter, G. A. (1975). Search for Optical Identifications in the 5C3 Radio Survey. II. Statistical Treatment and Results. *Astronomische Nachrichten*, 296(2), 65.

336. Ringermacher, H. I. & Mead, L. R. (2009). A new formula describing the scaffold structure of spiral galaxies. MNRAS, 397(1), 164–171.

337. Rodríguez-Mozos, J. M. & Moya, A. (2019). Erosion of an exoplanetary atmosphere caused by stellar winds. A&A, 630, A52.

338. Ruffini, R., Vereshchagin, G., & Xue, S.-S. (2010). Electron-positron pairs in physics and astrophysics: From heavy nuclei to black holes. Phys. Rep., 487(1-4), 1–140.

339. Rybicki, G. B. & Lightman, A. P. (1986). *Radiative Processes in Astrophysics*. Wiley-VCH.

340. Ryu, C.-M. (2008). Dynamo Mechanism by Transport Flow. In *Statistical Physics, High Energy, Condensed Matter and Mathematical Physics* (pp. 542–542).

341. Ryu, D., Kang, H., Cho, J., & Das, S. (2008). Turbulence and Magnetic Fields in the Large-Scale Structure of the Universe. *Science*, 320(5878), 909.

342. Ryu, D., Kang, H., Hallman, E., & Jones, T. W. (2003). Cosmological shock waves and their role in the large-scale structure of the universe. *The Astrophysical Journal*, 593(2), 599–610.

343. Sabater, J., Best, P. N., Hardcastle, M. J., Shimwell, T. W., Tasse, C., Williams, W. L., Brüggen, M., Cochrane, R. K., Croston, J. H., de Gasperin, F., Duncan, K. J., Gürkan, G., Mechev, A. P., Morabito, L. K., Prandoni, I., Röttgering, H. J. A., Smith, D. J. B., Harwood, J. J., Mingo, B., Mooney, S., & Saxena, A. (2019). The LoTSS view of radio AGN in the local Universe. The most massive galaxies are always switched on. A&A, 622, A17.

344. Safouris, V., Subrahmanyan, R., Bicknell, G. V., & Saripalli, L. (2009). MRCB0319-454: probing the large-scale structure with a giant radio galaxy. MNRAS, 393(1), 2–20.

345. Saikia, D. J., Konar, C., & Kulkarni, V. K. (2006). J0041+3224: a new double-double radio galaxy. MNRAS, 366(4), 1391–1398.

346. Salim, S., Boquien, M., & Lee, J. C. (2018). Dust Attenuation Curves in the Local Universe: Demographics and New Laws for Star-forming Galaxies and High-redshift Analogs. ApJ, 859(1), 11.

347. Salpeter, E. E. (1955). The Luminosity Function and Stellar Evolution. ApJ, 121, 161.

348. Salter, C. J. (1983). Loop-I the North Polar Spur - a Major Feature of the Local Interstellar Environment. *Bulletin of the Astronomical Society of India*, 11, 1.

349. Saripalli, L., Subrahmanyan, R., & Udaya Shankar, N. (2002). A Case for Renewed Activity in the Giant Radio Galaxy J0116-473. ApJ, 565(1), 256–264.

350. Saulder, C., van Kampen, E., Chilingarian, I. V., Mieske, S., & Zeilinger, W. W. (2016). The matter distribution in the local Universe as derived from galaxy groups in SDSS DR12 and 2MRS. A&A, 596, A14.

351. Saxena, A., Jagannathan, P., Röttgering, H. J. A., Best, P. N., Intema, H. T., Zhang, M., Duncan, K. J., Carilli, C. L., & Miley, G. K. (2018). A search for faint high-redshift radio galaxy candidates at 150 MHz. MNRAS, 475(4), 5041–5058.

352. Schaye, J., Crain, R. A., Bower, R. G., Furlong, M., Schaller, M., Theuns, T., Dalla Vecchia, C., Frenk, C. S., McCarthy, I. G., Helly, J. C., Jenkins, A., Rosas-Guevara, Y. M., White, S. D. M., Baes, M., Booth, C. M., Camps, P., Navarro, J. F., Qu, Y., Rahmati, A., Sawala, T., Thomas, P. A., & Trayford, J. (2015). The EAGLE project: simulating the evolution and assembly of galaxies and their environments. MNRAS, 446(1), 521–554.

353. Schaye, J., Dalla Vecchia, C., Booth, C. M., Wiersma, R. P. C., Theuns, T., Haas, M. R., Bertone, S., Duffy, A. R., McCarthy, I. G., & van de Voort, F. (2010). The physics driving the cosmic star formation history. MNRAS, 402(3), 1536–1560.

354. Scheuer, P. A. G. (1974). Models of extragalactic radio sources with a continuous energy supply from a central object. MNRAS, 166, 513–528.

355. Schlafly, E. F., Meisner, A. M., & Green, G. M. (2019). The unWISE Catalog: Two Billion Infrared Sources from Five Years of WISE Imaging. ApJS, 240(2), 30.

356. Schoenmakers, A. P., Mack, K. H., de Bruyn, A. G., Röttgering, H. J. A., Klein, U., & van der Laan, H. (2000). A new sample of giant radio galaxies from the WENSS survey. II. A multi-frequency radio study of a complete sample: Properties of the radio lobes and their environment. A&AS, 146, 293–322.

357. Schoenmakers, A. P., Mack, K. H., Lara, L., Röttgering, H. J. A., de Bruyn, A. G., van der Laan, H., & Giovannini, G. (1998). WNB 0313+683: analysis of a newly discovered giant radio galaxy. A&A, 336, 455–478.

358. Scott, D. & Tout, C. A. (1989). Nearest neighbour analysis of random distributions on a sphere. MNRAS, 241, 109–117.

359. Sebastian, B., Ishwara-Chandra, C. H., Joshi, R., & Wadadekar, Y. (2018). Discovery of a new, 2.2-Mpc giant radio galaxy at a redshift of 0.57. MNRAS, 473(4), 4926–4931.

360. Shabala, S. S. & Godfrey, L. E. H. (2013). Size Dependence of the Radio-luminosity-Mechanical-power Correlation in Radio Galaxies. ApJ, 769(2), 129.

361. Sharp, R., Saunders, W., Smith, G., Churilov, V., Correll, D., Dawson, J., Farrel, T., Frost, G., Haynes, R., Heald, R., Lankshear, A., Mayfield, D., Waller, L., & Whittard, D. (2006). Performance of AAOmega: the AAT multi-purpose fiber-fed spectrograph. In I. S. McLean & M. Iye (Eds.), *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 6269 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* (pp. 62690G).

362. Shim, H., Im, M., Lee, H. M., Lee, M. G., Kim, S. J., Hwang, H. S., Hwang, N., Ko, J., Lee, J. C., Lim, S., Matsuhara, H., Seo, H., Wada, T., & Goto, T. (2011). Merging Galaxy Cluster A2255 in Mid-infrared. ApJ, 727(1), 14.

363. Shimwell, T. W., Hardcastle, M. J., Tasse, C., Best, P. N., Röttgering, H. J. A., Williams, W. L., Botteon, A., Drabent, A., Mechev, A., Shulevski, A., van Weeren, R. J., Bester, L., Brüggen, M., Brunetti, G., Callingham, J. R., Chyży, K. T., Conway, J. E., Dijkema, T. J., Duncan, K., de Gasperin, F., Hale, C. L., Haverkorn, M., Hugo, B., Jackson, N., Mevius, M., Miley, G. K., Morabito, L. K., Morganti, R., Offringa, A., Oonk, J. B. R., Rafferty, D., Sabater, J., Smith, D. J. B., Schwarz, D. J., Smirnov, O., O'Sullivan, S. P., Vedantham, H., White, G. J., Albert, J. G., Alegre, L., Asabere, B., Bacon, D. J., Bonafede, A., Bonnassieux, E., Brienza, M., Bilicki, M., Bonato, M., Calistro Rivera, G., Cassano, R., Cochrane, R., Croston, J. H., Cuciti, V., Dallacasa, D., Danezi, A., Dettmar, R. J., Di Gennaro, G., Edler, H. W., Enßlin, T. A., Emig, K. L., Franzen, T. M. O., García-Vergara, C., Grange, Y. G., Gürkan, G., Hajduk, M., Heald, G., Heesen, V., Hoang, D. N., Hoeft, M., Horellou, C., Iacobelli, M., Jamrozy, M., Jelić, V., Kondapally, R., Kukreti, P., Kunert-Bajraszewska, M., Magliocchetti, M., Mahatma, V., Małek, K., Mandal, S., Massaro, F., Meyer-Zhao, Z., Mingo, B., Mostert, R. I. J., Nair, D. G., Nakoneczny, S. J., Nikiel-Wroczyński, B., Orrú, E., Pajdosz-Śmierciak, U., Pasini, T., Prandoni, I., van Piggelen, H. E., Rajpurohit, K., Retana-Montenegro, E., Riseley, C. J., Rowlinson, A., Saxena, A., Schrijvers, C., Sweijen, F., Siewert, T. M., Timmerman, R., Vaccari, M., Vink, J., West, J. L., Wołowska, A., Zhang, X., & Zheng, J. (2022). The LOFAR Two-metre Sky Survey. V. Second data release. A&A, 659, A1.

364. Shimwell, T. W., Röttgering, H. J. A., Best, P. N., Williams, W. L., Dijkema, T. J., de Gasperin, F., Hardcastle, M. J., Heald, G. H., Hoang, D. N., Horneffer, A., Intema, H., Mahony, E. K., Mandal, S., Mechev, A. P., Morabito, L., Oonk, J. B. R., Rafferty, D., Retana-Montenegro, E., Sabater, J., Tasse, C., van Weeren, R. J., Brüggen, M., Brunetti, G., Chyży, K. T., Conway, J. E., Haverkorn, M., Jackson, N., Jarvis, M. J., McKean, J. P., Miley, G. K., Morganti, R., White, G. J., Wise, M. W., van Bemmel, I. M., Beck, R., Brienza, M., Bonafede, A., Calistro Rivera, G., Cassano, R., Clarke, A. O., Cseh, D., Deller, A., Drabent, A., van Driel, W., Engels, D., Falcke, H., Ferrari, C., Fröhlich, S., Garrett, M. A., Harwood, J. J., Heesen, V., Hoeft, M., Horellou, C., Israel, F. P., Kapińska, A. D., Kunert-Bajraszewska, M., McKay, D. J., Mohan, N. R., Orrú, E., Pizzo, R. F., Prandoni, I., Schwarz, D. J., Shulevski, A., Sipior, M., Smith, D. J. B., Sridhar, S. S., Steinmetz, M., Stroe, A., Varenius, E., van der Werf, P. P., Zensus, J. A., & Zwart, J. T. L. (2017). The LOFAR Two-metre Sky Survey. I. Survey description and preliminary data release. A&A, 598, A104.

365. Shimwell, T. W., Tasse, C., Hardcastle, M. J., Mechev, A. P., Williams, W. L., Best, P. N., Röttgering, H. J. A., Callingham, J. R., Dijkema, T. J., de Gasperin, F., Hoang, D. N., Hugo, B., Mirmont, M., Oonk, J. B. R., Prandoni, I., Rafferty, D., Sabater, J., Smirnov, O., van Weeren, R. J., White, G. J., Atemkeng, M., Bester, L., Bonnassieux, E., Brüggen, M., Brunetti, G., Chyży, K. T., Cochrane, R., Conway, J. E., Croston, J. H., Danezi, A., Duncan, K., Haverkorn, M., Heald, G. H., Iacobelli, M., Intema, H. T., Jackson, N., Jamrozy, M., Jarvis, M. J., Lakhoo, R., Mevius, M., Miley, G. K., Morabito, L., Morganti, R., Nisbet, D., Orrú, E., Perkins, S., Pizzo, R. F., Schrijvers, C., Smith, D. J. B., Vermeulen, R., Wise, M. W., Alegre, L., Bacon, D. J., van Bemmel, I. M., Beswick, R. J., Bonafede, A., Botteon, A., Bourke, S., Brienza, M., Calistro Rivera, G., Cassano, R., Clarke, A. O., Conselice, C. J., Dettmar, R. J., Drabent, A., Dumba, C., Emig, K. L., Enßlin, T. A., Ferrari, C., Garrett, M. A., Génova-Santos, R. T., Goyal, A., Gürkan, G., Hale, C., Harwood, J. J., Heesen, V., Hoeft, M., Horellou, C., Jackson, C., Kokotanekov, G., Kondapally, R., Kunert-Bajraszewska, M., Mahatma, V., Mahony, E. K., Mandal, S., McKean, J. P., Merloni, A., Mingo, B., Miskolczi, A., Mooney, S., Nikiel-Wroczyński, B., O'Sullivan, S. P., Quinn, J., Reich, W., Roskowiński, C., Rowlinson, A., Savini, F., Saxena, A., Schwarz, D. J., Shulevski, A., Sridhar, S. S., Stacey, H. R., Urquhart, S., van der Wiel, M. H. D., Varenius, E., Webster, B., & Wilber, A. (2019). The LOFAR Two-metre Sky Survey. II. First data release. A&A, 622, A1.

366. Simonte, M., Andernach, H., Brüggen, M., Schwarz, D. J., Prandoni, I., & Willis, A. G. (2022). Giant radio galaxies in the LOw-Frequency ARray Two-metre Sky Survey Boötes deep field. MNRAS, 515(2), 2032–2052.

367. Sims, P. H., Lentati, L., Alexander, P., & Carilli, C. L. (2016). Contamination of the Epoch of Reionization power spectrum in the presence of foregrounds. MNRAS, 462(3), 3069–3093.

368. Singal, J., Petrosian, V., Haider, J., & Malik, S. (2019). Luminosity-Luminosity Correlations in Flux-limited Multiwavelength Data. ApJ, 877(1), 63.

369. Siponen, M. I., Legrand, P., Widdrat, M., Jones, S. R., Zhang, W.-J., Chang, M. C. Y., Faivre, D., Arnoux, P., & Pignol, D. (2013). Structural insight into magnetochrome-mediated magnetite biomineralization. *Nature*, 502(7473), 681–684.

370. Smith, D. J. B., Best, P. N., Duncan, K. J., Hatch, N. A., Jarvis, M. J., Röttgering, H. J. A., Simpson, C. J., Stott, J. P., Cochrane, R. K., Coppin, K. E., Dannerbauer, H., Davis, T. A., Geach, J. E., Hale, C. L., Hardcastle, M. J., Hatfield, P. W., Houghton, R. C. W., Maddox, N., McGee, S. L., Morabito, L., Nisbet, D., Pandey-Pommier, M., Prandoni, I., Saxena, A., Shimwell, T. W., Tarr, M., van Bemmel, I., Verma, A., White, G. J., & Williams, W. L. (2016). The WEAVE-LOFAR Survey. In C. Reylé, J. Richard, L. Cambrésy, M. Deleuil, E. Pécontal, L. Tresse, & I. Vauglin (Eds.), *SF2A-2016: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics* (pp. 271–280).

371. Solovyov, D. I. & Verkhodanov, O. V. (2011). A search for faint giant radio galaxies in the NVSS survey. *Astrophysical Bulletin*, 66(4), 416–423.

372. Solovyov, D. I. & Verkhodanov, O. V. (2014). Radio and optical identification of giant radio galaxies from NVSS radio survey. *Astrophysical Bulletin*, 69(2), 141–159.

373. Soltan, A. (1982). Masses of quasars. MNRAS, 200, 115–122.

374. Spoelstra, T. A. T. (1983). The influence of ionospheric refraction on radio astronomy interferometry. A&A, 120, 313–321.

375. Springel, V., Pakmor, R., Pillepich, A., Weinberger, R., Nelson, D., Hernquist, L., Vogelsberger, M., Genel, S., Torrey, P., Marinacci, F., & Naiman, J. (2018). First results from the IllustrisTNG simulations: matter and galaxy clustering. MNRAS, 475(1), 676–698.

376. Springel, V., White, S. D. M., Jenkins, A., Frenk, C. S., Yoshida, N., Gao, L., Navarro, J., Thacker, R., Croton, D., Helly, J., Peacock, J. A., Cole, S., Thomas, P., Couchman, H., Evrard, A., Colberg, J., & Pearce, F. (2005). Simulations of the formation, evolution and clustering of galaxies and quasars. Nature, 435(7042), 629–636.

377. Strauss, M. A., Weinberg, D. H., Lupton, R. H., Narayanan, V. K., Annis, J., Bernardi, M., Blanton, M., Burles, S., Connolly, A. J., Dalcanton, J., Doi, M., Eisenstein, D., Frieman, J. A., Fukugita, M., Gunn, J. E., Ivezić, Ž., Kent, S., Kim, R. S. J., Knapp, G. R., Kron, R. G., Munn, J. A., Newberg, H. J., Nichol, R. C., Okamura, S., Quinn, T. R., Richmond, M. W., Schlegel, D. J., Shimasaku, K., SubbaRao, M., Szalay, A. S., Vanden Berk, D., Vogeley, M. S., Yanny, B., Yasuda, N., York, D. G., & Zehavi, I. (2002). Spectroscopic Target Selection in the Sloan Digital Sky Survey: The Main Galaxy Sample. AJ, 124(3), 1810–1824.

378. Stuardi, C., O'Sullivan, S. P., Bonafede, A., Brüggen, M., Dabhade, P., Horellou, C., Morganti, R., Carretti, E., Heald, G., Iacobelli, M., & Vacca, V. (2020). The LOFAR view of intergalactic magnetic fields with giant radio galaxies. A&A, 638, A48.

379. Subrahmanyan, R., Hunstead, R. W., Cox, N. L. J., & McIntyre, V. (2006). SGRS J0515-8100: A Fat-Double Giant Radio Galaxy. ApJ, 636(1), 172–180.

380. Subrahmanyan, R., Saripalli, L., & Hunstead, R. W. (1996). Morphologies in megaparsec-size powerful radio galaxies. MNRAS, 279(1), 257–274.

381. Subrahmanyan, R., Saripalli, L., Safouris, V., & Hunstead, R. W. (2008). On the Relationship between a Giant Radio

Galaxy MSH 05-22 and the Ambient Large-Scale Galaxy Structure. ApJ, 677(1), 63–78.

382. Subramanian, K. (2016). The origin, evolution and signatures of primordial magnetic fields. *Reports on Progress in Physics*, 79(7), 076901.

383. Sun, M., Voit, G. M., Donahue, M., Jones, C., Forman, W., & Vikhlinin, A. (2009). Chandra Studies of the X-Ray Gas Properties of Galaxy Groups. ApJ, 693(2), 1142–1172.

384. Sutherland, W. & Saunders, W. (1992). On the likelihood ratio for source identification. MNRAS, 259, 413–420.

385. Tamhane, P., Wadadekar, Y., Basu, A., Singh, V., Ishwara-Chandra, C. H., Beelen, A., & Sirothia, S. (2015). J021659-044920: a relic giant radio galaxy at z $\sim$ 1.3. MNRAS, 453(3), 2438–2446.

386. Tang, H., Scaife, A. M. M., Wong, O. I., Kapińska, A. D., Rudnick, L., Shabala, S. S., Seymour, N., & Norris, R. P. (2020). Radio Galaxy Zoo: new giant radio galaxies in the RGZ DR1 catalogue. MNRAS, 499(1), 68–76.

387. Tanimura, H., Aghanim, N., Douspis, M., Beelen, A., & Bonjean, V. (2019a). Detection of intercluster gas in superclusters using the thermal Sunyaev-Zel'dovich effect. A&A, 625, A67.

388. Tanimura, H., Aghanim, N., Douspis, M., & Malavasi, N. (2022). X-ray emission from cosmic web filaments in SRG/eROSITA data. *arXiv e-prints*, (pp. arXiv:2206.00084).

389. Tanimura, H., Aghanim, N., Kolodzig, A., Douspis, M., & Malavasi, N. (2020). First detection of stacked X-ray emission from cosmic web filaments. A&A, 643, L2.

390. Tanimura, H., Hinshaw, G., McCarthy, I. G., Van Waerbeke, L., Aghanim, N., Ma, Y.-Z., Mead, A., Hojjati, A., & Tröster, T. (2019b). A search for warm/hot gas filaments between pairs of SDSS Luminous Red Galaxies. MNRAS, 483(1), 223–234.

391. Tasse, C., Hugo, B., Mirmont, M., Smirnov, O., Atemkeng, M., Bester, L., Hardcastle, M. J., Lakhoo, R., Perkins, S., & Shimwell, T. (2018). Faceting for direction-dependent spectral deconvolution. A&A, 611, A87.

392. Tasse, C., Shimwell, T., Hardcastle, M. J., O'Sullivan, S. P., van Weeren, R., Best, P. N., Bester, L., Hugo, B., Smirnov, O., Sabater, J., Calistro-Rivera, G., de Gasperin, F., Morabito, L. K., Röttgering, H., Williams, W. L., Bonato, M., Bondi, M., Botteon, A., Brüggen, M., Brunetti, G., Chyży, K. T., Garrett, M. A., Gürkan, G., Jarvis, M. J., Kondapally, R., Mandal, S., Prandoni, I., Repetti, A., Retana-Montenegro, E., Schwarz, D. J., Shulevski, A., & Wiaux, Y. (2021). The LOFAR Two-meter Sky Survey: Deep Fields Data Release 1. I. Direction-dependent calibration and imaging. A&A, 648, A1.

393. Tempel, E., Tuvikene, T., Kipper, R., & Libeskind, N. I. (2017). Merging groups and clusters of galaxies from the SDSS data. The catalogue of groups and potentially merging systems. A&A, 602, A100.

394. The Astropy Collaboration, Price-Whelan, A. M. Price-Whelan, A. M., Sipőcz, B. M., Günther, H. M., Lim, P. L., Crawford, S. M., Conseil, S., Shupe, D. L., Craig, M. W., & Dencheva, N. (2018). The Astropy Project: Building an Open-science Project and Status of the v2.0 Core Package. AJ, 156, 123.

395. Timmerman, R., van Weeren, R. J., Callingham, J. R., Cotton, W. D., Perley, R., Morabito, L. K., Gizani, N. A. B., Bridle, A. H., O'Dea, C. P., Baum, S. A., Tremblay, G. R., Kharb, P., Kassim, N. E., Röttgering, H. J. A., Botteon, A., Sweijen, F., Tasse, C., Brüggen, M., Moldon, J., Shimwell, T., & Brunetti, G. (2022). Origin of the ring structures in Hercules A. Sub-arcsecond 144 MHz to 7 GHz observations. A&A, 658, A5.

396. Tjoa, J. N. K. Y., Mueller, M., & van der Tak, F. F. S. (2020). The subsurface habitability of small, icy exomoons. A&A, 636, A50.

397. Treumann, R. A. (2006). The electron-cyclotron maser for astrophysical application. A&A Rev., 13(4), 229–315.

398. Tsagas, C. G. (2007). Magnetic fields in conformally flat spacetimes. In *Journal of Physics Conference Series*, volume 68 of *Journal of Physics Conference Series* (pp. 012051).

399. Tsaprazi, E., Jasche, J., Lavaux, G., & Leclercq, F. (2023). Higher-order statistics of the large-scale structure from photometric redshifts. *arXiv e-prints*, (pp. arXiv:2301.03581).

400. Tully, R. B. (2015). Galaxy Groups: A 2MASS Catalog. AJ, 149(5), 171.

401. Tuominen, T., Nevalainen, J., Tempel, E., Kuutma, T., Wijers, N., Schaye, J., Heinämäki, P., Bonamente, M., & Ganeshaiah Veena, P. (2021). An EAGLE view of the missing baryons. A&A, 646, A156.

402. Turner, M. S. & Widrow, L. M. (1988). Inflation-produced, large-scale magnetic fields. *Phys. Rev. D*, 37, 2743–2754.

403. Turner, R. J. & Shabala, S. S. (2015). Energetics and Lifetimes of Local Radio Active Galactic Nuclei. ApJ, 806(1), 59.

404. van Cittert, P. H. (1934). Die Wahrscheinliche Schwingungsverteilung in Einer von Einer Lichtquelle Direkt Oder Mittels Einer Linse Beleuchteten Ebene. *Physica*, 1(1), 201–210.

405. van der Tol, S. (2009). *Bayesian estimation for ionospheric calibration in radio astronomy*. PhD thesis, TU Delft.

406. van der Tol, S., Veenboer, B., & Offringa, A. R. (2018). Image Domain Gridding: a fast method for convolutional resampling

of visibilities. A&A, 616, A27.

407. van Haarlem, M. P., Wise, M. W., Gunst, A. W., Heald, G., McKean, J. P., Hessels, J. W. T., de Bruyn, A. G., Nijboer, R., Swinbank, J., Fallows, R., Brentjens, M., Nelles, A., Beck, R., Falcke, H., Fender, R., Hörandel, J., Koopmans, L. V. E., Mann, G., Miley, G., Röttgering, H., Stappers, B. W., Wijers, R. A. M. J., Zaroubi, S., van den Akker, M., Alexov, A., Anderson, J., Anderson, K., van Ardenne, A., Arts, M., Asgekar, A., Avruch, I. M., Batejat, F., Bähren, L., Bell, M. E., Bell, M. R., van Bemmel, I., Bennema, P., Bentum, M. J., Bernardi, G., Best, P., Bîrzan, L., Bonafede, A., Boonstra, A. J., Braun, R., Bregman, J., Breitling, F., van de Brink, R. H., Broderick, J., Broekema, P. C., Brouw, W. N., Brüggen, M., Butcher, H. R., van Cappellen, W., Ciardi, B., Coenen, T., Conway, J., Coolen, A., Corstanje, A., Damstra, S., Davies, O., Deller, A. T., Dettmar, R. J., van Diepen, G., Dijkstra, K., Donker, P., Doorduin, A., Dromer, J., Drost, M., van Duin, A., Eislöffel, J., van Enst, J., Ferrari, C., Frieswijk, W., Gankema, H., Garrett, M. A., de Gasperin, F., Gerbers, M., de Geus, E., Grießmeier, J. M., Grit, T., Gruppen, P., Hamaker, J. P., Hassall, T., Hoeft, M., Holties, H. A., Horneffer, A., van der Horst, A., van Houwelingen, A., Huijgen, A., Iacobelli, M., Intema, H., Jackson, N., Jelic, V., de Jong, A., Juette, E., Kant, D., Karastergiou, A., Koers, A., Kollen, H., Kondratiev, V. I., Kooistra, E., Koopman, Y., Koster, A., Kuniyoshi, M., Kramer, M., Kuper, G., Lambropoulos, P., Law, C., van Leeuwen, J., Lemaitre, J., Loose, M., Maat, P., Macario, G., Markoff, S., Masters, J., McFadden, R. A., McKay-Bukowski, D., Meijering, H., Meulman, H., Mevius, M., Middelberg, E., Millenaar, R., Miller-Jones, J. C. A., Mohan, R. N., Mol, J. D., Morawietz, J., Morganti, R., Mulcahy, D. D., Mulder, E., Munk, H., Nieuwenhuis, L., van Nieuwpoort, R., Noordam, J. E., Norden, M., Noutsos, A., Offringa, A. R., Olofsson, H., Omar, A., Orrú, E., Overeem, R., Paas, H., Pandey-Pommier, M., Pandey, V. N., Pizzo, R., Polatidis, A., Rafferty, D., Rawlings, S., Reich, W., de Reijer, J. P., Reitsma, J., Renting, G. A., Riemers, P., Rol, E., Romein, J. W., Roosjen, J., Ruiter, M., Scaife, A., van der Schaaf, K., Scheers, B., Schellart, P., Schoenmakers, A., Schoonderbeek, G., Serylak, M., Shulevski, A., Sluman, J., Smirnov, O., Sobey, C., Spreeuw, H., Steinmetz, M., Sterks, C. G. M., Stiepel, H. J., Stuurwold, K., Tagger, M., Tang, Y., Tasse, C., Thomas, I., Thoudam, S., Toribio, M. C., van der Tol, B., Usov, O., van Veelen, M., van der Veen, A. J., ter Veen, S., Verbiest, J. P. W., Vermeulen, R., Vermaas, N., Vocks, C., Vogt, C., de Vos, M., van der Wal, E., van Weeren, R., Weggemans, H., Weltevrede, P., White, S., Wijnholds, S. J., Wilhelmsson, T., Wucknitz, O., Yatawatta, S., Zarka, P., Zensus, A., & van Zwieten, J. (2013). LOFAR: The LOw-Frequency ARray. A&A, 556, A2.

408. van Leeuwen, H. J. (1919). *Vraagstukken uit de electronentheorie van het magnetisme*. Leiden: Eduard IJdo.

409. van Weeren, R. J., de Gasperin, F., Akamatsu, H., Brüggen, M., Feretti, L., Kang, H., Stroe, A., & Zandanel, F. (2019). Diffuse Radio Emission from Galaxy Clusters. Space Sci. Rev., 215(1), 16.

410. van Weeren, R. J., Röttgering, H. J. A., Intema, H. T., Rudnick, L., Brüggen, M., Hoeft, M., & Oonk, J. B. R. (2012). The "toothbrush-relic": evidence for a coherent linear 2-Mpc scale shock wave in a massive merging galaxy cluster? A&A, 546, A124.

411. van Weeren, R. J., Shimwell, T. W., Botteon, A., Brunetti, G., Brüggen, M., Boxelaar, J. M., Cassano, R., Di Gennaro, G., Andrade-Santos, F., Bonnassieux, E., Bonafede, A., Cuciti, V., Dallacasa, D., de Gasperin, F., Gastaldello, F., Hardcastle, M. J., Hoeft, M., Kraft, R. P., Mandal, S., Rossetti, M., Röttgering, H. J. A., Tasse, C., & Wilber, A. G. (2021). LOFAR observations of galaxy clusters in HETDEX. Extraction and self-calibration of individual LOFAR targets. A&A, 651, A115.

412. van Weeren, R. J., Williams, W. L., Hardcastle, M. J., Shimwell, T. W., Rafferty, D. A., Sabater, J., Heald, G., Sridhar, S. S., Dijkema, T. J., Brunetti, G., Brüggen, M., Andrade-Santos, F., Ogrean, G. A., Röttgering, H. J. A., Dawson, W. A., Forman, W. R., de Gasperin, F., Jones, C., Miley, G. K., Rudnick, L., Sarazin, C. L., Bonafede, A., Best, P. N., Bîrzan, L., Cassano, R., Chyży, K. T., Croston, J. H., Ensslin, T., Ferrari, C., Hoeft, M., Horellou, C., Jarvis, M. J., Kraft, R. P., Mevius, M., Intema, H. T., Murray, S. S., Orrú, E., Pizzo, R., Simionescu, A., Stroe, A., van der Tol, S., & White, G. J. (2016). LOFAR Facet Calibration. ApJS, 223, 2.

413. Vazza, F., Brüggen, M., Gheller, C., Hackstein, S., Wittor, D., & Hinz, P. M. (2017). Simulations of extragalactic magnetic fields and of their observables. *Classical and Quantum Gravity*, 34(23), 234001.

414. Vazza, F., Brunetti, G., & Gheller, C. (2009). Shock waves in Eulerian cosmological simulations: main properties and acceleration of cosmic rays. MNRAS, 395(3), 1333–1354.

415. Vazza, F., Dolag, K., Ryu, D., Brunetti, G., Gheller, C., Kang, H., & Pfrommer, C. (2011). A comparison of cosmological codes: properties of thermal gas and shock waves in large-scale structures. MNRAS, 418(2), 960–985.

416. Vazza, F., Ettori, S., Roncarelli, M., Angelinelli, M., Brüggen, M., & Gheller, C. (2019). Detecting shocked intergalactic gas with X-ray and radio observations. A&A, 627, A5.

417. Vazza, F. & Feletti, A. (2020). The quantitative comparison between the neuronal network and the cosmic web. *Frontiers in Physics*, 8, 491.

418. Vazza, F., Ferrari, C., Brüggen, M., Bonafede, A., Gheller, C., & Wang, P. (2015). Forecasts for the detection of the magnetised cosmic web from cosmological simulations. A&A, 580, A119.

419. Vazza, F., Locatelli, N., Rajpurohit, K., Banfi, S., Domínguez-Fernández, P., Wittor, D., Angelinelli, M., Inchingolo, G.,

Brienza, M., Hackstein, S., Dallacasa, D., Gheller, C., Brüggen, M., Brunetti, G., Bonafede, A., Ettori, S., Stuardi, C., Paoletti, D., & Finelli, F. (2021a). Magnetogenesis and the Cosmic Web: A Joint Challenge for Radio Observations and Numerical Simulations. *Galaxies*, 9(4), 109.

420. Vazza, F., Paoletti, D., Banfi, S., Finelli, F., Gheller, C., O'Sullivan, S. P., & Brüggen, M. (2021b). Simulations and observational tests of primordial magnetic fields from Cosmic Microwave Background constraints. MNRAS, 500(4), 5350–5368.

421. Vedantham, H. K. & Koopmans, L. V. E. (2015). Scintillation noise in widefield radio interferometry. MNRAS, 453, 925–938.

422. Vernstrom, T., Gaensler, B. M., Brown, S., Lenc, E., & Norris, R. P. (2017). Low-frequency radio constraints on the synchrotron cosmic web. MNRAS, 467(4), 4914–4936.

423. Vernstrom, T., Gaensler, B. M., Rudnick, L., & Andernach, H. (2019). Differences in Faraday Rotation between Adjacent Extragalactic Radio Sources as a Probe of Cosmic Magnetic Fields. ApJ, 878(2), 92.

424. Vernstrom, T., Heald, G., Vazza, F., Galvin, T. J., West, J. L., Locatelli, N., Fornengo, N., & Pinetti, E. (2021). Discovery of magnetic fields along stacked cosmic filaments as revealed by radio and X-ray emission. MNRAS, 505(3), 4178–4196.

425. Vernstrom, T., West, J., Vazza, F., Wittor, D., Riseley, C. J., & Heald, G. (2023). Polarized accretion shocks from the cosmic web. *Science Advances*, 9(7), eade7233.

426. Voges, W., Aschenbach, B., Boller, T., Bräuninger, H., Briel, U., Burkert, W., Dennerl, K., Englhauser, J., Gruber, R., Haberl, F., Hartner, G., Hasinger, G., Kürster, M., Pfeffermann, E., Pietsch, W., Predehl, P., Rosso, C., Schmitt, J. H. M. M., Trümper, J., & Zimmermann, H. U. (1999). The ROSAT all-sky survey bright source catalogue. A&A, 349, 389–405.

427. Wandelt, B. D., Larson, D. L., & Lakshminarayanan, A. (2004). Global, exact cosmic microwave background data analysis using Gibbs sampling. Phys. Rev. D, 70(8), 083511.

428. Wang, H., Mo, H. J., Yang, X., Jing, Y. P., & Lin, W. P. (2014). ELUCID—Exploring the Local Universe with the Reconstructed Initial Density Field. I. Hamiltonian Markov Chain Monte Carlo Method with Particle Mesh Dynamics. ApJ, 794(1), 94.

429. Wang, H., Mo, H. J., Yang, X., & van den Bosch, F. C. (2013). Reconstructing the Initial Density Field of the Local Universe: Methods and Tests with Mock Catalogs. ApJ, 772(1), 63.

430. Wang, S., Liu, J., Qiu, Y., Bai, Y., Yang, H., Guo, J., & Zhang, P. (2016). CHANDRA ACIS Survey of X-Ray Point Sources: The Source Catalog. ApJS, 224(2), 40.

431. Wang, W., Dai, J., Chen, Z., Huang, Z., Li, Z., Zhu, X., Hu, X., Lu, T., Lu, L., Li, H., et al. (2022). Internimage: Exploring large-scale vision foundation models with deformable convolutions. *arXiv preprint arXiv:2211.05778*.

432. Ward, S. R., Harrison, C. M., Costa, T., & Mainieri, V. (2022). Cosmological simulations predict that AGN preferentially live in gas-rich, star-forming galaxies despite effective feedback. MNRAS, 514(2), 2936–2957.

433. Warner, T. (2012). *Synthesis, Properties and Mineralogy of Important Inorganic Materials*. Wiley.

434. Wasilewski, P. & Kletetschka, G. (1999). Lodestone: Nature's only permanent magnet — What it is and how it gets charged. *Geophysical Research Letters*, 26(15), 2275–2278.

435. Weiss, Y. & Freeman, W. T. (2001). Correctness of belief propagation in gaussian graphical models of arbitrary topology. *Neural Computation*, 13(10), 2173–2200.

436. Wen, Z. L. & Han, J. L. (2015). Calibration of the Optical Mass Proxy for Clusters of Galaxies and an Update of the WHL12 Cluster Catalog. ApJ, 807(2), 178.

437. Wen, Z. L., Han, J. L., & Liu, F. S. (2012). A Catalog of 132,684 Clusters of Galaxies Identified from Sloan Digital Sky Survey III. ApJS, 199(2), 34.

438. Werner, N., Simionescu, A., Million, E. T., Allen, S. W., Nulsen, P. E. J., von der Linden, A., Hansen, S. M., Böhringer, H., Churazov, E., Fabian, A. C., Forman, W. R., Jones, C., Sanders, J. S., & Taylor, G. B. (2010). Feedback under the microscope-II. Heating, gas uplift and mixing in the nearest cluster core. MNRAS, 407(4), 2063–2074.

439. Wilcox, C. H. (1962). Wave propagation in a random medium (lev a. chernov). *SIAM Review*, 4(1), 55–55.

440. Williams, W. L., Calistro Rivera, G., Best, P. N., Hardcastle, M. J., Röttgering, H. J. A., Duncan, K. J., de Gasperin, F., Jarvis, M. J., Miley, G. K., Mahony, E. K., Morabito, L. K., Nisbet, D. M., Prandoni, I., Smith, D. J. B., Tasse, C., & White, G. J. (2018). LOFAR-Boötes: properties of high- and low-excitation radio galaxies at $0.5 < z < 2.0$. MNRAS, 475(3), 3429–3452.

441. Williams, W. L., Hardcastle, M. J., Best, P. N., Sabater, J., Croston, J. H., Duncan, K. J., Shimwell, T. W., Röttgering, H. J. A., Nisbet, D., Gürkan, G., Alegre, L., Cochrane, R. K., Goyal, A., Hale, C. L., Jackson, N., Jamrozy, M., Kondapally, R., Kunert-Bajraszewska, M., Mahatma, V. H., Mingo, B., Morabito, L. K., Prandoni, I., Roskowinski, C., Shulevski, A., Smith, D. J. B., Tasse, C., Urquhart, S., Webster, B., White, G. J., Beswick, R. J., Callingham, J. R., Chyży, K. T., de Gasperin,

F., Harwood, J. J., Hoeft, M., Iacobelli, M., McKean, J. P., Mechev, A. P., Miley, G. K., Schwarz, D. J., & van Weeren, R. J. (2019). The LOFAR Two-metre Sky Survey. III. First data release: Optical/infrared identifications and value-added catalogue. A&A, 622, A2.

442. Willis, A. G., Strom, R. G., & Wilson, A. S. (1974). 3C236, DA240; the largest radio sources known. Nature, 250(5468), 625–630.

443. Willott, C. J., Rawlings, S., Blundell, K. M., & Lacy, M. (1999). The emission line-radio correlation for radio sources using the 7C Redshift Survey. MNRAS, 309(4), 1017–1033.

444. Wiltschko, R. & Wiltschko, W. (2013). The magnetite-based receptors in the beak of birds and their role in avian navigation. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol*, 199(2), 89–98.

445. Wolf, E. (1969). Three-dimensional structure determination of semi-transparent objects from holographic data. *Optics Communications*, 1, 153–156.

446. Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., Ressler, M. E., Cutri, R. M., Jarrett, T., Kirkpatrick, J. D., Padgett, D., McMillan, R. S., Skrutskie, M., Stanford, S. A., Cohen, M., Walker, R. G., Mather, J. C., Leisawitz, D., Gautier, Thomas N., I., McLean, I., Benford, D., Lonsdale, C. J., Blain, A., Mendez, B., Irace, W. R., Duval, V., Liu, F., Royer, D., Heinrichsen, I., Howard, J., Shannon, M., Kendall, M., Walsh, A. L., Larsen, M., Cardon, J. G., Schick, S., Schwalm, M., Abid, M., Fabinsky, B., Naes, L., & Tsai, C.-W. (2010). The Wide-field Infrared Survey Explorer (WISE): Mission Description and Initial On-orbit Performance. AJ, 140(6), 1868–1881.

447. Wu, L., Wang, D., & A. Evans, J. (2019). Large teams develop and small teams disrupt science and technology. *Nature*, 566, 1.

448. Xu, H., O'Shea, B. W., Collins, D. C., Norman, M. L., Li, H., & Li, S. (2008). The Biermann Battery in Cosmological MHD Simulations of Population III Star Formation. ApJ, 688(2), L57.

449. Xu, R., Spitkovsky, A., & Caprioli, D. (2020). Electron Acceleration in One-dimensional Nonrelativistic Quasi-perpendicular Collisionless Shocks. ApJ, 897(2), L41.

450. Yahil, A. & Vidal, N. V. (1977). The Velocity Distribution of Galaxies in Clusters. ApJ, 214, 347–350.

451. Yang, H. Y. K., Gaspari, M., & Marlow, C. (2019a). The Impact of Radio AGN Bubble Composition on the Dynamics and Thermal Balance of the Intracluster Medium. ApJ, 871(1), 6.

452. Yang, Y., Wan, M., Matthaeus, W. H., Shi, Y., Parashar, T. N., Lu, Q., & Chen, S. (2019b). Role of magnetic field curvature in magnetohydrodynamic turbulence. *Physics of Plasmas*, 26(7), 072306.

453. Yeh, K. C. (1962). Propagation of spherical waves through an ionosphere containing anisotropic irregularities. *Journal of Research of National Bureau of Standards*, 5, 621–636.

454. Yeh, K. C. & Swenson, Jr., G. W. (1959). The scintillation of radio signals from satellites. J. Geophys. Res., 64, 2281–2286.

455. York, D. G., Adelman, J., Anderson, John E., J., Anderson, S. F., Annis, J., Bahcall, N. A., Bakken, J. A., Barkhouser, R., Bastian, S., Berman, E., Boroski, W. N., Bracker, S., Briegel, C., Briggs, J. W., Brinkmann, J., Brunner, R., Burles, S., Carey, L., Carr, M. A., Castander, F. J., Chen, B., Colestock, P. L., Connolly, A. J., Crocker, J. H., Csabai, I., Czarapata, P. C., Davis, J. E., Doi, M., Dombeck, T., Eisenstein, D., Ellman, N., Elms, B. R., Evans, M. L., Fan, X., Federwitz, G. R., Fiscelli, L., Friedman, S., Frieman, J. A., Fukugita, M., Gillespie, B., Gunn, J. E., Gurbani, V. K., de Haas, E., Haldeman, M., Harris, F. H., Hayes, J., Heckman, T. M., Hennessy, G. S., Hindsley, R. B., Holm, S., Holmgren, D. J., Huang, C.-h., Hull, C., Husby, D., Ichikawa, S.-I., Ichikawa, T., Ivezić, Ž., Kent, S., Kim, R. S. J., Kinney, E., Klaene, M., Kleinman, A. N., Kleinman, S., Knapp, G. R., Korienek, J., Kron, R. G., Kunszt, P. Z., Lamb, D. Q., Lee, B., Leger, R. F., Limmongkol, S., Lindenmeyer, C., Long, D. C., Loomis, C., Loveday, J., Lucinio, R., Lupton, R. H., MacKinnon, B., Mannery, E. J., Mantsch, P. M., Margon, B., McGehee, P., McKay, T. A., Meiksin, A., Merelli, A., Monet, D. G., Munn, J. A., Narayanan, V. K., Nash, T., Neilsen, E., Neswold, R., Newberg, H. J., Nichol, R. C., Nicinski, T., Nonino, M., Okada, N., Okamura, S., Ostriker, J. P., Owen, R., Pauls, A. G., Peoples, J., Peterson, R. L., Petravick, D., Pier, J. R., Pope, A., Pordes, R., Prosapio, A., Rechenmacher, R., Quinn, T. R., Richards, G. T., Richmond, M. W., Rivetta, C. H., Rockosi, C. M., Ruthmansdorfer, K., Sand ford, D., Schlegel, D. J., Schneider, D. P., Sekiguchi, M., Sergey, G., Shimasaku, K., Siegmund, W. A., Smee, S., Smith, J. A., Snedden, S., Stone, R., Stoughton, C., Strauss, M. A., Stubbs, C., SubbaRao, M., Szalay, A. S., Szapudi, I., Szokoly, G. P., Thakar, A. R., Tremonti, C., Tucker, D. L., Uomoto, A., Vanden Berk, D., Vogeley, M. S., Waddell, P., Wang, S.-i., Watanabe, M., Weinberg, D. H., Yanny, B., Yasuda, N., & SDSS Collaboration (2000). The Sloan Digital Sky Survey: Technical Summary. AJ, 120(3), 1579–1587.

456. Zel'dovich, Y. B. (1970). Reprint of 1970A&A.....5...84Z. Gravitational instability: an approximate theory for large density perturbations. A&A, 500, 13–18.

457. Zernike, F. (1938). The concept of degree of coherence and its application to optical problems. *Physica*, 5(8), 785–795.

458. Zhang, H., Li, F., Liu, S., Zhang, L., Su, H., Zhu, J., Ni, L. M., & Shum, H.-Y. (2022). DINO: DETR with Improved

DeNoising Anchor Boxes for End-to-End Object Detection. *arXiv e-prints*, (pp. arXiv:2203.03605).

459. Zheng, H., Tegmark, M., Dillon, J. S., Kim, D. A., Liu, A., Neben, A. R., Jonas, J., Reich, P., & Reich, W. (2017). An improved model of diffuse galactic radio emission from 10 MHz to 5 THz. MNRAS, 464(3), 3486–3497.

460. Zou, H., Gao, J., Xu, X., Zhou, X., Ma, J., Zhou, Z., Zhang, T., Nie, J., Wang, J., & Xue, S. (2021). Galaxy Clusters from the DESI Legacy Imaging Surveys. I. Cluster Detection. ApJS, 253(2), 56.

461. Zou, H., Sui, J., Xue, S., Zhou, X., Ma, J., Zhou, Z., Nie, J., Zhang, T., Feng, L., Shen, Z., & Wang, J. (2022). Photometric Redshifts and Galaxy Clusters for DES DR2, DESI DR9, and HSC-SSP PDR3 Data. *Research in Astronomy and Astrophysics*, 22(6), 065001.