



Universiteit
Leiden
The Netherlands

Dancing with the stars

Albert, J.G.

Citation

Albert, J. G. (2020, October 28). *Dancing with the stars*. Retrieved from <https://hdl.handle.net/1887/137988>

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/1887/137988>

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

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/137988> holds various files of this Leiden University dissertation.

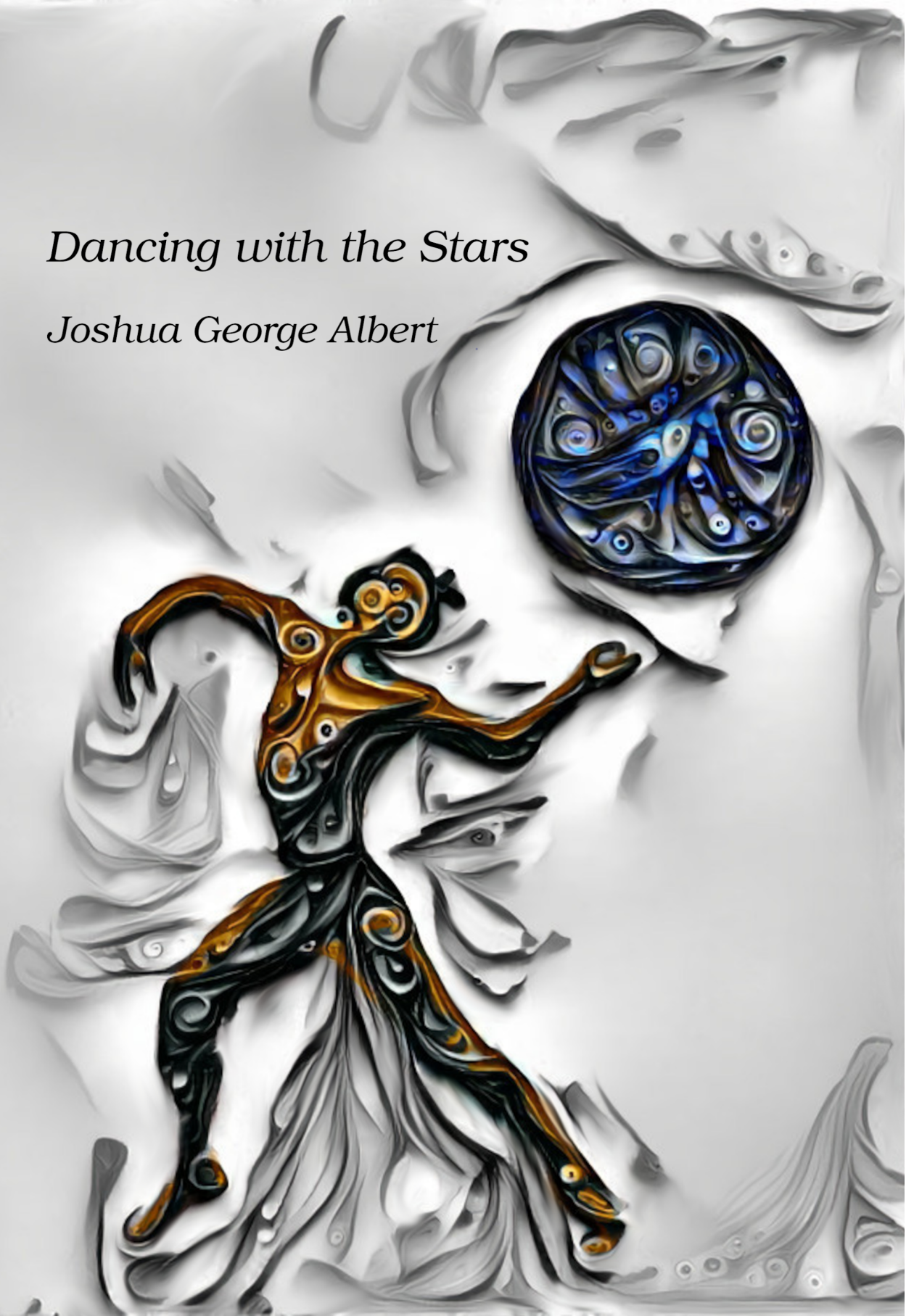
Author: Albert, J.G.

Title: Dancing with the stars

Issue Date: 2020-10-28

Dancing with the Stars

Joshua George Albert



Dancing with the Stars

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 28 oktober 2020
klokke 13:45 uur

door

Joshua George Albert
geboren te London, Ontario, Canada
in 1990

Promotor: Prof.dr. H.J.A. Röttgering (Sterrewacht Leiden)
Co-Promotor: Dr. R.J. van Weeren (Sterrewacht Leiden)

Promotiecommissie:

Voorzitter: Prof.dr. H.J. van Langevelde (Sterrewacht Leiden)
Secretaris: Prof.dr. C.U. Keller (Sterrewacht Leiden)
Overige leden: Prof.dr. K.J. Batenburg (Mathematisch Instituut Leiden)
Prof.dr. L.V.E. Koopmans (Rijksuniversiteit Groningen)
Prof.dr. O. Smirnov (Rhodes University)
Dr.ing. H.T. Intema (Sterrewacht Leiden)
Dr. C. Tasse (L'Observatoire de Paris)
Dr. M. Mevius (Stichting ASTRON, Dwingeloo)

ISBN: 978-94-6332-685-8

© 2020 Joshua George Albert

Omslagontwerp: Joshua George Albert

Hoofdstuk illustraties: Ainsley Bernice Albert

Dit werk is gelicenseerd onder een Creative Commons Naamsvermelding-GelijkDelen 3.0 Unported licentie. Ga naar <http://creativecommons.org/licenses/by-sa/3.0/> om een kopie van de licentie te kunnen lezen.

Dit proefschrift werd ondersteund door het European Research Council Advanced Investigator programme NewClusters 321271.

For Alina, my twinkling star.

Contents

Preface	vii
Introduction	1
Radio emission mechanisms	1
Diffuse emission in galaxy clusters	4
Radio interferometry	5
History	5
Technical introduction to scattering and radio interferometry	8
Next-generation radio interferometry	17
About this thesis	19
Future directions	20
1 Complex diffuse emission in PLCK G004.5-19.5	23
1.1 Introduction	25
1.2 Observations and data reduction	26
1.2.1 GMRT observations of PLCK G004.5-19.5	26
1.2.2 <i>XMM-Newton</i> reduction	28
1.2.3 Spectroscopic data	28
1.3 Results and discussion	29
1.3.1 X-ray structure	29
1.3.2 Radio emission	31
1.3.3 Nature of the diffuse emission	35
1.4 Conclusion	36
2 Probabilistic ionospheric calibration	37
2.1 Introduction	39
2.2 Ionospheric effects on interferometric visibilities	40
2.3 Probabilistic relation between FED and ΔTEC	42
2.4 Method	45
2.4.1 Data generation	46
2.4.2 Competitor models	47
2.4.3 Model comparison	48
2.5 Results	52
2.6 Discussion	55
2.6.1 Model selection bias	55

2.6.2	Implicit tomography	55
2.6.3	Temporal differential TEC correlations	56
2.6.4	Structure function turnover and anisotropic diffractive scale	56
2.6.5	Low-accuracy numerical integration	58
2.7	Conclusion	58
3	Probabilistic ionospheric calibration for LOFAR-HBA	61
3.1	Introduction	63
3.2	Doubly differential phase screens	64
3.3	Methods	68
3.3.1	Subtract and solve	68
3.3.2	Smooth and slow-resolve	71
3.3.3	Measure DDTEC	73
3.3.4	Infer DDTEC screen	78
3.3.5	Image screen	82
3.4	Results	83
3.5	Discussion	84
3.6	Conclusion	85
4	Feasibility of ionospheric screens for LoTSS	91
4.1	Introduction	93
4.2	Method	94
4.2.1	An improved outlier detection	96
4.3	Image improvements and ionospheric robustness	101
4.4	A non-ionospheric systematic and correction	105
4.5	Information in DDTEC screens	111
4.6	Conclusion	113
	Appendix A Derivation of tomographic equivalence	135
	Appendix B ΔTEC variance function	137
	Appendix C Factoring commutative DI dependence from the RIME	141
	Appendix D Recursive Bayesian estimation	143
	Appendix E Jones scalar variational expectation	147
	Appendix F Derivation of the M-step	151
	English Summary	155
	Nederlandse Samenvatting	159
	Curriculum Vitae	163
	Acknowledgements	167



"You're a smart guy, but tell me this. Why is it that nobody talks about what was there before all this?" asked my grand father waving his old weathered hands around at the gently whispering trees. We sat in two chairs perched on top of an old green hill that looked out over the water of a place called Camp. An unlit cigar adorned my grand father's finger tips, and I leaned over to help shield it from the wind as he lit it.

I answered back, "What do you mean by *this*, exactly? Do you mean the space that we live in, and perhaps also the time that we occupy?"

"All of it," he replied, "If we're to believe what people say, the entire Universe started with a Big Bang, and has been expanding ever since. I mean, if you imagine the Universe is expanding, shooting out in all directions, then it must be expanding into something. What is that?" I paused for a moment. I loved these conversations with my grand father very much. Even in his eighth decade, he always had a way of stepping far outside his immediate surroundings into his imagination. Furthermore, he is a very stubborn man, and this combination – imagination and steadfast persistence – is a wonderful mindset to have when it comes to discussing the finer points of nature unknown.

"Ah, you mean if you could magically be at the edge of the Universe, would there be some sort of shock wave as the Universe expanded into something? And then, what would that something be?" He nodded. I attempted to light my own cigar in the gentle breeze, and gratefully grunted a thanks as my grand father leaned over to shield the wind. "Well, unfortunately we are stuck here on Earth, with only our minds to wonder. But let's say we were there during the Big Bang, and were riding one of the first packets of energy – let's call it a photon – as it seemingly raced out into the void. Since we were riding the first photon, we wouldn't be able to see anything in front of us, and we might shout 'watch out for the void!' as we feared collision with something we could not see. If this little photon just kept going, with us on its back, then we would always seem to be flying into the void. So I fear, whatever it is the Universe expands into, we could never describe it with things that can be seen."

"Very interesting," said my grand father, "You see why no one ever talks about this, because no one knows, but you have a good way of putting it into words. Surely there must be something we could measure of this void,

like a temperature perhaps?"

"You know what temperature is, right? It's just a measure of how fast molecules are bouncing around. If there is no matter – or energy – in the void, then it's not possible to measure temperature." I smiled at my grand father's face.

"Ah, your right. Maybe one day we'll know, and I think you're the right guy for it." The next few minutes passed without words between us, although it was by no means silent. Around us whispered the leaves, and below us the roll of the waves against the shore concealed their meanings. As it were, we'd return to this topic again some time later.

These are the sort of memories, going back as far as I can recall, that I cherish. Not only of conversations with my grand father, but with others, and myself.

It is through the childhood that a person first gains their sight of the world – that magical thing we call life. We naturally accept the guidance of those around us, who walk with us through the difficulties ahead. It seems such a messy and beautiful dance, and all lifeforms partake. Interestingly, we messy creatures have devised scientific inquiry as a tool for surveying that which lays before us. To us, we look not into a void, but into a rich landscape. And yet, we too emerge from a singularity at birth and dance forth into the unknown, so some prudence to caution is advisable.

Can we use our senses and reasoning, that which we investigate the world with, to investigate existence itself? I would argue the only such hope we have of understanding existence through reasoning is that it's a tautology, 'cogito ergo sum' in the words of René Descartes. However, outside the walls of the mind we quickly lose objectivity, if it ever existed, and is properly put by Meister Eckhart von Hochheim (ca. 1260 – 1328),

If reason presents certain actions as likely to improve your condition but your conscience does not approve, then leave them undone. Reason never grasps the world in its entirety and the means it dictates to achieving its limited aim will ultimately, and in some inscrutable way, only cause damage.

If in this life we always had a mirror in front of us in which we could see things at a glance, grasp everything in one image, acting and knowing would not cause us any problem. But, since in our viewing we must turn from one thing to another we cannot concentrate on one without obstructing the other.

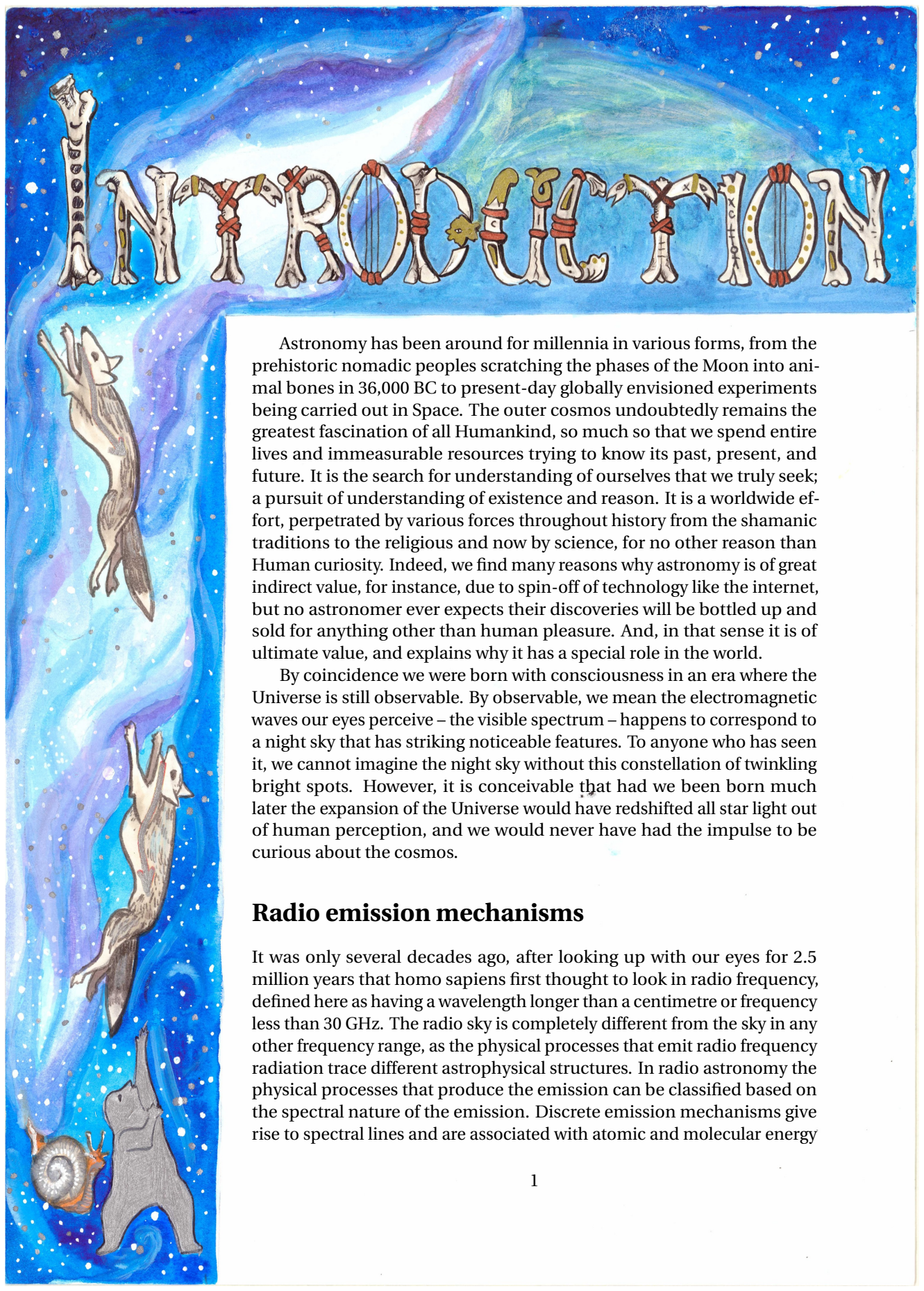
In a sense, Meister Eckhart paints a type of uncertainty principle between reason and conscience, acting and knowing, epistemology and ontology. To turn toward one you must turn from the other. We may juxtapose the bumbling child's discovery of life with the ideal scientist, into whom the child yearns to become. Regarding the end state of this transformation, science may be seen as immaculate and homogeneous, and is often given titles of truth. Yet, look closer and the messy dance of the child never truly stops, like Zeno's tortoise never quite reaching the finish line. There is never time to pause and look around and feel certainty, 'a masterpiece is never finished, only abandoned' in the words of Leonardo da Vinci. Thus, I must insist, science *is* an interpretive dance, where no one may violate Meister Eckhart's uncertainty principle.

Why are we dancing with the stars? Well, I'm afraid *the medium is the message*, and that's all there is to it. Twentieth century philosopher, Michael Serres, metaphysically discusses two rats eating rich left overs on a Persian rug in the Mayor's house,

The two companions scurry off when they hear a noise at the door. It was only a noise, but it was also a message, a bit of information producing panic: an interruption, a corruption, a rupture of information. Was the noise really a message? Wasn't it, rather, static, a parasite? A parasite who has the last word, who produces disorder and who generates a different order. Let's go to the country where we eat only soup, but quietly and without interruption.

To parasite means to eat next to. The noise partakes in the meal next to the rats and causes panic. In the country quietness is the parasite.

We astronomers look up at night on a clear sky and watch the stars twinkle and consider the twinkle a noise, a corruption, a rupture of information. Is the twinkle really a message? The twinkle is a movement of light, spurred by Earth's atmosphere, as natural a thing as ever existed. We isolate the movement, and, following it, we dance our telescopes until we see a stationary shape. Oh, but such a strange shape! The struts of the telescope leave these arms of light seemingly coming out of the star, another corruption. Are these arms of light really a message? To be sure they are not, but merely a parasite. And, we may go on and on in such a way always trying to interpret what is beyond the noise without knowing. And, that is the essence of this work: we perform an interpretive dance with twinkling lights, with an impeccably powerful tool of human inquiry, trying not to confuse the message for the medium.



INTRODUCTION

Astronomy has been around for millennia in various forms, from the prehistoric nomadic peoples scratching the phases of the Moon into animal bones in 36,000 BC to present-day globally envisioned experiments being carried out in Space. The outer cosmos undoubtedly remains the greatest fascination of all Humankind, so much so that we spend entire lives and immeasurable resources trying to know its past, present, and future. It is the search for understanding of ourselves that we truly seek; a pursuit of understanding of existence and reason. It is a worldwide effort, perpetrated by various forces throughout history from the shamanic traditions to the religious and now by science, for no other reason than Human curiosity. Indeed, we find many reasons why astronomy is of great indirect value, for instance, due to spin-off of technology like the internet, but no astronomer ever expects their discoveries will be bottled up and sold for anything other than human pleasure. And, in that sense it is of ultimate value, and explains why it has a special role in the world.

By coincidence we were born with consciousness in an era where the Universe is still observable. By observable, we mean the electromagnetic waves our eyes perceive – the visible spectrum – happens to correspond to a night sky that has striking noticeable features. To anyone who has seen it, we cannot imagine the night sky without this constellation of twinkling bright spots. However, it is conceivable that had we been born much later the expansion of the Universe would have redshifted all star light out of human perception, and we would never have had the impulse to be curious about the cosmos.

Radio emission mechanisms

It was only several decades ago, after looking up with our eyes for 2.5 million years that homo sapiens first thought to look in radio frequency, defined here as having a wavelength longer than a centimetre or frequency less than 30 GHz. The radio sky is completely different from the sky in any other frequency range, as the physical processes that emit radio frequency radiation trace different astrophysical structures. In radio astronomy the physical processes that produce the emission can be classified based on the spectral nature of the emission. Discrete emission mechanisms give rise to spectral lines and are associated with atomic and molecular energy

transitions, while continuum emission mechanisms give rise to a broad band spectra.

In the realm of discrete spectra emission mechanisms we enumerate two main types, and their birth in radio astronomy. The first is the 21 cm emission from the hyperfine transition in neutral hydrogen (HI), which was discovered as a ‘hiss’ coming from the interstellar medium by Ewen and Purcell [1951] and corroborated by Muller and Oort [1951]. Neutral hydrogen consists of a proton and electron, both of which have a quantum angular momentum spin. Correspondingly, each has a quantised magnetic dipole moment which is proportional to the charge times spin. When the angular momentum spins are aligned then the magnetic dipole moments are anti-align – due to the opposite charges – and the energy of the HI molecule is higher than if the magnetic dipoles were aligned. Occasionally, with a decay rate of 2.9×10^{-15} seconds, an HI molecule with anti-aligned magnetic dipoles will spontaneously undergo a spin-flip to the lower energy level, and emit a photon at approximately 1.4 GHz. Due to the vanishingly small decay rate, in order to see the occurrence of 21 cm emission there must be an extremely large volume of HI. This makes it an ideal candidate to study the cosmic dark ages during the few hundred million years that followed the Big Bang, in which there is abundant HI and little star formation [e.g. Patil et al., 2017].

Another important branch of discrete emission mechanism radio astronomy is radio recombination lines (RRLs). A RRL is the emission associated with the transition of an ion from higher to lower principle quantum numbers, $n \rightarrow n - N$, where N is the number of energy levels transitioned. By far the highest probability transitions are of the form $n \rightarrow n - 1$. Furthermore these transitions of principle quantum numbers are orders of magnitudes more likely than fine and hyperfine transitions such as the one involved in 21 cm HI emission. They were first hypothesised to be observable in interstellar ionised hydrogen (HII) by Kardashev [1959], and have since been observed in galactic HII regions over a range of frequencies 36.5 GHz ($56 \rightarrow 55$) to 404 MHz ($254 \rightarrow 253$). In Ball et al. [1970] the interpretation of RRLs suggested that molecules other than HII were needed to fully explain their data, and since has lead to a rich exploration of the galaxy’s composition in the heavier nuclei [Brocklehurst and Seaton, 1972]. Recently, Emig et al. [2019] have observed the first RRLs at cosmological distances, opening a new field for RRLs in studying distant galactic nuclei.

Continuum spectra radio emission mechanisms form a broad class of non-discrete physical process which generally are to do with moving charges. Here we enumerate two classes, and the physical processes that generate them. The first is emission related to deceleration of charged particles during a scattering event. The typical scenario is an electron scattering off of a proton ion, where kinetic energy is converted into emission. This free-free scattering – so called because the electron begins and ends unbound – is a thermal process when driven by the internal energy of the medium. It can be shown [e.g. Rybicki and Lightman, 1986] that the thermal free-free emission flux density from optically thin systems has very little frequency dependence ($S \propto \nu^{-0.1}$), whereas for optically thick systems there is a steep dependence ($S \propto \nu^2$). Thus, thermal free-free emission can shed light on the density of plasma environments.

The second class of continuum emission mechanisms is called synchrotron emission named after the first observation in 1947 by Robert Langmuir and Herb Pollock at the General Electric synchrotron accelerator. Synchrotron emission is any emission related to acceleration of charged particles perpendicular to their velocity, i.e. $|\mathbf{a} \times \mathbf{v}| \neq 0$, and has a continuous spectrum spanning all frequencies. In laboratory experiments, synchrotron radiation is produced by forcing electrons to move in circular orbits using a magnetic field, and is seen as

a nuisance source of energy loss.

The history of astrophysical synchrotron emission is more complicated as noted by Breus [2001],

In particular, the Russian physicist V.L. Ginzburg broke his relationships with I.S. Shklovsky and did not speak with him for 18 years. In the West, Thomas Gold and Sir Fred Hoyle were in dispute with H. Alfven and N. Herlofson, while K.O. Kiepenheuer and G. Hutchinson were ignored by them.

While the prediction of astrophysical synchrotron *may* be attributed to Shklovsky [1953], the first observation seems to be by Burbidge [1956] of Messier 87, whose jet is shown in Figure 1.

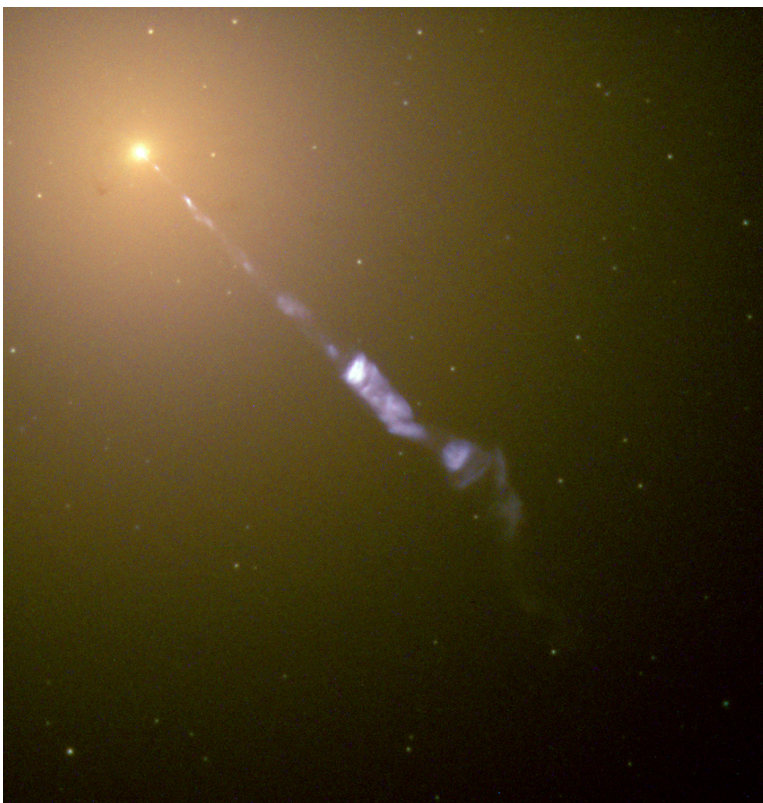


Figure 1: Messier 87 and the radio jet produced by synchrotron. The jet was the first observed astrophysical synchrotron emission.

Astrophysical radio-frequency synchrotron emission comes from charged particles – typically electrons – moving through a magnetic field, where Lorentz force causes acceleration perpendicular to the velocity. This causes particles to trace out circular spirals as they move forward with a gyro-frequency $\nu_e \propto B_{\perp}$, where $B_{\perp}$ is the strength of the magnetic field perpendicular to the velocity of the particle. For $B_{\perp} = 1 \mu\text{Gauss}$ the gyro-frequency of the

electron is approximately 3 Hz. For particles at relativistic speeds, the frequency of emitted photons is boosted by two factors of the Lorentz factor γ due to the relativistic Doppler effect and length contraction. Thus, in the relativistic regime¹, the frequency of maximal emission is then related to the gyro-frequency by $\gamma^2 \nu_e$. For an ultra-relativistic electron with energy of 10 GeV, the Lorentz factor is about 1.96×10^4 and the frequency of maximal radiation is 1.14 GHz.

There are several types of physical processes that can boost an electron to nearly the speed of light. One is a jet stemming from rapidly spinning neutron stars and super massive black holes forming the basis of pulsars [Hewish et al., 1968] and active galactic nuclei [Ambartsumian, 1956, Osterbrock, 1989, Begelman et al., 1984], respectively. Another is diffusive shock acceleration [DSA; Drury, 1983, Blandford and Eichler, 1987], a mechanism by which a charged particle is repeatedly reflected between magnetic ‘walls’ with increasing velocity, which plays an important role in understanding merging galaxy clusters, super nova shocks, and solar flares.

Diffuse emission in galaxy clusters

The cosmic web is the structure of the matter, magnetic field, and kinetic energy on the largest scales of the observable Universe [e.g. van de Weygaert and Schaap, 2009]. The web-like structure evolved from very homogeneous Gaussian primordial conditions and is the perfect example of how complexity can emerge from simple rules and non-complex initial conditions [Vazza, 2020]. Supported by cosmological simulations, the current modern perspective on assembly of the cosmic web is hierarchical in nature [Aragón-Calvo et al., 2010]. Large clusters of galaxies formed at slight over-densities in the matter distribution, from which tenuous filamentary structures extended for many Mpc. Smaller structures formed in the filaments and were drawn toward the clusters, with velocities exceeding 1000 km s^{-1} . Due to hierarchical merging of clusters with accretion from the filaments, galaxy clusters grow to be very massive with masses on the order of $10^{14-15} M_\odot$, and the mass of nearby filamentary structures can be equally massive [Metzler et al., 2001].

Occasionally, galaxy clusters merge converting 10^{56} J of gravitational energy into kinetic energy through a variety of processes including DSA of electrons to ultra relativistic energies in the intra cluster medium (ICM) as well as turbulence. The resulting ultra relativistic cosmic ray electrons (CRe) give off diffuse synchrotron emission in the presence of an ICM magnetic field [for a review see van Weeren et al., 2019]. Diffusive shock acceleration can accelerate non-relativistic thermal populations of CRe, or reaccelerate older aged populations of CRe. The synchrotron emission from the first type is referred to as a radio relic [e.g. Ensslin et al., 1998, van Weeren et al., 2010], while that of the latter is referred to as radio fossil emission [e.g. de Gasperin et al., 2017, Mandal et al., 2020]. There are a number of physical parameters that can be inferred from observations of DSA driven synchrotron emission, including information on the CRe population and ICM magnetic field [Brüggen and Vazza, 2020]. One of the important factors is the speed of the shock which is ultimately coupled to the dynamics of the merger, and premerger initial conditions. Another is the acceleration efficiency which is a phenomenological number describing how much energy is transferred

¹ There are non-relativistic astrophysical sources of synchrotron, which, instead of depending on Lorentz boosting, rely on extremely large magnetic fields. This is known as cyclotron emission and we pay it no further attention here.

to an electron diffusing through the shock front once.

In order to perform inference with non-thermal synchrotron emission – whether from radio halos, relics, or fossils – an electron ageing model must be chosen, which describes how the spectral index of synchrotron emission from a population of ultra-relativistic CRe evolves over time. Two standard models exist, colloquially known as the Jaffe and Perola [JP; Jaffe and Perola, 1973] model and the Kardashev and Pacholczyk [KP; Kardashev, 1962, Pacholczyk, 1970] model. The key difference between the models is that the JP model assumes an anisotropic pitch-angle distribution of the CRe, and the KP model assumes an isotropic pitch-angle distribution of the CRe. The initial formulation of both the JP and KP models assumed that energy was injected into the CRe suddenly at single moment in time after which the electrons age in isolation. Then in Komissarov and Gubanov [1994] both models were extended to incorporate a finite period of continuous energy injection and the effect of energy loss due to inverse Compton scattering.

Another important aspect is the magnetic field of the ICM. Galaxy clusters have magnetic fields on the order of $1 \mu\text{G}$ [e.g. Roettiger et al., 1999], however the magnetic field is highly dependent on the dynamical state of the cluster [Bonafede et al., 2010]. Galaxy clusters can have many dynamical states ranging from disturbed to relaxed [Cassano et al., 2010b], which is characterised by the X-ray morphology of the intra cluster medium (ICM) [Poole et al., 2006, Maughan et al., 2008, Santos et al., 2008]. Difusive shock acceleration driven emission is highly polarised, reflecting the fact that synchrotron emission emissivity is proportional to $|\mathbf{a} \times \mathbf{v}|$ and that the Lorentz force produces acceleration perpendicular to the direction of velocity. Radio relics are primarily observed near the periphery of the galaxy cluster, therefore they probe magnetic fields far from the densest part of the ICM [e.g. Wittor et al., 2019].

Turbulence is another mechanism for kinetic energy dissipation into the ICM following a merger, accounting for approximately 20% – 30% of the thermal energy in merging clusters. Thermal populations of electrons also produce emission due to thermal free-free scattering. In contrast with DSA driven synchrotron emission, turbulence driven emission should have zero average polarisation. Such emission is referred to as a radio halo. Coupled with simulations they can still provide constraints on the magnetic field structure, however due to the dispersive nature of turbulence only statistical studies are possible [Vazza et al., 2011]. **Figure 2** shows an example slice of temperature from a galaxy cluster merger simulation with turbulence.

Radio interferometry

History

One thing that becomes clear is that the radio sky is rich in information on a wide variety of physical processes, so how was radio astronomy born? This account can be found in a number of sources [e.g. Southworth, 1956, Sullivan, 2009]. The answer goes back to 1839 when Carl Gauss postulated that a conducting layer at atmospheric heights should account for the variations in a compass reading of magnetic north – an effect that navigators of the time well knew. This mysterious electric current had supporters throughout the 1800's, such as Lord Kelvin. Around sixty years later, in 1901, Guglielmo Marconi transmitted a radio signal across the Atlantic ocean, which should only have been possible if some reflecting layer bounced the radio signal around Earth's curvature. An electric current surrounding

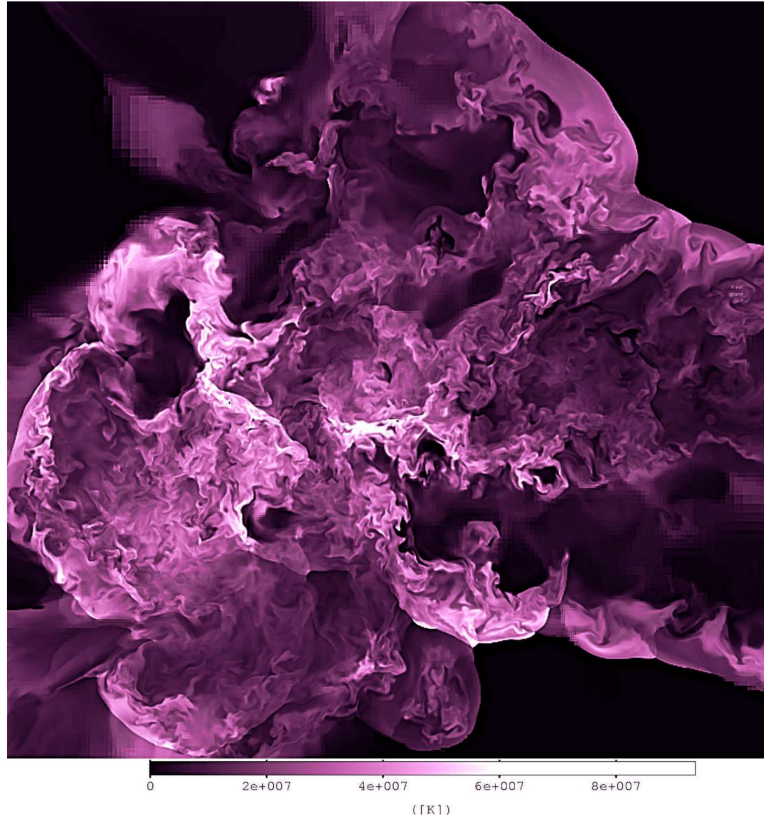


Figure 2: 2-dimensional slice showing the gas temperature for the innermost region of a simulated galaxy cluster, during its main merger event ($z = 0.6$). The side of the slice is $8.8 \text{ Mpc } h^{-1}$ and the depth along the line of sight is $25 \text{ kpc } h^{-1}$. Taken from [Vazza et al., 2011].

Earth seemed like a plausible explanation. For the next three decades astronomers assumed that, while there should be radio emission from Space, it would be bounced back by the same electric layer around the Earth, just as Marconi's signal had been bounced back down. There were few attempts to look for extraterrestrial radio emission, and the development of trans-Atlantic radio communications flourished.

In 1928 a young 23 year old Karl Jansky, newly having attained his Bachelors in physics, was hired by Bell Labs. He was stationed at their Holmdel, New Jersey site and tasked with identifying sources of atmospheric radio interference in telecommunications. Karl built a 100-ft in diameter steerable antenna, which he called the 'merry-go-round' for its construction included four sets of Ford model-T tires on which it could be rotated azimuthally. Over the course of months Karl measured the various components of radio noise with his rig as a function of time of day and azimuth. He identified two types of static-like noise related to thunderstorms, which were well-known at the time to cause radio interference: one due to local thunderstorms, and the second due to distant thunderstorms reflecting off of the ionosphere. In August of 1931, such thunderstorms and associated interference were

common occurrences of night-time, however as Karl continued irregular monitoring into the autumn and winter he began to notice a third hiss-like component to the interference, which he referred to as a ‘very steady continuous interference... that changes direction continuously throughout the day, going completely around the compass in 24 hours’. The direction of this component seemed entirely aligned with the Sun causing him to refer to the hiss as ‘Sun static’. Despite equipment damages due to ferocious winter weather, he was able to monitor into the spring at which point he noticed the hiss direction appeared to drifted in time ‘in accordance with the lengthening day’. This further convinced him that the hiss was somehow related to the vernal Northerly swings of the Sun.

The idea that the source of interference might be sidereal, fixed to the stars, came some time in December of 1932. By this time the reality of the Great Depression had hit Bell Labs, and working hours had been severely cut back limiting his time available to study the hiss. Karl had instead spent most of 1932 studying general methods of measurement of noisy signals, still taking daily measurements but not deeply analysing them. He had noticed that the signal continued to advance in time relative to the Sun, and during a partial solar eclipse in August Karl noted that the hiss intensity had little dependence on the solar occlusion. Despite these observations, Karl must have still held the notion that the hiss was still somehow related to the Sun. For it wasn’t until December 1932 that George C. Southworth, a correspondent working for AT&T in New York, asked Karl to plot his entire years’ data to see if it correlated with his study of ‘diurnal changes in Earth currents’. Upon performing this task, Karl, not an astronomer, was shocked and puzzled to find that the signal had slipped by exactly one day. To an astronomer such a pattern would be obvious, for the stars rise four minutes earlier each day due to the Earth’s rotation about the Sun, adding up to one day over a year. Then Karl’s good friend and bridge partner, Albert Melvin Skellett, helped make the realisation that the period of signal was remarkably the same as the sidereal day. (Albert lead a dual life, being both a radio engineer for Bell Labs, and a graduate student of astronomy at the near-by Princeton University.) This lead Karl to learn about astronomical coordinates, and by Chirstmas 1932 he had determined the signal was fixed in space with a direction of 18 hours right ascension, and -4 degrees declination,

Since I was home [early November] I have taken more data which indicates definitely that the stuff, whatever it is, comes from something not only extraterrestrial, but from outside the solar system. It comes from a direction that is fixed in space and the surprising thing is that... [it] is the direction towards which the solar system is moving in space. According to Skellett, there are clouds of ‘cosmic dust’ in that direction through which the Earth travels.

Karl had fixated on a region of sky in the Sagittarius constellation [Jansky, 1933], which we now know to be the galactic centre. Karl Jansky was subsequently ordered by Bell Labs back to more pressing tasks. The fundamental unit of flux in radio astronomy is named after Karl, the jansky (Jy), $1 \text{ Jy} = 10^{-26} \text{ W Hz}^{-1} \text{ m}^{-2}$. The electric layer that surrounds Earth is the ionosphere, which is the central topic of this thesis.

Following World War II, and the debut of the atomic bomb, there was a great desire to make progress in radio technology. Radio astronomy gained much momentum in the tail winds of radar technology, and indeed there was a great deal of shared knowledge between the two fields. While we can only surmise, radio technology flourished in large part for better methods of surveillance and communication, and perhaps ultimately for the prevention

of another man-made global catastrophe. The renewed interest in radio technologies was a great boon to radio astronomy which lead to radio interferometry being born [Ryle and Vonberg, 1946]. In the ‘West’, radio research institutes sprang up in Europe, especially in Britain and the Netherlands, the USA at the National Radio Astronomy Observatory (NRAO), and Australia at Commonwealth Scientific and Industrial Research Organisation (CSIRO). The Iron Curtain, separating ‘East and West’, provided an – arguably *ideal* – opportunity for independent progress on many topics of mathematics, physics, and engineering [Pesch and Plail, 2009, Gillespie et al., 1977], however in radio astronomy the result was quite the opposite. Just like in the ‘West’, radar and radio astronomy flourished in post-war USSR. However, unlike their counterparts in Europe, USA, and Australia, who primarily worked out of universities, all USSR radio astronomy was conducted at military institutes and was heavily censored. Most discoveries were never published in academic journals, and those that did get published were often missing key data [Braude et al., 2012]. This policy of secrecy ultimately lead to a lack of trust by their contemporaries, and hence today there are very few references to early-day radio astronomy out of the ‘East’.

Eventually, radio astronomy spread around the world to places like India, South America, Canada, South Africa, and China. Today, there are global collaborations undertaking some of the most massive, in terms of data, human experiments, such as the Low-Frequency Array (LOFAR), with even bigger experiments in the near future such as the Square Kilo-Metre Array (SKA). Aperture synthesis radio interferometry, which will shortly be described, naturally breeds unification among countries; for the need to cover long baselines – in order to resolve the most distant objects in the Universe – necessitates an international community. The irony is profound as radio interferometry, which sprang from war, promotes peace.

We’ll now proceed with explanation of interferometry from a modern perspective in a way that is foundational to Chapters 2–4.

Imagine yourself lying on the bottom of a swimming pool, and looking up through the water. You are the radio telescope and the water is the Earth’s atmosphere and ionosphere. Now imagine that there’s a ceiling up overhead, with faint frescoes that you’re trying to make out through the water. The frescoes are the faint objects of the universe. Then, somebody goes and puts a thousand bright lights on the ceiling (the radio-bright objects of the universe) – you’re now trying to make out the frescoes behind the lights’ glare. And then the wind picks up so the water gets choppy.

OLEG SMIRNOV

Technical introduction to scattering and radio interferometry

Most of the rich information in astrophysical radio emission requires good angular resolution, which a single antenna is unable to provide. Interferometry is a way of inferring the shape and location of a source by measuring the interference pattern of light which has simultaneously travelled two different optical pathways. Interference is a property of waves and is completely described by Maxwell’s equations of electromagnetism, thus we start there. We present the

theory of scattering as it relates to the ionosphere and the observables of radio interferometry, which sets the stage for Chapters 2 to 4. We treat the topic with slightly more detail than a typical introduction to the ionosphere and radio interferometry, though we skip rigour where appropriate for flow. The aim is that this might also serve as a summary reference for future work on the topic.

The ionosphere is a low-density multispecies ion plasma situated between 50km and 1000km from Earth's surface [for a detailed description see Kivelson and Russell, 1995]. The plasma precipitation is fed by extreme ultraviolet radiation from the Sun, and the bulk motions are driven by a complex network of forces. The ionosphere is likely a turbulent medium [Mevius et al., 2016], and displays many discrete behaviours [Jordan et al., 2017]. The spatially and temporally varying free electron density (FED) in the plasma alters the local dielectric properties of ionosphere, and this leads to a refractive index less than unity that decreases linearly with FED and wavelength of the radiation to second order. The refractive index is given by the Lassen-Appleton-Hartree equation [Hutchinson, 2005]. The ionosphere thus acts as a random scattering medium, and its main effects on the image quality are a reduction of dynamic range, distortion of the shapes of sources, and decoherence of the visibilities. We present a description of the effect of the ionosphere from first principles.

To begin, consider monochromatic electric and magnetic fields of the form $\mathbf{E}(\mathbf{r}, t) = e^{-i\omega t} \mathbf{E}(\mathbf{r})$ and $\mathbf{B}(\mathbf{r}, t) = e^{-i\omega t} \mathbf{B}(\mathbf{r})$ respectively. Going forward we'll drop the spatial functional argument for clarity. Using Fourier representation we can build any electric or magnetic field from various monochromatic components and the following derivation thus applies to generic fields. Recall that in an isotropic linear medium that the electric field is related to the electric displacement field by $\mathbf{E} = \epsilon^{-1} \mathbf{D}$ and the magnetic field is related to the magnetic flux density by $\mathbf{B} = \mu \mathbf{H}$, where ϵ and μ are the permittivity and permeability respectively. We assume that both ϵ and μ do not change over the timescale of propagation, and that μ is also constant over space. The FED in the ionosphere induces a spatially² varying permittivity ϵ but does not alter permeability μ . Furthermore, we assume there are no sources of charge or current, i.e. that macroscopically the ionosphere is charge neutral on scales comparable to the wavelength. Maxwell's equations of electromagnetism are then,

$$\nabla \cdot (\epsilon \mathbf{E}) = 0 \quad (1)$$

$$\nabla \cdot (\mu \mathbf{H}) = 0 \quad (2)$$

$$\nabla \times \mathbf{E} = i\omega \mu \mathbf{H} \quad (3)$$

$$\nabla \times \mathbf{H} = -i\omega \epsilon \mathbf{E}, \quad (4)$$

where we've used the fact that all temporal dependence is contained in the unitary factor $e^{-i\omega t}$. We now define the refractive index as $n^2 \triangleq c^2 \mu \epsilon$, and the wavenumber as $k = (2\pi)n/\lambda$ where λ is the wavelength in vacuum. In general, the wavenumber changes spatially due to the refractive index. Taking the curl of Eq. 3, inserting Eq. 4, and using a vector calculus identity we get,

$$\nabla^2 \mathbf{E} + k^2 \mathbf{E} = \nabla(\nabla \cdot \mathbf{E}). \quad (5)$$

By making use of Eq. 1, the right-hand side is found to be equivalent to $\nabla(\nabla \cdot \mathbf{E}) = -\nabla(\mathbf{E} \cdot \nabla \log \epsilon_r)$, and is zero if $\mathbf{E} \perp \nabla \log \epsilon_r$ or $\log \epsilon_r = 0$ everywhere, where $\epsilon_r = \epsilon/\epsilon_0$ is the relative

²The FED is assumed to be constant in time over the timescale of propagation of light through the medium.

permittivity. These conditions are true if the wave is normal incident on a scattering surface, or if there is no scatterer. More generally, we can neglect $\nabla(\nabla \cdot \mathbf{E})$ if the length scale over which ϵ_r changes is much larger than the wavelength. In the ionosphere, the irregularities vary over kilometres which is thousands of times larger than the wavelength. Since we assume $\nabla(\nabla \cdot \mathbf{E}) = 0$ we can apply the linear operator to each electric field component and treat the electric field as a scalar. We therefore drop the vector notation for the electric field. In this case Eq. 5 turns into,

$$(\nabla^2 + k_m^2)E = -k_m^2 \Phi E, \quad (6)$$

where $k_m = \frac{2\pi}{\lambda}$ is the constant wavenumber of the background medium which we take to be vacuum, and $\Phi = (n^2 - 1)$ is the scattering potential.

Green's function. The general solution to Eq. 6 can be written in terms of the Green's function for the Helmholtz equation, which is a function satisfying $(\nabla^2 + k_m^2)g(\mathbf{r}, \mathbf{r}') = -\delta(\mathbf{r} - \mathbf{r}')$. Since the coefficients on the left side are coordinate-independent the Green's function is translation-invariant, thus Green's function can be found by taking the Fourier transform³ and solving,

$$\tilde{g}(\mathbf{s}) = \frac{1}{4\pi^2 s^2 - k_m^2} \quad (7)$$

$$\Rightarrow g(\mathbf{r}, \mathbf{r}') = \frac{e^{ik_m|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|}. \quad (8)$$

We have skipped a large amount of rigour in deriving Eqs. 7 and 8, notably we have not even mentioned generalised functions. For a very clear account of all assumptions including the Sommerfeld radiation condition see Schmalz et al. [2010]. Note, it is possible to derive the Green's function including the $\nabla(\nabla \cdot \mathbf{E})$ term [Krüger et al., 2017]. The additional terms in this complete Green's function are negligible when $k_m^2 r^2 \gg O(25)$, which is where the condition to drop the $\nabla(\nabla \cdot \mathbf{E})$ term comes from. Both the Green's function and its Fourier transform will be useful in the following.

The general solution to Eq. 6 is then found by [Evans et al., 1999],

$$E = k_m^2 \int_V g \Phi E \, dV + \oint_S (g \nabla E - E \nabla g) \cdot d\mathbf{A} \quad (9)$$

Equation 9 is an implicit solution since the solution appears also in the right-hand side, therefore further assumptions are required to explicitly solve. Importantly, in order for the solution to be defined we must stipulate boundary conditions on a surface S enclosing our volume V .

Interestingly, George Green (ca. 1793 – 1841) was at least several decades ahead of his time in formulating his theory on the solution to differential equations, despite having no formal education in – or indeed any real exposure to – mathematics [Challis and Sheard, 2003]. He was not even really known about until years after his death, such that this day no

³Note that in this work we use the unitary definition of the Fourier transform with frequency variable in units of cycles.

one knows what he looked like. He grew up as the son of a miller near Nottingham, which was rural farmland in the early 19th Century, and had little in the way of education. It would have been hard to find access to the then cutting-edge mathematics that were happening in places like Cambridge, and Paris. At the time Nottingham was mostly farmland, and the only resources he had access to were a local library where there certainly was nothing in the way of advanced mathematics to be found. The only possible connections to non-banal mathematics would have been through a local teacher who had studied mathematics before becoming a teacher. Nonetheless, George Green's works are the foundation upon which all equations of interest in physics are solved, such as Maxwell's equations, the elastic wave equation, the Helmholtz equation, the Schrödinger equation, the Dirac equation, the viscous fluid equation, the Klein-Gordon equation, and Einstein's equations to name but a few.

There are a number of ways to solve Eq. 9 which differ in the functional basis taken and in the assumed conditions on the scattering potential Φ . In this work we discuss the Born and Rytov approximations whose first-order approximations are equivalent [e.g. Born and Wolf, 1999, Blackledge, 2005]. As will become apparent, the two approximations differ in the space where solutions are additive – in the sense of linear differential operators – and therefore each has its own advantages in interpretation.

Born approximation. In the Born approximation we assume additivity in the complex wave function, i.e. we solve the linear equation Eq. 9 given boundary conditions for E . We choose the boundary conditions to be a super position of temporally incoherent plane waves on the celestial sphere,

$$E_i(\mathbf{r}) = \int_{\mathbb{S}^2} a(\hat{\mathbf{s}}) e^{ik_m \hat{\mathbf{s}} \cdot \mathbf{r}} d\hat{\mathbf{s}} \quad (10)$$

where $|a(\hat{\mathbf{s}})|^2$ is the specific intensity, or surface brightness, from the direction $\hat{\mathbf{s}}$ on the celestial sphere, in units Jysr^{-1} . It gives the power per area per frequency per solid angle. Clearly, E_i satisfies the homogeneous Helmholtz equation $(\nabla^2 + k_m^2)E_i = 0$. By temporally incoherent we mean that each $a(\hat{\mathbf{s}})$ has a unitary factor with random phase, η , of the form $e^{i\eta t}$ such that $\langle e^{i2\pi(\eta_1 - \eta_2)t} \rangle = \delta(\eta_1 - \eta_2)$ where averaging is over a time much longer than λ/c [Mandel and Wolf, 1965]. Using Green's theorem the second integral in Eq. 9 simplifies to E_i [Evans et al., 1999] and the Born solution becomes,

$$E^{\text{Born}} = k_m^2 \int_V g \Phi E dV + E_i \quad (11)$$

$$\triangleq E_s + E_i. \quad (12)$$

Equation 11 is still an implicit solution, however a series representation known as the Born series is available,

$$E_{n+1}^{\text{Born}} = E_i + k_m^2 \int_V g \Phi E_n^{\text{Born}} dV, \quad (13)$$

with $E_0^{\text{Born}} = E_i$. The series defined by Eq. 13 converges if the maximal absolute eigenvalue of the matrix representation – the spectral radius – of the convolution $g \star \Phi$ is less than unity [Osnabrugge et al., 2016]. Equivalently, $k_m^2 \int_V g \Phi dV < 1$ everywhere in V [Blackledge, 2005].

In the ionosphere this is almost always true, though periods of intense solar activity can decrease the refractive index which increases the absolute value of the scattering potential. The assumptions of radio interferometry are violated during these periods and the data should be thrown away.

Going forward we will assume the first-order Born approximation, $E^{\text{Born}} \approx E_1^{\text{Born}}$, which has a convenient representation via Fourier transforms. Since Green's function, Eq. 8, is translation-invariant the first-order Born approximation can be written as a convolution, $E^{\text{Born}} = E_i + k_m^2 (g \star \Phi E_i)$. We apply the convolution theorem and write the first-order Born approximation as,

$$E^{\text{Born}}(\mathbf{r}) = E_i(\mathbf{r}) + \underbrace{k_m^2 \int e^{i2\pi\mathbf{s}\cdot\mathbf{r}} \tilde{g}(\mathbf{s}) \int_{\mathbb{S}^2} a(\hat{\mathbf{s}}) \tilde{\Phi}(\mathbf{s} - \lambda^{-1}\hat{\mathbf{s}}) d\hat{\mathbf{s}} d\mathbf{s}}_{E_s(\mathbf{r})}, \quad (14)$$

where $\tilde{\Phi}$ is the 3D Fourier transform of the scattering potential. Equation 14 is the scattering equation first written down by Max Born in the 1920s. It implies that if we can measure the scattered wave, $E_s = E^{\text{Born}} - E_i$, then this allows us to reconstruct the scattering potential by solving an inverse problem. Such problems are called inverse scattering problems and are some of the hardest problems in mathematics and physics. Indeed, Richard Feynman's creation in the 1960s of Feynman diagrams is all about solving scattering problems, and is also based on Green's functions.

The Born approximation has a nice connection to the plane wave spectrum of Fourier optics [Born and Wolf, 1999], which will be useful for us. Let the coordinate frame be Cartesian $\mathbf{r} = x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}}$ with $\hat{\mathbf{z}}$ pointing towards the observer, and the origin fixed to the centroid of the scattering potential. In this coordinate system we parametrise the plane wave directions with $\hat{\mathbf{s}} = (l, m, n \triangleq \sqrt{1-l^2-m^2})$ with $l^2 + m^2 < 1$. We will use notation that subscript 2D denotes the $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ components of the vector, e.g. $\hat{\mathbf{s}}_{2D} = l\hat{\mathbf{x}} + m\hat{\mathbf{y}}$. Consider the 2D Fourier transform of Eq. 14 in a plane spanned by $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$. Since the incoming wave, E_i , is a superposition of plane waves, its 2D Fourier transform can easily be derived using the definition of the Fourier transform of the Dirac delta, and is,

$$\tilde{E}_i(\mathbf{s}_{2D}; z) = \int_{\mathbb{S}^2} a(\hat{\mathbf{s}}) \delta(\mathbf{s}_{2D} - \lambda^{-1}\hat{\mathbf{s}}_{2D}) e^{ik_m n z} d\hat{\mathbf{s}}. \quad (15)$$

The 2D Fourier transform of the scattered wave term, E_s , requires a bit more work. From Eq. 14 we can read off the 3D Fourier transform of E_s ,

$$\tilde{E}_s(\mathbf{s}) = \frac{k_m^2}{4\pi^2 s^2 - k_m^2} \int_{\mathbb{S}^2} a(\hat{\mathbf{s}}) \tilde{\Phi}(\mathbf{s} - \lambda^{-1}\hat{\mathbf{s}}) d\hat{\mathbf{s}}, \quad (16)$$

where we substituted the Green's function Fourier transform from Eq. 7. The 2D Fourier transform is then found by inverse transforming s_z back to the spatial domain. In order to do this we use the residue theorem to integrate around the singularity at $s_z^0 = +\sqrt{\frac{k_m^2}{4\pi^2} - s_x^2 - s_y^2}$ using the partial fraction expansion $\frac{1}{4\pi^2 s^2 - k_m^2} = \frac{1}{2s_z^0} \left(\frac{1}{s_z - s_z^0} - \frac{1}{s_z + s_z^0} \right)$,

$$\tilde{E}_s(\mathbf{s}_{2D}; z) = ik_m^2 \frac{e^{i2\pi s_z^0 z}}{4\pi s_z^0} \int_{\mathbb{S}^2} a(\hat{\mathbf{s}}) \tilde{\Phi}(\mathbf{s}^0 - \lambda^{-1}\hat{\mathbf{s}}) d\hat{\mathbf{s}}, \quad (17)$$

where $\mathbf{s}^0 = (s_x, s_y, s_z^0)$. Note this requires that $\tilde{\Phi}$ contains no singularities, which is true of the ionosphere.

Equations 15 and 17 are exact for the first-order Born approximation. Let us examine them qualitatively. Both equations contain exponential terms of a similar form which encode the propagation of the plane wave spectrum along $\hat{\mathbf{z}}$. Notice that in the plane wave spectrum of the scattered wave, Eq. 17, for directions of propagation where $s_x^2 + s_y^2 > \frac{k_m^2}{4\pi^2}$ the wave decays with z . Such waves are called evanescent waves and are not detectable by radio interferometers.

Equation 17 is a weighted sum of the incoming wave amplitudes, $a(\hat{\mathbf{s}})$, with the scattering potential spectrum, $\tilde{\Phi}$, evaluated on a sphere of radius λ^{-1} with origin offset by $\mathbf{s}^0$. For the purpose of illustration, imagine the scattering potential is an isotropic Gaussian blob with a characteristic length scale l . This imposes a characteristic length scale in the Fourier domain of $\sigma = (2\pi l)^{-1}$ beyond which the spectrum amplitude decays to zero, i.e. $\tilde{\Phi}(\mathbf{s}) \approx 0$ when $|\mathbf{s}| \gg \sigma$. This implies that observers looking in direction $\hat{\mathbf{s}}^0$ will only see scattered emission from celestial sources with directions, $\hat{\mathbf{s}}$, satisfying $\arccos(\hat{\mathbf{s}}^0 \cdot \hat{\mathbf{s}}) \lesssim \frac{\lambda}{2\pi l}$. The smallest ionospheric structures often range between 1–5 km with sub-km structures occasionally occurring [Yeh and Swenson, 1959, Yeh, 1962, Mevius et al., 2016]. At a radio frequency of 150 MHz this implies light from a single source is scattered into a cone of approximately $1.1[l/\text{km}]^{-1}$. This implies that diffraction effects are weak.

A useful property of an optical system is whether or not it is isoplanatic [Fellgett and Linfoot, 1955]. In general, when the electric field plane wave spectrum (Eqs. 15 and 17) at the detector surface can be sufficiently approximated by a planar convolution of the plane wave spectrum on the object surface ($\mathbb{S}^2$ in our case), then the optical system is called isoplanatic. We note that both Eqs. 15 and 17 are not planar convolutions but rather convolutions over the 2-sphere. Therefore, in general, an astronomical radio interferometer is not an isoplanatic optical system. In radio interferometry there are two sources of non-isoplanaticity: one from the curvature of the celestial sphere, and one from scattering. This can be seen by interpreting the convolution kernels in Eqs. 15 and 17 as the sources of non-isoplanaticity. Various methods exist in radio interferometry for handling non-isoplanaticity due to curvature of the celestial sphere in the case of non-coplanar arrays [e.g. Cornwell and Perley, 1992, Cotton et al., 2004, Cornwell et al., 2008, van Weeren et al., 2016, Tasse et al., 2018]. The methods introduced in Chapters 2 and 3, and applied in Chapter 4, are about handling non-isoplanaticity due to scattering. In general, non-isoplanaticity can be handled by dividing the object surface into isoplanatic patches [Fellgett and Linfoot, 1955].

In radio interferometry it is not the electric field that is measured but rather the electric field coherence, $V(\mathbf{r}, \mathbf{r}') = \langle E(\mathbf{r}, t) E^*(\mathbf{r}', t) \rangle$, where $\langle \cdot \rangle$ is averaging over a time period much longer than λ/c . For simplicity, we assume a perfect receptive sensitivity, i.e. the radio antennas measure an unbiased electric field⁴. Using the inverse 2D Fourier transform of the plane wave spectrum the electric field coherence is given by,

$$V(\mathbf{r}, \mathbf{r}') = \int \langle (\tilde{E}_i(\mathbf{s}_{2D}; z) + \tilde{E}_s(\mathbf{s}_{2D}; z)) (\tilde{E}_i^*(\mathbf{s}'_{2D}; z') + \tilde{E}_s^*(\mathbf{s}'_{2D}; z')) \rangle e^{i2\pi(\mathbf{s}_{2D} \cdot \mathbf{r}_{2D} - \mathbf{s}'_{2D} \cdot \mathbf{r}'_{2D})} d\mathbf{s}_{2D} d\mathbf{s}'_{2D} \quad (18)$$

⁴Incorporating a receptive sensitivity is simply a matter of multiplying the plane wave spectrum by the directional sensitivity gain of the antenna.

The standard equation for electric field coherence seen in radio interferometry is stationary [e.g. Perley et al., 1989], i.e. depends only on the separation between two positions, however the assumptions arriving there are often absent. In general, the electric field coherence is non-stationary [e.g. Mandel and Wolf, 1965, Garcia-Sucerquia et al., 2002], i.e. depends on two positions, and is only truly stationary for incoherent emission satisfying the homogeneous Helmholtz equation. This is true of the celestial radiation, E_i , but not true of the scattered wave, E_s . The scattered wave gains spatial coherence similar to how spatially incoherent monochromatic radiation passing through two closely separated pin holes gains spatial coherence and produces an interference pattern that is not translation-invariant. However, the electric field coherence can be approximated as stationary using the same observation noted earlier; that the sparse incoherent celestial emission is scattered into a narrow cone of diameter approximately $1'$, so that two scattered waves interfere only if their wave vectors are nearly perfectly aligned. Under this assumption, using Eqs. 15 and 17, Eq. 18 can be worked out,

$$V(\mathbf{r}-\mathbf{r}') = k_m^2 \int_{\mathbb{S}^2} \frac{e^{i2\pi\lambda^{-1}\hat{\mathbf{s}}\cdot(\mathbf{r}-\mathbf{r}')}}{4s_z^2} \left(\int_{\mathbb{S}^2} |\tilde{\Phi}(\lambda^{-1}\hat{\mathbf{s}} - \lambda^{-1}\hat{\mathbf{s}}')|^2 \langle |a(\hat{\mathbf{s}}')|^2 \rangle d\hat{\mathbf{s}}' \right) d\hat{\mathbf{s}} \\ + \int_{\mathbb{S}^2} e^{i2\pi\lambda^{-1}\hat{\mathbf{s}}\cdot(\mathbf{r}-\mathbf{r}')} \langle |a(\hat{\mathbf{s}})|^2 \rangle d\hat{\mathbf{s}}. \quad (19)$$

Note that due to the stationary assumption there are no cross-coherence terms, i.e. $\langle \tilde{E}_i(\mathbf{s}_{2D}; z) \tilde{E}_s^*(\mathbf{s}_{2D}; z') \rangle + \langle \tilde{E}_s(\mathbf{s}_{2D}; z) \tilde{E}_i^*(\mathbf{s}_{2D}; z') \rangle = 0$, which can be verified directly. Equation 19 comprises inverse Fourier transforms on the 2-sphere, which are in the class of restricted Fourier transforms [e.g. Barcelo et al., 1997, Tao, 2003]. Efficient methods of computing (inverse) Fourier transforms and convolutions on the 2-sphere exist [e.g. Driscoll and Healy, 1994, Jarosz, 2008]. The use of such methods result is an exact treatment of non-isoplanatic systems [Eastwood et al., 2018], yet still remain relatively unexplored in radio interferometry. Equation 19 is an oscillatory integral, so another way to exactly treat non-isoplanaticity is via oscillatory integral methods such as the method of stationary phase. Such methods are completely unexplored in radio interferometry.

For incoherent radiation the isoplanatic assumption is equivalent to the condition that the specific intensity in the detector surface can be written as a planar convolution of the specific intensity on the object surface. The kernel of the convolution in this case is called the point spread function (PSF). Again, there are two sources of non-isoplanaticity due to celestial sphere curvature, and scattering. For a coplanar array (detector surface of constant z) an optical system with incoherent radiation becomes isoplanatic. However, recall that scattering introduces coherence, and therefore a system with scattering is not isoplanatic even for coplanar arrays.

For illustrative purposes, assume the special case of a coplanar array $z = z'$, and let the electric field coherence be measured everywhere in the plane of the interferometer. We observe that the second term in Eq. 19 reduces to a 2D Fourier transform relation between the electric field coherence and the 2D power spectrum of the incoming electric field⁵, which we call the specific intensity, sky-brightness, or image, in units of Jy sr^{-1} . This is known as the Wiener-Khinchin theorem. Observe that the first term does not simplify for a coplanar

⁵This technically requires changing the domain of integration to $\mathbb{R}^2$ and defining the integrands as zero outside the unit disk.

array due to the s_z^2 in the denominator. However, if we additionally add the assumption that the celestial emission is restricted to a narrow field of view, such that $s_z^2 \approx 1$, then we can also apply the Wiener-Khinchin theorem. In this case, we see that there are two images; one corresponding to the celestial emission and one to scattered emission respectively,

$$\mathcal{I}_i(\hat{\mathbf{s}}) = \langle |a(\hat{\mathbf{s}})|^2 \rangle \quad (20)$$

$$\mathcal{I}_s(\hat{\mathbf{s}}) = \frac{k_m^2}{4s_z^2} \int_{\mathbb{S}^2} |\tilde{\Phi}(\lambda^{-1}\hat{\mathbf{s}} - \lambda^{-1}\hat{\mathbf{s}}')|^2 \langle |a(\hat{\mathbf{s}}')|^2 \rangle d\hat{\mathbf{s}}'. \quad (21)$$

The interpretation of Eq. 21 can be thought of with a toy model. Imagine the ionosphere is a flat layer perpendicular to zenith with a sinusoidal wave of refractive index with a spatial frequency, $\mathbf{s}_{\text{ion}}$, and let there be one source at zenith. Then, clearly $\tilde{\Phi}(\mathbf{s})$ will have a delta peak at $\pm \mathbf{s}_{\text{ion}}$. Now, Eq. 21, says that a fraction of the source flux will be shifted symmetrically by $\pm \mathbf{s}_{\text{ion}}$. This would look like two speckles offset from the source at zenith. Therefore, the intuition behind Eq. 21 is that speckles are symmetrically displaced radially from sources according to the magnitude and direction of strong spatial modes in the ionosphere. Over the course of an observation the integral effect of scattering would be perceived as a halo of speckles, which grows brighter proportional to the specific intensity of the source. Since the aperture of an Earth rotation synthesis radio interferometer [Högbom, 1974] is sparsely covered, a displaced speckle would appear as a displaced instantaneous PSF [Born and Wolf, 1999]. Another characteristic comes from the property of temporal continuity in the ionosphere. Since wind speed in the ionosphere can be several hundreds of metres per second [Phillips, 1952] we expect temporal continuity of the FED and the speckles should move continuously. Therefore, during periods of high ionosphere activity the speckles should ‘draw’ radial spokes emanating from bright sources, e.g. as seen in Figure 3.3 in Chapter 3.

Finally, observe that Eq. 21 is a relation between the 3D scattering potential spectrum and the 2D scattered intensity. The relation is invertible given knowledge of the celestial sky intensity, $\langle |a(\hat{\mathbf{s}})|^2 \rangle$, a measured scattered intensity $\mathcal{I}_s(\hat{\mathbf{s}})$, and a prior over the scattering power spectrum. Given the above restriction on diffraction to a cone of approximately 1° , without priors the relation is only invertible if $\langle |a(\hat{\mathbf{s}})|^2 \rangle$ is dense with respect to this separation. The fact that the complete information of a 3D object can be stored in a 2D object is the principle of holography, first introduced by Dennis Gabor for microscopic wavefront reconstruction that bypasses the diffraction limit of a lens [Gabor, 1949].

Incoherent light is radiation of which we know only certain data of a statistical nature: the energy or intensity, the polarization, the spectral composition. If we interrupt an optical tract at an arbitrary plane, we see only an almost-uniform blur. It must strike one as something of a miracle that a lens can sort out the intermingling wavelets and restore a picture of an object – at least in one plane.

DENNIS GABOR, 1969

Rytov approximation. The Rytov approximation gives a different interpretation of the effects of scattering in terms of what are called *gains*. Gains form the basis of most radio interferometry calibration procedures, and are closely related to Chapters 2 to 4.

The Rytov approximation makes the eikonal transformation, $E^{\text{Rytov}}(\mathbf{r}) = E_0(\mathbf{r})e^{\phi(\mathbf{r})}$, where E_0 satisfies the homogeneous Helmholtz equation $(\nabla^2 + k_m^2)E_0 = 0$ and $(\nabla E_0)/E_0 = \nabla\phi_0$, for some function ϕ_0 [Born and Wolf, 1999]. We choose $E_0(\mathbf{r}, \hat{\mathbf{s}}) = a(\hat{\mathbf{s}})e^{ik_m\hat{\mathbf{s}}\cdot\mathbf{r}}$ which clearly satisfies the homogeneous Helmholtz equation and $(\nabla E_0)/E_0 = \nabla(ik_m\hat{\mathbf{s}}\cdot\mathbf{r})$. If we substitute the eikonal transformation into Eq. 9 then we get a non-linear Riccatian equation for ϕ . If $|\nabla\phi \cdot \nabla\phi| \ll |k_m^2\Phi|$ everywhere in V , then the non-linear Riccatian equation can also be solved by making the substitution $w = \phi/E_0$ yielding [Blackledge, 2005],

$$\phi = \frac{k_m^2}{E_0} \int_V g\Phi E_0 dV. \quad (22)$$

In the Rytov approximation we choose additivity – in the sense of linear operators – in phase rather than electric field. In other words, by making the eikonal transformation, we get a linear differential equation for phase, and not electric field as in the Born approximation. Nonetheless, the principle of super position of electric fields applies and the scattered electric field is given by the difference between the solution and the incoming electric field, $E_s^{\text{Rytov}} = E_0(e^\phi - 1)$. We note that Eq. 22 looks similar to the scattered wave in the first-order Born approximation,

$$E_s^{\text{Born}} = E_0\phi, \quad (23)$$

when we make the choice $E_i = E_0$. Equation 23 in combination with Eq. 14 implies that $E^{\text{Rytov}} = E_0[\exp(E_1^{\text{Born}}/E_0) - 1]$, and when the spectral radius of $g \star \Phi$ is less than unity then, $E^{\text{Rytov}} \approx E_1^{\text{Born}}$. This shows that the two approaches are equivalent in the regime of the typical ionosphere.

Let us now make the connection between the Rytov approximation and classical non-diffractive tomography as introduced by Johann Radon [Radon, 1917]. From the discussion on the Born approximation we already know that diffraction is approximately limited to a cone on the angular scale of $1'$. Given that the average separation between bright celestial sources is greater than this, it makes sense to consider the non-diffractive limit. In this setting we choose to align the coordinate axes with $\hat{\mathbf{z}}$ pointing in the direction of the single plane wave $\hat{\mathbf{s}}$. If we take the 2D Fourier transform of Eq. 23, use the definition of the plane wave spectrum of the Born scattered wave, Eq. 17, and rearrange we get,

$$\tilde{\Phi}(\mathbf{s}^0 - \lambda^{-1}\hat{\mathbf{s}}) = -i \frac{4\pi}{k_m^2} s_z^0 e^{-i2\pi s_z^0 z} \tilde{\phi}(\mathbf{s}_{2D} - \lambda^{-1}\hat{\mathbf{s}}_{2D}; z). \quad (24)$$

where $\mathbf{s}^0 = (s_x, s_y, s_z^0 = \sqrt{\lambda^{-2} - s_x^2 - s_y^2})$ and $\hat{\mathbf{s}} = (0, 0, 1)$. Equation 24 is an exact relation between the Rytov and first-order Born approximations. Let us make the substitution $\mathbf{s}' = \mathbf{s}^0 - \lambda^{-1}\hat{\mathbf{s}}$. Now, from our discussion in the Born approximation we have that the ionosphere is only weakly diffracting and therefore $|\mathbf{s}'|^2 = 2\lambda^{-2}(1 - \hat{\mathbf{s}}^0 \cdot \hat{\mathbf{s}}) \ll \lambda^{-2}$ from which we have that $s_z^0 \approx \lambda^{-1}$ and $s_z' \approx 0$. Equation 24 becomes,

$$\tilde{\Phi}(\mathbf{s}') \approx -i \frac{4\pi}{\lambda k_m^2} \tilde{\phi}(\mathbf{s}'_{2D}; z). \quad (25)$$

If we take the inverse Fourier transform of both sides, using that $s_z' \approx 0$, and rearrange we get

the classic non-diffractive tomography relation,

$$\int e^{-i2\pi s'_z z} \Phi(\mathbf{r}) dz = -i \frac{4\pi}{\lambda k_m^2} \phi(\mathbf{r}) \quad (26)$$

$$\implies \phi(\mathbf{r}) \approx i \frac{\lambda k_m^2}{4\pi} \int \Phi(\mathbf{r}) dz. \quad (27)$$

This is precisely the equation that is foundational to Chapter 2. Another way to arrive at this result is via the Jeffreys-Wentzel-Kramers-Brillouin (JWKB) approximation [e.g. Jeffreys, 1925]. In the JWKB approximation one makes the extra assumption on the eikonal phase, that $\nabla^2 \phi$ can also be neglected, which results in the geometric optics regime. The eikonal representation in the JWKB approximation gives rise to Jones and Müller algebra treatment of ray optics [Jones, 1941, Parke, 1949].

Given Eq. 27 the Rytov solution for a super position of incoming incoherent plane waves, Eq. 10, is given by,

$$E^{\text{Rytov}}(\mathbf{r}) = \int G(\hat{\mathbf{s}}, \mathbf{r}) a(\hat{\mathbf{s}}) e^{ik_m \hat{\mathbf{s}} \cdot \mathbf{r}} d\hat{\mathbf{s}}, \quad (28)$$

where we define the *gains*, $G(\hat{\mathbf{s}}, \mathbf{r}) = \exp \left[i \frac{\lambda k_m^2}{4\pi} \int_{\hat{\mathbf{r}}=\hat{\mathbf{s}}} \Phi(\mathbf{r}) dV(\mathbf{r}) \right]$. This implies that the electric field coherence is given by,

$$V(\mathbf{r}, \mathbf{r}') = \int G(\hat{\mathbf{s}}, \mathbf{r}) G^*(\hat{\mathbf{s}}, \mathbf{r}') (|a(\hat{\mathbf{s}})|^2) e^{ik_m \hat{\mathbf{s}} \cdot (\mathbf{r} - \mathbf{r}')} d\hat{\mathbf{s}}. \quad (29)$$

Equation 29 is the radio interferometry measurement equation for scalar electric fields and perfect antenna sensitivity [Hamaker et al., 1996]. Importantly, we note that in the non-diffractive limit the gains contain no amplitude information. When the full Eq. 22 is used then the resulting gains would contain amplitude information, and this is possibly an avenue of future research.

In comparing Eq. 29 with Eq. 19 we can clearly see the different approximations lead to different interpretations. For the Born approximation, the effect of the scattering on the image is intuitive in terms of the plane wave spectrum. For the Rytov approximation, the effect on the electric field coherence is intuitive in terms of gains. It should be made clear that both are valid – indeed equivalent – approaches. It is the author's opinion that there is still much to gain in radio interferometry by better understanding the foundational equations. Often improvements and understanding come from returning to the basics of a theory, e.g. understanding ‘ghosts’ in images due to over-calibration [Grobler et al., 2014].

Next-generation radio interferometry

The next generation of radio interferometry is pushing technology to its limit. In particular, science with the near-future Square Kilometre Array [SKA; Dewdney et al., 2009] at full capacity is not technologically feasible with current techniques. The key parameters of a radio interferometer are its sensitivity, angular and spectral resolution, bandwidth, and field of view. In terms of instrumental attributes these parameters correspond to surface area, baseline coverage, number of channels, and the type of antenna.

In an aperture synthesis radio interferometer [Högbom, 1974], as collecting surface area increases the volume of data grows quadratically as this typically corresponds to adding more antennas, which adds one extra source of data for every pre-existing antenna. The SKA is a near-future radio interferometer that has been envisioned for over two decades [Braun, 1996, van de Weygaert and van Albada, 1996], and aims to provide 1 km^2 of effective collecting area [Hollitt et al., 2017]. The SKA is the "pie in the sky" of radio interferometry, since with current technology it would only be possible to utilise a mere fraction of its capacity for science. The expected data rate of the SKA is tens of petabytes per day, or about 1 Internet – the same volume as the entire internet. This challenge has spawned a huge interest in algorithms [e.g. Lindner et al., 2015, Offringa, 2016, Wijnholds et al., 2018, Cai et al., 2018, 2019, Lyon et al., 2019] and technology [e.g. Hassan et al., 2013, Isaev and Tarasov, 2016] for handling huge data volumes.

The demand for higher angular resolution means longer baselines are required, and antennas become further separated. This raises the challenge of keeping antennas synchronised and transferring massive amounts of data to a common correlator. This ultimately would result in improved internet infrastructure around the world. Another challenge with long baseline arrays is dealing with the curvature of the Earth, which breaks the coplanarity assumption. This challenge was already beginning to be addressed several decades ago [e.g. Cornwell and Perley, 1992], and was largely solved within the last decade [Cornwell et al., 2008]. The divergence of the radio interferometer from the classical optical detector plane description, wherein fast Fourier transforms do everything, also sparked much work in more general treatments of interferometric imaging [e.g. Tasse et al., 2013].

In this thesis we work with two world-class radio interferometers: Low-Frequency Array [LOFAR; van Haarlem et al., 2013], and the Giant Metre-wave Radio Telescope [GMRT; Gupta et al., 2017].

LOFAR is a high-sensitivity, high-resolution, wide-band, wide-field, low-frequency, fiber-optic connected, phased-array aperture synthesis radio interferometer that spans Europe, from Ireland to Latvia, and France to northern Sweden. It has two operating bands, the high-band antennae (HBA; 115–240 MHz) and the low-band antennae (LBA; 10–80 MHz). In this thesis, we make use of the Dutch LOFAR array, with baselines out to 80 km, and an effective collecting area of 0.04 km^2 at 150 MHz. The expected sensitivity of a standard 8.5 hour observation with the full Dutch LOFAR-HBA and 24 subbands is $72.8 \mu\text{Jy beam}^{-1}$. The original primary aim of LOFAR was to survey the entire Northern sky with unprecedented resolution, depth, and fidelity. The science that comes out of LOFAR now spans all radio fields and domains of astronomy including: cosmic dawn [e.g. Patil et al., 2017, Mertens et al., 2020], early star-formation [e.g. Saxena et al., 2018], active galactic nuclei [e.g. Williams et al., 2018], radio transients [e.g. Kuiack et al., 2019], galaxy clusters [e.g. van Weeren et al., 2019, Mandal et al., 2019], radio exo-planets [e.g. Callingham et al., 2020], and more. LOFAR is a pathfinder for SKA-low, and comes with a great number of technological challenges [Noordam, 2004]. In some ways the Murchison Widefield Array [MWA; Tingay et al., 2013] in Australia is a Southern-hemisphere counterpart to LOFAR, acting as a precursor to the SKA-low.

The GMRT is an aperture synthesis radio interferometer composed of 30 massive steerable 45 m parabolic dishes situated just 80 km north of Pune, India. It had three narrow bands (32 MHz wide) situated at 150 MHz, 320 MHz, and 610 MHz. Given the large collecting area of the dishes, approximately 0.03 km^2 , the GMRT was one of the most sensitive radio

interferometers prior to LOFAR, being approximately three times larger than the Very Large Array [VLA; Condon et al., 1998] and equally large as the single dish giant Arecibo Radio Telescope [Heiles et al., 2001]. It has enabled a host of science ranging from cosmic dawn [e.g. Paciga et al., 2013], radio transients [e.g. Bhattacharyya et al., 2016], radio exo-planets [e.g. Hess and Zarka, 2011], galaxy clusters [e.g. Venturi et al., 2008], to stellar radio emission [Quiroga-Nuñez et al., 2020]. The GMRT was recently upgraded to a wideband version [uGMRT; Gupta et al., 2017], however in this thesis we only made use of the GMRT.

About this thesis

In some sense the SKA, being a far-off intangible goal, has provided the motivation to come together as a community and solve the technological challenges of today's telescopes. Radio interferometry has seen a huge boost in its algorithms in the last decade as we prepare to meet the challenges of SKA. One of the many challenges is direction dependent ionospheric calibration. There are several common methods for ionospheric calibration, e.g. field-based calibration [Cotton et al., 2004] which models image domain apparent source shifts, and facet-based calibration [Intema et al., 2009, van Weeren et al., 2016, Tasse et al., 2018] which performs calibration on disjoint patches of the sky called facets. In this thesis we focus entirely on the facet-based approach.

Facet-based calibration has several orders of complexity. The simplest method is to form an independent data set per facet, by subtracting a model of the rest of the sky, and then perform calibration independently per facet. The result solutions can then be spatially filtered to remove noise and to try to make them spatially consistent with some model. This is the basis of Source Peeling and Atmospheric Modelling [SPAM; Intema et al., 2009]. An extension on this is to also enforce a linear phase model during solve, e.g. that the ionospheric phase distortions are $\propto \nu^{-1}$. This makes forms the basis of Facet Calibration for LOFAR [FACTOR; van Weeren et al., 2016]. The more complicated method is to simultaneously and jointly calibrate all facets on all the visibilities [Tasse, 2014b, Smirnov and Tasse, 2015]. Following this post-processing can enforce a linear phase model, e.g. $\propto \nu^{-1}$ dependence for ionospheric terms. This forms the basis of killMS and DDFacet [Tasse et al., 2018]. All facet-based methods make use of self-calibration [Pearson and Readhead, 1984].

Common to all of these methods is that 1) they are limited by the density of infield-calibrators to form facets, and 2) they do not use any spatial information of the ionosphere. One of the big open questions in radio interferometry is: Is it possible to use the physics of a weakly scattering random medium to improve on these methods of ionospheric calibration? In this thesis we answer this question in theory and in practice.

The thesis is split into two parts. In the Chapter 1 we investigate diffuse synchrotron emission in a high redshift galaxy cluster. We use this investigation to gain a working experience of radio interferometry, as well as to explore an interesting field in astronomy, i.e. the study of the largest scale structure in the Universe. In Chapters 2 to 4 we then tackle the problem of probabilistic calibration of the ionosphere using a physical model of weakly scattering random media. The scope of the project is aimed at investigating ways to solve the problem of the ionosphere for LOFAR, as a path finder for SKA.

In Chapter 2 we apply well-known physics of non-diffractive, or non-dispersive, phase distortions to derive a probabilistic tomographic prior for the ionospheric distortions. The

result is a mean and covariance function for a Gaussian process [Rasmussen and Williams, 2006b], with hyper parameters that encode information about the geometry of the ionosphere including height, thickness, and spectral density. We show, using simulated ionospheres, that the resulting Gaussian process surpasses all other common Gaussian processes in modelling ionospheric distortions and is computationally feasible.

Following this, in Chapter 3, we devise a method that can apply this to real interferometric data. Using DDFacet we acquire calibration data which we use as data in our model. There are a number of challenges we face. Firstly, the calibrator data is low signal-to-noise. This complicates extracting ionospheric information from them. We devise a variational hidden Markov model [HMM; e.g. Rabiner and Juang, 1986] which provides a Bayesian estimate of the distribution of the ionospheric distortions, which forms the input to our model. In applying the method to real data we are forced to make several approximations to the model for computational feasibility. Ultimately, we apply it to a single LOFAR observation and find we are able to significantly reduce ionospheric artefacts.

Finally, in Chapter 4 we assess the feasibility of the method on the deep Lockman Hole data set comprising 12 8.5 hour LOFAR observations. We find that the method significantly improves 10/12 data set, and that 2/12 data sets only show subtle improvements. In no cases it the image quality degraded over the current state-of-the-art. We also jointly image all 100 hours of the deep Lockman Hole data set and find that the improvements extend to multi-epoch observations. In the process we uncover a modelling systematic, which we are able to isolate and characterise. We find that there is a direction-dependent non-ionospheric component that limits the resulting noise level around sources to a few times thermal noise. We propose a correction to the method which shows, in principle, that we can account for it, although we leave the extension of the method to a following paper.

Future directions

Given the success of the model at LOFAR-HBA frequencies, a natural extension is to apply it to LOFAR-LBA. This likely comes with a number of challenges. For example, at lower frequencies the number of suitable calibrators will be far fewer because sensitivity over appropriate time scales will be much lower. In Chapter 2 we showed that the full tomographic version of the model can handle LOFAR-LBA conditions. Therefore, a possible future avenue will be making the full tomographic version computationally feasible to handle this sparse and noisy regime.

In this thesis we focus purely on the ionospheric modelling aspect, and do not attempt to alter the way in which calibration and imaging occurs. However, the author recognises the potential for new calibration and imaging techniques that implicitly account for the ionosphere. Currently, there are no methods that take into account the spatial information and physics of the ionosphere when inverting the visibilities. Furthermore, currently the standard description of the ionosphere in radio astronomy excludes the amplitude information of the ionosphere from the model. The model could be extended to incorporate a model for amplitudes, however this would likely require modelling the ionosphere directly on visibilities.

Another possible future direction has to do with the definition of the radio interferometry measurement equation. The current definition [Hamaker et al., 1996] was formed in a time prior to big-data, tomographic ionosphere corrections, and all development of the past two

decades. Currently, the definition stems from the use of Jones algebra which means that we choose additivity in phase space and that wave-fronts never distort. We therefore see the visibilities as a convex sum of complex exponentials weighted by the celestial specific intensity. There are a several side-effects of this including non-linearity of the visibility model, and phase-wrapping of all propagation effects. However, if we were to take the Born approximation, we would naturally gain linearity of the visibility model, which would make least-squares calibration much more well-founded. This could have extensive implications for calibrating the ionosphere phase and amplitude effects directly from visibilities.



J. G. Albert,
C. Sifón, A. Stroe, F. Mernier,
H. T. Intema, H. J. A. Röttgering,
and G. Brunetti

Complex diffuse emission in the $z = 0.52$ cluster PLCK G004.5-19.5

J. G. ALBERT, C. SIFÓN, A. STROE, F. MERNIER, H. T. INTEMA, H. J. A. RÖTTGERING,
AND G. BRUNETTI

Published in A&A

Received 25 January 2017 / Accepted 3 July 2017

We present radio observations of the galaxy cluster PLCK G004.5-19.5 ($z = 0.52$) using the Giant Metrewave Radio Telescope at 150 MHz, 325 MHz, and 610 MHz. We find an unusual arrangement of diffuse radio emission in the center and periphery of the cluster, as well as several radio galaxies with head-tail emission. A patch of peripheral emission resembles a radio relic, and central emission resembles a radio halo. Reanalysis of archival *XMM-Newton* X-ray data shows that PLCK G004.5-19.5 is disturbed, which has a known correlation with the existence of radio relics and halos. Given that the number of known radio halos and radio relics at $z > 0.5$ is very limited, PLCK G004.5-19.5 is an important addition to understanding merger-related particle acceleration at higher redshifts.

Acknowledgements. J. A., A. S., and H. R. acknowledge financial support from an NWO top subsidy (614.001.006). C. S. acknowledges support from the European Research Council under FP7 grant number 279396. We thank the staff of the GMRT, who have made these observations possible. The GMRT is run by the National Centre for Radio Astrophysics of the Tata Institute of Fundamental Research. This work is based on observations obtained with XMM-Newton, an ESA science mission with instruments and contributions directly funded by ESA member states and the USA (NASA). The SRON Netherlands Institute for Space Research is supported financially by NWO, the Netherlands Organisation for Space Research. G. B. acknowledges partial support from grant PRIN-INAF2014. J. A. thanks R. Cassano for private discussion on use of morphological parameters at higher redshifts.

1.1 Introduction

Galaxy clusters are the most massive gravitationally bound structures in the Universe with masses and volumes of the order of $10^{14-15} M_{\odot}$ and 100 Mpc^3 . Hierarchical mergers between clusters at the intersections of the cosmic web, with relative velocities near 1000 km s^{-1} , can release gravitational energy of the order of 10^{63} erg . A fraction of this energy is dissipated into shocks and turbulence, which in turn accelerate cosmic ray electrons (CRE) and hadrons in the intracluster medium (ICM). The shock acceleration might occur via diffusive shock acceleration [DSA; for a comprehensive review see Drury, 1983, Malkov and O’C Drury, 2001, Kang and Ryu, 2013]. Turbulence can reaccelerate relativistic particles via second-order Fermi mechanisms [for an updated review see Brunetti and Jones, 2014]. The relativistic CRE give rise to synchrotron radiation in the presence of magnetic fields, which are typically of the order of μG [Roettiger et al., 1999].

Emission tracing the shock fronts is referred to as radio relics, which are typically diffuse, elongated, and polarised sources located near the cluster periphery. Central unpolarized, diffuse radio sources are known as radio halos and likely result from turbulent acceleration. Clusters hosting radio relics and halos have a disturbed morphology [Cassano et al., 2010b], but not all merging clusters exhibit halos, suggesting that halos originate from a hierarchy of complex mechanisms [Donnert et al., 2013, Brunetti, 2016].

The current catalogues of radio relics and halos are far from complete, mainly because of the limited sensitivities of existing radio surveys. Nuza et al. [2012] performed an analysis of MareNostrum, a high-resolution cosmological simulation, and modelled the abundance of radio relics across redshift. By assuming a DSA efficiency, they estimate $\gtrsim 100(800)$ to be found at $0.5 < z < 1$ by the upcoming Tier-1 survey of LOFAR [van Haarlem et al., 2013] at 60 MHz (120 MHz). Currently, there are only four relics observed beyond $z = 0.5$ (MACS J1149.5+2223 at $z = 0.54$ Bonafede et al., 2012; MACS J0717.5+3745 at $z = 0.55$ Bonafede et al., 2009, van Weeren et al., 2009; El Gordo at $z = 0.87$ Lindner et al., 2014, Botteon et al., 2016 and MACS J0025.4-1222 at $z = 0.586$ Riseley et al., 2017).

The number of observed halos beyond $z = 0.5$ has been limited by the sensitivity of low-frequency radio observations. It is estimated that there are about 500–1000 radio halos for $z < 0.6$ that should be found at 150 MHz by LOFAR Tier-1 Cassano et al. [2006]. This is because inverse Compton (IC) losses (scaling with $(1+z)^4$) become more efficient than DSA at higher redshifts [Cassano et al., 2006]. Thus IC losses are able to suppress high-energy CRE, giving rise to an increasing fraction of radio halos with very steep spectra, which are only bright at low frequencies [e.g. Brunetti et al., 2008].

One as yet unanswered question is how efficient DSA and turbulence are at electron acceleration. In particular, it is unclear how the two mechanisms evolve over cosmic time, and what their impact is on the underlying magnetic field. Furthermore, we eventually wish to understand how shocks and turbulence affect host galaxies following a merger.

We present galaxy cluster PLCK G004.5-19.5, which lies at a redshift ($z = 0.516$) with few known radio relics and halos; this makes it a tantalizing target to search for diffuse emission. It was discovered by the *Planck* satellite through the Sunyaev-Zel’dovich (SZ) effect and confirmed with X-ray observations [Planck Collaboration et al., 2011, 2014]. This cluster is hot, $10.2 \pm 0.5 \text{ keV}$, and very massive, $M_{500}^{\text{SZ}} = (10.4 \pm 0.7) \times 10^{14} M_{\odot}$, and it hosts strong lensing arcs [Sifón et al., 2014]. Initial follow-up studies in low-resolution low-frequency archival data found that PLCK G004.5-19.5 hosted strong radio emission [Sifón et al., 2014]. However,

since the radio sources were unresolved, the nature of the radio emission could not be solved.

In this paper we offer new high-resolution 150 MHz, 325 MHz, and 610 MHz observations of the cluster PLCK G004.5-19.5 using the Giant Metrewave Radio Telescope (GMRT), and a reanalysis of *XMM-Newton* data for a morphology study of the X-ray emission. In Section 1.2 we explain our reduction method and how we corrected *XMM-Newton* X-ray data of PLCK G004.5-19.5 for background and vignetting. In Section 1.3 we analyse the X-ray and radio emission.

In this work, we assume a flat Λ CDM cosmology with $h = 0.70$, and $\Omega_m = 0.30$, which gives an angular scale at $z = 0.516$ of $6.2 \text{ kpc}''$.

1.2 Observations and data reduction

1.2.1 GMRT observations of PLCK G004.5-19.5

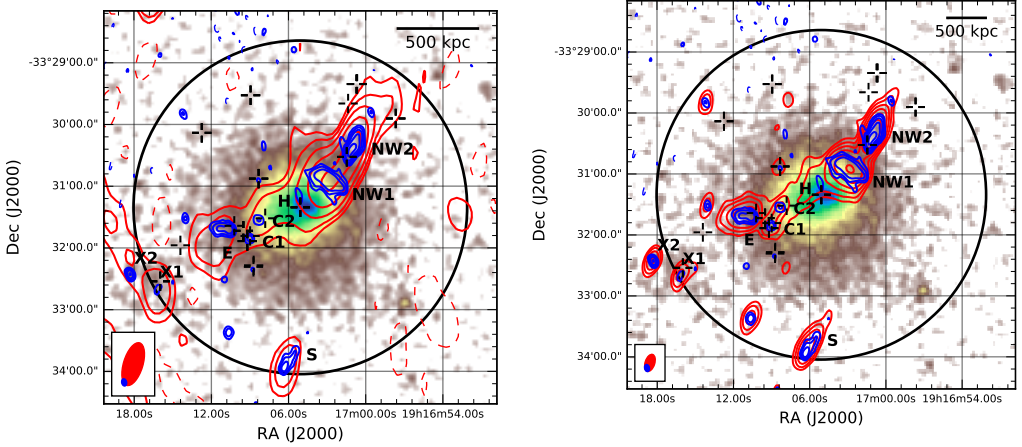
We carried out a set of observations on PLCK G004.5-19.5 using the GMRT at 150 MHz, 325 MHz (PI: A. Stroe, Project: 27_051), and 610 MHz (PI: C. Sifón, Project: 25_036) with total times on source of 445 min., 390 min., and 798 min. We performed 15 min. observations on flux calibrators 3C 48 and 3C 286 at the start and end of the observation session, and 10 min. interleaved observations of the nearby and bright phase calibrator J1924-292 that has a flat spectrum amplitude [Healey et al., 2007].

We removed the initial radio frequency interference (RFI) of the obvious corrupt time and frequency data by flagging by hand. We then used the source peeling and atmospheric modelling pipeline [SPAM; for a full description see Intema et al., 2009, Intema, 2014]. SPAM performs antenna delay, bandpass, phase, and amplitude calibration, and then multiple rounds of self-calibration and further flagging of bad data. It then iteratively subtracts all sources except for bright calibrators, and calibrates direction dependently. It then spatially fits a model to the direction-dependent calibration products and derives interpolated ionospheric corrections.

For our choice of initial sky model for self-calibration we used a bootstrap method. We first used the 150 MHz model from the TIFR GMRT Sky Survey Alternative Data Release (TGSSADR; Intema et al. [2016]) to start self-calibration for our 150 MHz observations, and then used our deeper 150 MHz model as the initial model for our 325 MHz self-calibration, likewise using the 325 MHz image for the 610 MHz initial sky model. The TGSSADR provides an excellent starting model for self-calibration in this lower frequency range and is available to the public¹.

We performed multi-scale (`multiscale` = (0,4,8,32) pixels) multi-frequency deconvolution with CASA on the SPAM-calibrated products with a slightly uniform Briggs weighting [`robust` = -0.3, Briggs, 1995] and four pixels per beam. We also corrected for a non-coplanar array using w -projection [Cornwell et al., 2008]. The multi-scale clean avoids the issues related to representing resolved diffuse emission with a sum of point-like sources, and this choice of weighting balances beam uniformity in side lobes with beam resolution. The highest minimum baseline of the three frequencies is 120λ , corresponding to a highest sensitivity scale of $\sim 20'$.

¹<http://tgssadr.strw.leidenuniv.nl/>



(a) 150 MHz contours (red) and 610 MHz contours (blue). The 150 MHz beam is $43.3'' \times 18.9''$, and 610 MHz is $7.2'' \times 4.9''$. (b) 325 MHz contours (red) and 610 MHz contours (blue). The 325 MHz beam is $17.5'' \times 9.5''$, and 610 MHz is $7.2'' \times 4.9''$.

Figure 1.1: Radio contour maps and corrected *XMM-Newton* EPIC image map. 150 MHz and 325 MHz contours (red) overlaid on the 610 MHz contours (blue), with radio beams in the bottom left corner. Contours are at $(5, 10, 20, 40, 80, 160)\sigma_{\text{rms}}$ levels. Thinner dashed contours mark $-5\sigma_{\text{rms}}$ (they are only noticeable for 150 MHz). Black crosses show GMOS spectroscopic cluster members within the redshift range $z = 0.516 \pm 0.015$. R_{500} is shown by the black circle.

The background root-mean-squared noises for the individual radio maps (**Figures 1.1a** and **1.1b**), σ_{rms} , were calculated as the standard deviation of the residual after Gaussian source fitting and subtraction using Python Blob Detection and Source Measurement (PyBDSM) v1.8. Our reduction yielded $\sigma_{\text{rms}} = 1.40 \text{ mJy beam}^{-1}$ at 150 MHz, $\sigma_{\text{rms}} = 120 \text{ } \mu\text{Jy beam}^{-1}$ at 325 MHz, and $\sigma_{\text{rms}} = 90 \text{ } \mu\text{Jy beam}^{-1}$ at 610 MHz. The resolutions of the maps are $43.3'' \times 18.9''$, $17.5'' \times 9.5''$, and $7.2'' \times 4.9''$, respectively. We note that the side lobes from a nearby bright source ($S_{1.4 \text{ GHz}} = 230 \text{ mJy}$ from the NRAO VLA Sky Survey) cause a spatially varying background noise in our target field.

1.2.2 XMM-Newton reduction

PLCK G004.5-19.5 was observed by *XMM-Newton* on 23 March, 2010 (ObsID: 0656201001), for a total duration of 14.2 ks. We reduced the EPIC MOS 1, MOS 2, and pn data using the XMM Science Analysis Software (SAS) v14. We ran the standard pipeline tasks `emproc` and `epproc` for MOS and pn data, respectively. For each of the three detectors, we filtered the data for solar flare events, following the same method as described in Mernier et al. [2015]. In the 10–12 keV band of MOS and the 12–14 keV band of pn, we fitted count-rate histograms binned in 100 s intervals with a Poissonian curve. We excluded all the time intervals in which the count-rate lay above $\mu + 2\sqrt{\mu}$, where μ is the mean of the Poissonian distribution. We repeated the procedure for the 0.3–10 keV band in the three instruments with histograms binned in 10 s intervals, since De Luca and Molendi [2004] reported that soft flare events can also affect soft X-ray bands. As recommended by the calibration reports, we kept only the highest quality events (`flag=0`) for the three instruments. Only single events were allowed in pn (`pattern=0`), while we allowed single, double, triple, and quadruple events in MOS (`pattern≤12`).

We extracted the MOS 1, MOS 2, and pn images of PLCK G004.5-19.5 within four distinct energy bands (0.3–2 keV, 2–4.5 keV, 4.5–7.5 keV, and 7.5–12 keV), and extracted similar “background” images, taken from stacked filter-wheel closed observations². We scaled the average count-rate of the “background” images to the average count-rate of the EPIC images within 10–12 keV, where negligible cluster emission is expected. The EPIC images in each band, and for each instrument, were then subtracted from this instrumental background and corrected for vignetting effects, after dividing them by their respective exposure maps generated using `eexppmap`.

The background- and vignetting-corrected images were then combined into one full EPIC image which can be seen in Figure 1.2.

1.2.3 Spectroscopic data

We complement our X-ray and radio analysis with spectroscopic measurements of galaxy redshifts derived from two sets of observations. In addition to the six redshifts published by Sifón et al. [2014], we retrieved spectroscopic observations performed with the Focal reducer and low-dispersion spectrograph (FORS2) on the Very Large Telescope (VLT; Program ID: 090.A-0925(A), PI: H. Böhringer) as part of a spectroscopic follow-up of *Planck* $z > 0.5$ cluster candidates. We reduced these data with the ESO REFLEX pipeline, which performs standard reduction steps such as bias subtraction, flat-field correction and wavelength calibration

²XMM-Newton SOC website (<http://xmm.esac.esa.int>).

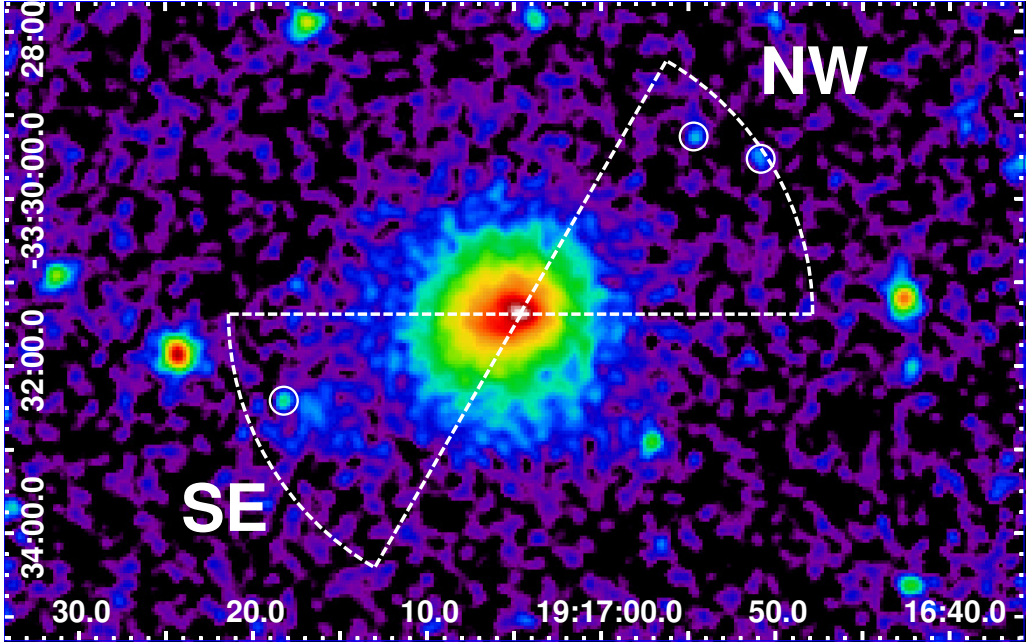


Figure 1.2: *XMM-Newton* EPIC (MOS+pn) background- and exposure-map corrected image of PLCK G004.5-19.5 (0.3–12 keV). The point-spread function of the EPIC cameras is $\approx 6''$. Overlaid are the extraction regions (wedges) and point-source masks (circles) used for the radial profile measurement.

[Freudling et al., 2013]. We measured galaxy redshifts by cross-correlating the resulting spectra with template galaxy spectra from the Sloan Digital Sky Survey using the RVSAO software [Kurtz and Mink, 1998]. Combining these two data sets, we find 16 galaxies in the redshift range 0.50–0.53 (corresponding to a range of $3,000 \text{ km s}^{-1}$) that are isolated in redshift space and have a velocity dispersion of $\sim 1200 \text{ km s}^{-1}$. The cluster redshift from the combined 16 galaxies is $z_{\text{cl}} = 0.519$.

1.3 Results and discussion

1.3.1 X-ray structure

Our processed EPIC image of PLCK G004.5-19.5 (**Figure 1.2**) suggests that the cluster is disturbed. Assuming $z = 0.516$, we find an offset $\sim 100 \text{ kpc}$ between the X-ray peak and the X-ray centroid, estimated from a circular aperture of $R_{\text{ap}} = 500 \text{ kpc}$ on the X-ray peak.

The degree of cluster disturbance can be provided in a more quantitative way, by estimating at least two morphological parameters [described in detail in Cassano et al., 2010b].

1. The centroid shift, w [Poole et al., 2006, Maughan et al., 2008], is estimated from fitting 2D β -profiles in a series of N circular apertures $\xi_i R_{\text{ap}}$ (with $0.05 \leq \xi_i \leq 1$) centred on

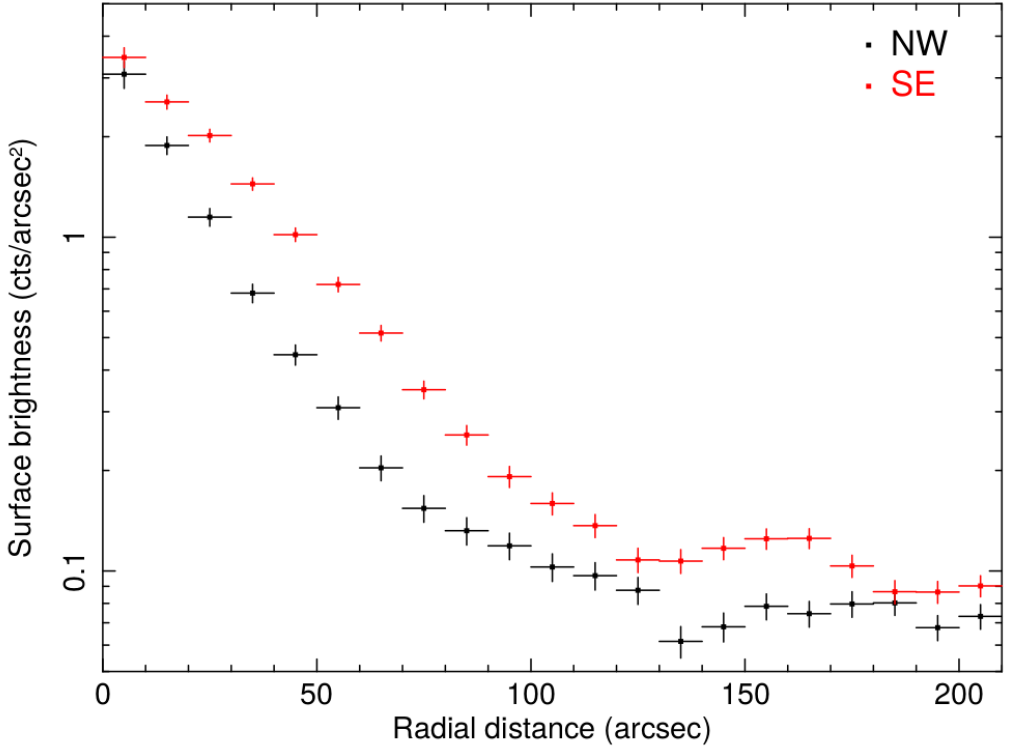


Figure 1.3: Radial surface brightness profile of *XMM-Newton* EPIC background- and exposure-map corrected image of PLCK G004.5-19.5 (0.3–12) keV in the NW and SE directions, centred on the X-ray brightness peak.

the X-ray peak, and can be expressed as

$$w = \frac{1}{R_{\text{ap}}} \times \left[\frac{1}{N-1} \sum_{i=1}^N (\Delta_i - \langle \Delta \rangle)^2 \right]^{1/2}, \quad (1.1)$$

where $\Delta_i - \langle \Delta \rangle$ is the difference between the X-ray peak and the centroid estimated in the i th aperture. Although Cassano et al. [2010b] proceeded from $0.05R_{\text{ap}}$ to R_{ap} in 5% steps, here the larger PSF of EPIC does not allow us to explore the image with such precision. Therefore, here we chose larger aperture radii of $0.2R_{\text{ap}}$, $0.4R_{\text{ap}}$, $0.6R_{\text{ap}}$, $0.8R_{\text{ap}}$, and R_{ap} .

2. The concentration parameter, c [Santos et al., 2008], is defined as the ratio of the peak over the ambient surface brightness S :

$$c = \frac{S(r < 100 \text{ kpc})}{S(r < 500 \text{ kpc})}. \quad (1.2)$$

Based on our EPIC image, we find $w = 0.058 \pm 0.007$ and $c = 0.079 \pm 0.004$. Referring to **Figure 1a** of Cassano et al. [2010b], it appears that PLCK G004.5-19.5 is situated within the

lower-right quadrant where disturbed galaxy clusters are found that host radio halos. This suggests that PLCK G004.5-19.5 is disturbed, although we note that PLCK G004.5-19.5 lies further in redshift than all of the clusters used in the derivation of the relation.

We note a small nodule of previously unidentified X-ray emission to the south-east of the main X-ray profile. A clustering of red elliptical galaxies, shown in **Figure 1.5a**, as well as an associated spectroscopic member, suggests that this nodule is at the cluster redshift. Assuming the nodule is purely thermal emission, with kT ranging from (3–10) keV, we find unabsorbed luminosities in the (0.1–2.4) keV energy band, within a circular region of $31''$ (192 kpc) and excluding a point-like source and main cluster X-ray emission³,

$$\begin{aligned} L_X^{\text{MOS1}} &= (6.7 \pm 1.4) \times 10^{43} \text{ erg s}^{-1} \\ L_X^{\text{MOS2}} &= (5.8 \pm 1.3) \times 10^{43} \text{ erg s}^{-1}, \end{aligned}$$

These luminosities are consistent with typical values for small galaxy groups and sub-clusters [Reiprich and Böhringer, 2002]. Taking the same aperture, $31''$, we calculate a main-cluster luminosity of $(9.2 \pm 1.4) \times 10^{44} \text{ erg s}^{-1}$. This is 14–16 times brighter than the luminosity of south-eastern nodule.

In **Figure 1.3** we show the radial X-ray (0.3–12) keV surface brightness profile of PLCK G004.5-19.5 in the south-east and north-west, centred on the X-ray brightness peak (see the mask in **Figure 1.2**). We see a significant uniform over-brightness in the south-east relative to north-west. This supports that the ICM is disturbed. We also note an excess located $160''$ to the south-east that is associated with the X-ray nodule.

1.3.2 Radio emission

We are able to detect several diffuse and compact radio sources, which are labelled in bold in Figures 1.1a and 1.1b. We performed Gaussian source extraction using PyBDSM v1.8.7 with a detection threshold of $5\sigma_{\text{rms}}$. For point sources lacking a detection at one frequency, we measured the flux within an aperture of one beam. For one resolved source, we measured the flux within a common aperture, and we discuss this procedure below. The flux measurements are given in Table 1.1, with nondetection measurements labelled accordingly.

We take the flux variance to be the sum of the measurement variance and a systematic variance equal to 15% of the flux, that is, $\sigma_S^2 = \sigma_{\text{rms}}^2 A/A_{\text{beam}} + (0.15S)^2$, where S , A , and A_{beam} are the measured flux, source area, and beam area.

We modelled the flux with two parameters, the spectral index α and $S_{1.4 \text{ GHz}}$, according to

$$S(\nu) = S_{1.4 \text{ GHz}} \left(\frac{\nu}{1.4 \text{ GHz}} \right)^\alpha. \quad (1.3)$$

We used the Metropolis-Hastings algorithm to sample the posterior distribution of the model. This is a common straightforward way to sample from the posterior distribution of a model given only knowledge of the a priori distribution and a method of computing the observables. We report the resulting posterior model parameters and uncertainties in Table 1.1 with Gaussian α and log-normal $S_{1.4 \text{ GHz}}$. The resulting model for each source with 1σ bands is shown in **Figure 1.4**.

³The south-east nodule falls between two CCD chips for the EPIC pn instrument, therefore we were unable to calculate its luminosity.

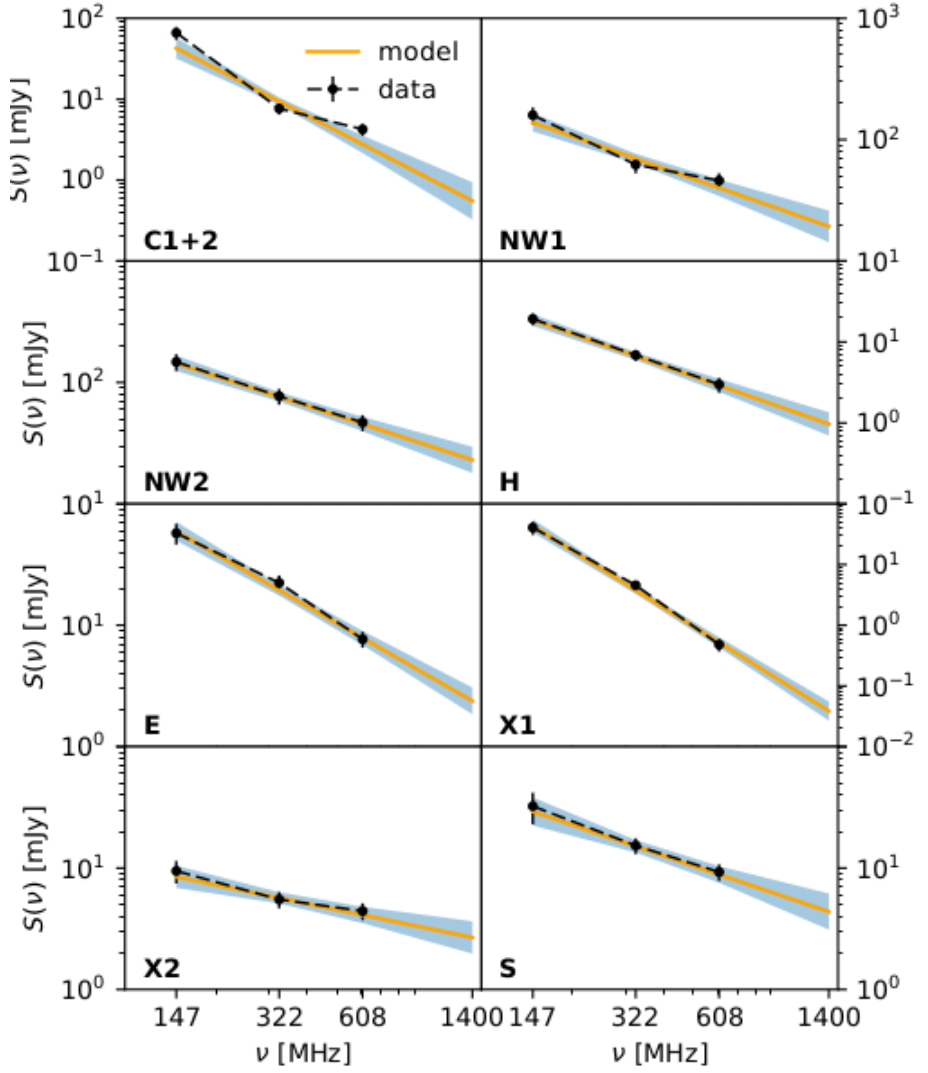


Figure 1.4: The measured fluxes, $S(\nu)$, for the sources with are in blue. The mean posterior model fluxes are with orange dashed lines with 1σ shaded intervals.

Table 1.1: Integrated spectral fluxes, spectral index, and rest-frame $P_{1.4 \text{ GHz}}$ with Λ CDM and $z = 0.516$.

Source	$S_{150 \text{ MHz}}$ (mJy)	$S_{325 \text{ MHz}}$ (mJy)	$S_{610 \text{ MHz}}$ (mJy)	α	$S_{1.4 \text{ GHz}}$ (mJy)	$P_{1.4 \text{ GHz}}$ (10^{24} W/Hz)	Nature
C1+2	65 ± 10	8 ± 1	4.2 ± 0.7^b	-1.9 ± 0.4	$0.6^{+0.4}_{-0.2}$	$0.6^{+0.3}_{-0.2}$	lobes?
NW1	160 ± 30	62 ± 9	46 ± 7	-0.9 ± 0.2	19^{+7}_{-5}	13^{+3}_{-3}	relic?
NW2	150 ± 20	75 ± 10	47 ± 7	-0.8 ± 0.2	23^{+7}_{-5}	15^{+3}_{-3}	head-tail
H	10 ± 2^c	4.0 ± 0.7^c	1.2 ± 0.5^c	-1.2 ± 0.4	$0.7^{+0.7}_{-0.3}$	$0.5^{+0.4}_{-0.2}$	halo?
E	55 ± 10	22 ± 3	8 ± 1	-1.4 ± 0.2	$2.4^{+0.7}_{-0.6}$	$2.0^{+0.4}_{-0.4}$	head-tail
X1	40 ± 10	4.6 ± 0.7	0.5 ± 0.1^a	-3.1 ± 0.3	$0.04^{+0.02}_{-0.01}$	$0.07^{+0.02}_{-0.02}$	AGN?
X2	10 ± 2^a	5.5 ± 0.9	4.4 ± 0.7	-0.5 ± 0.2	$2.7^{+1.0}_{-0.7}$	$1.6^{+0.4}_{-0.3}$	point?
S	30 ± 10	15 ± 2	9 ± 1	-0.8 ± 0.2	4^{+2}_{-1}	$3.0^{+0.8}_{-0.6}$	head-tail

^a Point-source, peak flux below detection limit $5\sigma_{\text{rms}}$, and aperture flux measurement.

^b Both resolved at 610 MHz; $S_{610 \text{ MHz}}^{\text{C1}} = 2.4 \pm 0.2 \text{ mJy}$ and $S_{610 \text{ MHz}}^{\text{C2}} = 1.9 \pm 0.2 \text{ mJy}$.

^c Resolved source, measurement within $17''$ circular aperture.

We labelled three sources S, X1, and X2 near R_{500} . Source S has an extended radio morphology that is indicative of a tailed radio galaxy, where the tail direction implies its projected relative velocity with respect to the ICM [see for e.g. Miley et al., 1974]. The tail of source S implies that it is moving away from the projected cluster centre. It has a spectral index of -0.8 ± 0.2 , which is in line with a radio lobe.

Sources X1 and X2 are situated above the X-ray nodule of the SE sub-cluster, and they have quite differing spectral indices of -3.1 ± 0.3 and -0.5 ± 0.2 , respectively. Figure 1.5a shows the optical region around the sources. Red elliptical galaxies cluster strongly around source X1. This suggests that X1 is associated with one of the giant red elliptical galaxies that make up the sub-cluster, and its steep spectral index suggests that it might be a merger-reactivated AGN.

To the east of the main cluster we have labelled an extended source, E, which has a tailed morphology similar to source S, and also a spectroscopically confirmed cluster member source near the radio surface brightness peak. The spectral index of source E is -1.4 ± 0.2 , which suggests that there is some significant electron ageing. This suggests that there might be some merger-induced reacceleration of aged electrons from a previous era of activity, or alternatively, some flux contamination from the central cluster region. The orientation of the tail implies that source E is moving towards the projected centre of the main cluster.

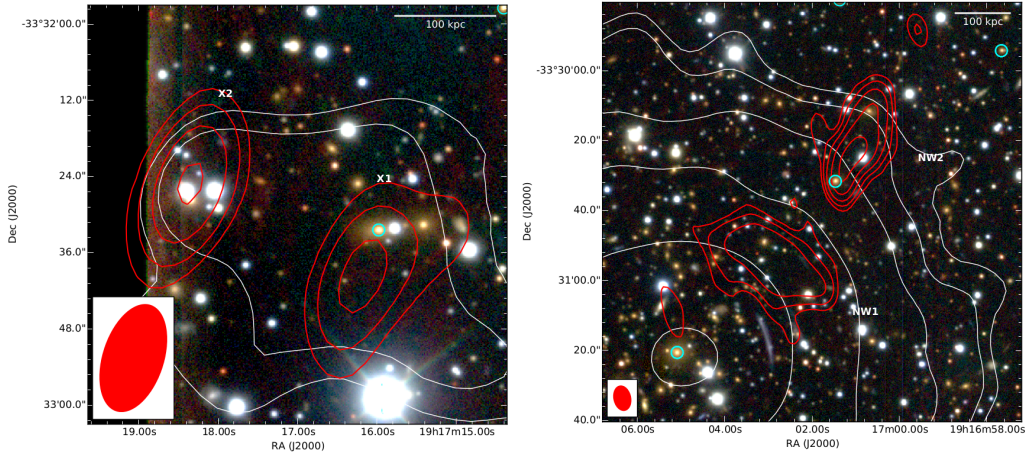
Two compact sources, labelled C1 and C2, are located between source E and the cluster centre. At 610 MHz, both patches are resolved and have similar fluxes of 2.4 ± 0.2 mJy and 1.9 ± 0.2 mJy, respectively. However, they are unresolved at 150 MHz and 325 MHz. We therefore calculated the integrated spectral index of the sum of the two sources (in Table 1.1 as C1+2), and found it to be very steep, -1.9 ± 0.4 . There are spectroscopic cluster members near the two brightness peaks of C1 and C2, although it is unclear whether they are counterparts, as there are also galaxies between C1 and C2 that are not confirmed as cluster members. The steep spectral index supports that C1 and C2 are lobes of a galaxy.

To the north-west of the main cluster are two extended sources labelled NW1 and NW2. Figure 1.5b shows a zoom of this region with 610 MHz contours over the *Gemini gri*-colour image.

Source NW2 has an orientation perpendicular to the X-ray profile, with the largest linear scale (LLS) of $40''$ (250 kpc). Its southernmost 610 MHz contour is $62''$ (380 kpc) away from the brightest cluster galaxy (BCG; 19h17m05.09s, $-33^{\circ}31'20.9''$). There is a slight brightening at the southern tip and a nearby spectroscopically confirmed cluster member (a giant red elliptical galaxy, shown by the cyan circle in **Figure 1.5b**). The integrated spectral index of NW2 is -0.8 ± 0.2 .

Source NW1 is oriented along the tangent of the X-ray profile, with LLS of $44''$ (270 kpc). Its nearest 610 MHz contour is $28''$ (170 kpc) from the BCG. The surface brightness along its length is uniform, with little variation compared to the brightness of source NW2. The integrated spectral index is -0.9 ± 0.2 .

In the central region, at 150 MHz and 325 MHz, there is a diffuse bridge of radio emission between the north-western and eastern parts of the cluster, which we have labelled H because its central location and extended morphology suggest that it is a radio halo. This source is noticeably extended in a region containing other compact sources (C1 and C2), therefore we did not use Gaussian source extraction to measure fluxes for H. After fitting all other sources with Gaussians and subtracting them from the field we took the flux within a $17''$ circular aperture.



(a) 320 MHz contours (red).

(b) 610 MHz contours (red). We also show a strong lensing arc south of NW1. The brightest cluster galaxy is the southernmost cyan circle.

Figure 1.5: *Gemini gri*-colour image with contours at $(5, 10, 20, 40, 80, 160)\sigma_{\text{rms}}$, and GMOS spectroscopic sources as cyan circles. X-ray contours are white.

We chose this aperture for two reasons. The first is that we required the aperture to fit completely within the region between C1+2 and NW1 to avoid flux contamination. In the 150 MHz map, the angular separation between C1+2 and NW1 is approximately $23''$ because of the lower resolution. At 325 MHz the separation is approximately $45''$. The second reason is that we wish to have at least one beam of collecting area at all frequencies. The aperture area is 900 arcsec^2 , which is equivalent to 1 beam at 150 MHz, 5 beams at 325 MHz, and 23 beams at 610 MHz. These criteria fixed the circular aperture size to $17''$. The spectral index of this extended source is measured to be -1.2 ± 0.4 .

1.3.3 Nature of the diffuse emission

In this disturbed galaxy cluster, the extended morphology and tangential orientation of source NW1 is in line with a radio relic tracing a shock front, and its spectral index is consistent with recently accelerated electrons. The X-ray data exposure time, 14.2 ks, is too short to detect a counterpart temperature jump, and there are no polarisation data. The projected distance of NW1 from the BCG is relatively small for a radio relic [Vazza et al., 2012], since shock structure is typically found farther out in the cluster periphery. A possible explanation is that we are viewing this system from a partially head-on perspective. In addition to the creation of relics and halos, shocks in the ICM can revive fossil radio plasma from a previous episode of activity [Enßlin and Gopal-Krishna, 2001] which can have filamentary structure. On the other hand, we might expect the spectra of NW1 to be steeper if this were the case [see e.g. sources in van Weeren et al., 2011].

Source NW2 has a morphology similar to sources E and S, with a tail direction implying it is an infalling cluster member. The spectroscopic cluster member could be related to the

slight brightening in the southern tip of NW2. We remark that the red elliptical galaxy at the northern tip of NW2 is likely also a cluster member that might likewise contribute to the extended emission.

The diffuse source H is consistent with a radio halo, in line with the evidence that these sources are found in dynamically disturbed systems. Current data suggest that the halo has a spectrum of -1.2 ± 0.4 , although higher sensitivity observations are required to derive better constraints. Measuring halo spectra at this redshift is very important since at $z > 0.5$ IC losses are expected to dominate. Simple modelling predicts that only one-third of very massive $\sim 1 - 2 \times 10^{15} M_{\odot}$ merging systems at redshift ~ 0.5 (such as PLCK G004.5-19.5) are predicted to have spectra flatter than $\alpha = -1.5$, with the majority of cases generating halos with much steeper spectra [Cassano et al., 2010a].

Radio halos can also, if less frequently, have filamentary structure [e.g. Govoni et al., 2005]. Since NW1 has a rather small projected distance on the sky from the BCG, and since we have no polarisation data and not enough data for an X-ray shock, we cannot rule out that NW1 is part of the radio halo that has survived IC losses. However, we rate this as unlikely since we expect IC losses, which scale as $(1+z)^4$, to heavily diminish the flux at high frequencies, which leads to the USS halo nature. The brightness of NW1 at 610 MHz indeed seems to argue against this. Furthermore, Sifón et al. [2014] found very bright emission in the region of NW1 in archival 1.4 GHz NVSS. Given that PLCK G004.5-19.5 is at a mean redshift of 0.516, we would not expect remnant halo flux to be bright at higher frequencies.

1.4 Conclusion

We have performed high-resolution 610 MHz, 325 MHz, and 150 MHz observations of the cluster PLCK G004.5-19.5 with the GMRT. We found a complex arrangement of central and peripheral diffuse emission, as well as several head-tail galaxies.

We reprocessed archival *XMM-Newton* data and found morphological parameters that imply that PLCK G004.5-19.5 is a disturbed system [Cassano et al., 2010b]. Furthermore, a previously unidentified south-eastern sub-cluster is observed with X-ray unabsorbed luminosity $(6.7 \pm 1.4) \times 10^{43} \text{ erg s}^{-1}$, approximately a factor 14–16 lower than the luminosity of their main cluster in the same aperture.

Most remarkably, we found a partially face-on radio relic (NW1) and centrally located radio halo (H) with spectral index (-1.2 ± 0.4) . Our interpretation of this finding favours the dynamically young state suggested by the X-ray morphology parameters, as well as the morphology of the radio emission.

PLCK G004.5-19.5 is the fifth cluster beyond $z = 0.5$ known to host a radio relic, and it is one of the few known instances of a ($z > 0.5$) radio halo.

The book cover features a watercolor-style illustration. At the top, a large, pale, cloudy moon or planet dominates the upper half. Below it, a central, bright yellow sun or star is partially obscured by a large, purple, irregular shape. The background is a deep blue with lighter blue and purple washes. On either side of the central purple shape, there are several smaller, blue, circular objects, some of which are connected by thin white lines, suggesting a network or constellation. The title 'COSMOS' is written in a large, red, gothic-style font at the top, with 'Two' in a smaller, black, gothic-style font below it. The subtitle 'A PROBABILISTIC APPROACH TO DIRECTION-DEPENDENT IONOSPHERIC CALIBRATION' is written in a smaller, white, sans-serif font below the title. The authors' names 'J.G. ALBERT', 'M.S.S.L. OEI', 'R.J. VAN WEEREN', 'H.T. INTEMA', and 'H.J.A. RÖTTGERING' are listed in a white, sans-serif font at the bottom.

COSMOS

Two

A PROBABILISTIC APPROACH
TO DIRECTION-DEPENDENT
IONOSPHERIC
CALIBRATION

J.G. ALBERT
M.S.S.L. OEI
R.J. VAN WEEREN
H.T. INTEMA
AND
H.J.A. RÖTTGERING

A probabilistic approach to direction-dependent ionospheric calibration

J. G. ALBERT, M. S. S. L. OEI, R. J. VAN WEEREN, H. T. INTEMA, AND H. J. A. RÖTTGER-ING

Published in A&A

Received 12 April 2019 / Accepted 10 October 2019

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.

Acknowledgements. J. G. A. and H. T. I. acknowledge funding by NWO under 'Nationale Roadmap Grootchalige Onderzoeksfaciliteiten', as this research is part of the NL SKA roadmap project. J. G. A. and H. J. A. R. acknowledge support from the ERC Advanced Investigator programme NewClusters 321271. R. J. vW. and M. S. S. L. O. acknowledge support of the VIDI research programme with project number 639.042.729, which is financed by the Netherlands Organisation for Scientific Research (NWO). M. S. S. L. O. thanks Jesse van Oostrum for helpful discussions.

If we know that our individual errors and fluctuations follow the magic bell-shaped curve exactly then the resulting estimates are known to have almost all the nice properties that people have been able to think of.

JOHN W. TUKEY, 1965

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 minimize risk – could be fruitful and disrupt the status quo [Wu et al., 2019]. The second approach is to treat the piecewise independent calibration solutions as data and predict calibration solutions for missing geodesics [e.g. Intema et al., 2009, 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 and 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 $O(\nu^{-5})$:

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

Here $\nu_p(\mathbf{x}) = \left(\frac{n_e(\mathbf{x})q^2}{4\pi^2\epsilon_0 m} \right)^{1/2}$ 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* propagation 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 $\mathcal{R}_x^{\mathbf{k}}$ be a functional of n , so that the geodesic $\mathcal{R}_x^{\mathbf{k}}[n] : [0, \infty) \rightarrow \mathbb{R}^3$ 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

$$\phi_x^{\mathbf{k}} = \frac{2\pi\nu}{c} \int_0^\infty n\left(\mathcal{R}_x^{\mathbf{k}}[n](s)\right) - 1 \, ds, \quad (2.2)$$

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, $\phi_x^{\mathbf{k}} \approx \frac{-q^2}{4\pi\epsilon_0 mc\nu} \tau_x^{\mathbf{k}}$, where,

$$\tau_x^{\mathbf{k}} \triangleq \int_0^\infty n_e\left(\mathcal{R}_x^{\mathbf{k}}[n](s)\right) ds. \quad (2.3)$$

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 $\mathcal{R}_i^{\mathbf{k}}[n]$ is used as shorthand for $\mathcal{R}_{x_i}^{\mathbf{k}}[n]$. Correspondingly, we introduce the notion of differential total electron content (ΔTEC),

$$\tau_{ij}^{\mathbf{k}} \triangleq \tau_i^{\mathbf{k}} - \tau_j^{\mathbf{k}}, \quad (2.4)$$

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

2.3 Probabilistic relation between FED and Δ TEC: Gaussian process layer model

In this section we derive the probability distribution of Δ TEC 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 $G_i^{\mathbf{k}} : \mathcal{V} \rightarrow \mathbb{R}$ mapping from the space of all scalar-valued functions over $\mathbb{R}^3$ to a scalar value according to,

$$G_i^{\mathbf{k}} f \triangleq \int_0^\infty f(\mathcal{R}_i^{\mathbf{k}}[n](s)) ds, \quad (2.5)$$

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

$$\Delta_{ij}^{\mathbf{k}} f \triangleq (G_i^{\mathbf{k}} - G_j^{\mathbf{k}}) f. \quad (2.6)$$

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

$$\tau_{ij}^{\mathbf{k}} = \Delta_{ij}^{\mathbf{k}} n_e. \quad (2.7)$$

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

$$n_e \sim \mathcal{N}[\mu, K], \quad (2.8)$$

where $\mu : \mathcal{X} \rightarrow \mathbb{R}_{>0}$ is the mean function, and $K : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ 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 $G_i^{\mathbf{k}}$ well-defined. To further simplify the model, we assume that the mean FED in the layer is constant; 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

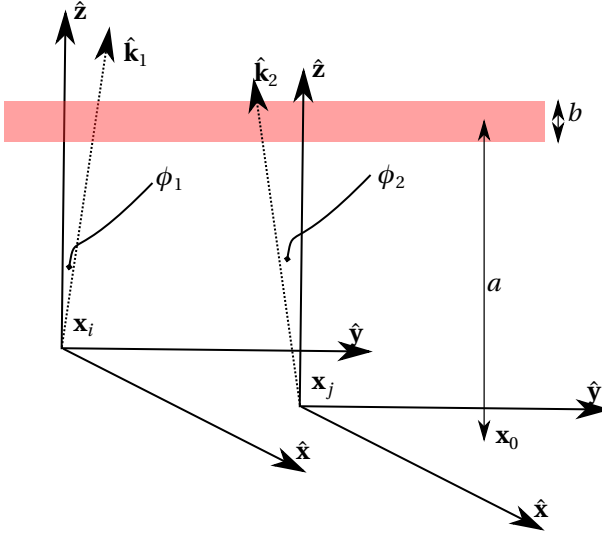


Figure 2.1: Geometry of the toy model. The ionosphere is a layer of thickness b at height a above a reference location $\mathbf{x}_0$. In general, ΔTEC is the TEC along one geodesic minus the TEC along another parallel geodesic. Usually, these geodesics are originating at antennae i and j (locations $\mathbf{x}_i$ 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 .

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 and Williams, 2006a].

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 $\rho(\mathbf{x}) \rightarrow 0$, 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 and 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 $\mathcal{R}_{\mathbf{x}}^{\hat{\mathbf{k}}}[n](s) = \mathbf{x} + s\hat{\mathbf{k}}$. 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 and Koopmans, 2015]. For frequencies $\lesssim 50$ MHz however, this simplification becomes problematic. Under the straight-ray assumption, Equation 2.7 becomes

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

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

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

where $\sec \phi = (\hat{\mathbf{k}} \cdot \hat{\mathbf{z}})^{-1}$ 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, $\mathcal{S} = \{(\mathbf{x}, \hat{\mathbf{k}}) \mid \mathbf{x} \in \mathbb{R}^3, \hat{\mathbf{k}} \in \mathbb{S}^2\}$. This is analogous to saying that the coordinates of the ΔTEC GP are a tuple of antenna location and calibration direction. Thus, given any $\mathbf{y} = (\mathbf{x}, \hat{\mathbf{k}}) \in \mathcal{S}$, the ΔTEC is denoted by $\tau_{x0}^{\hat{\mathbf{k}}}$. 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 $b \sec \phi$. This results in the mean TEC of two parallel rays being equal, and thus the first moment of ΔTEC is,

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

where $\mathbf{y} = (\mathbf{x}_i, \hat{\mathbf{k}}) \in \mathcal{S}$. 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 **Figure 2.1**.

$$K_{\Delta\text{TEC}}(\mathbf{y}, \mathbf{y}') = \mathbb{E} \left[\tau_{i0}^{\hat{\mathbf{k}}} \tau_{j0}^{\hat{\mathbf{k}}'} \right] \quad (2.12)$$

$$= \mathbb{E} \left[(G_i^{\hat{\mathbf{k}}} n_e - G_0^{\hat{\mathbf{k}}} n_e) (G_j^{\hat{\mathbf{k}}'} n_e - G_0^{\hat{\mathbf{k}}'} n_e) \right] \quad (2.13)$$

$$= I_{ij}^{\hat{\mathbf{k}}\hat{\mathbf{k}}'} + I_{00}^{\hat{\mathbf{k}}\hat{\mathbf{k}}'} - I_{i0}^{\hat{\mathbf{k}}\hat{\mathbf{k}}'} - I_{0j}^{\hat{\mathbf{k}}\hat{\mathbf{k}}'}, \quad (2.14)$$

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

$$I_{ij}^{\hat{\mathbf{k}}\hat{\mathbf{k}}'} = \int_{s_i^{\hat{\mathbf{k}}-}}^{s_i^{\hat{\mathbf{k}}+}} \int_{s_j^{\hat{\mathbf{k}}'-}}^{s_j^{\hat{\mathbf{k}}'+}} K(\mathbf{x}_i + s\hat{\mathbf{k}}, \mathbf{x}_j + s'\hat{\mathbf{k}}') ds ds'. \quad (2.15)$$

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 and 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.

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

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 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 and 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 **Figure 2.2**. For the dawn ionosphere we choose the Matérn-3/2 (M32) kernel,

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

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

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

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 and 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 $h \approx 1.177l_{\text{EQ}}$ for the EQ and $h \approx 0.969l_{\text{M32}}$ 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 perform a comparative analysis.

2.4.2 Competitor models

For the comparison with competitor models, we compare the physical kernel with: exponential quadratic (EQ), Matérn-5/2 (M52), Matérn-3/2 (M32), and Matérn-1/2 (M12) [Rasmussen and Williams, 2006a]. 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 $(\mathbf{x}, \hat{\mathbf{k}}), (\mathbf{x}', \hat{\mathbf{k}}') \in \mathcal{S}$, then we form the kernel $K_C((\mathbf{x}, \hat{\mathbf{k}}), (\mathbf{x}', \hat{\mathbf{k}}')) = K_C^1(\mathbf{x}, \mathbf{x}') K_C^2(\hat{\mathbf{k}}, \hat{\mathbf{k}}')$ thereby giving each subspace of the index set, $\mathcal{S}$, its own kernel with associated hyperparameters.

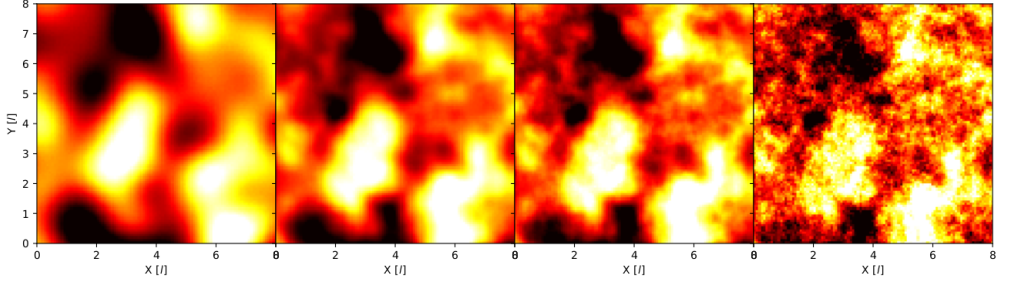


Figure 2.2: Example realisations from exponential quadratic, Matérn-5/2, Matérn-3/2, and Matérn-1/2 kernels. The same HPD was used in all kernels, however the smoothness of the resulting process realisation is different for each.

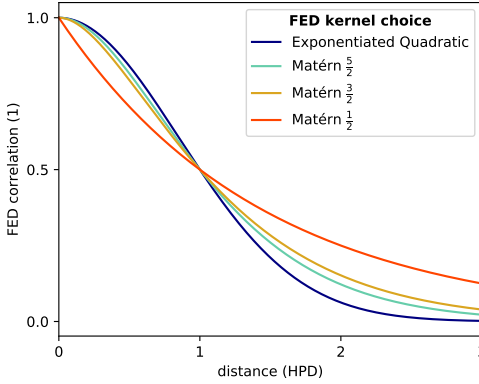


Figure 2.3: Shape of several kernels as a function of separation in units of the HPD of the kernel.

Figure 2.3 shows each kernel profile with the same HPD and Figure 2.2 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-Katta 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.

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 hyper parameters, 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 $\tau_{\text{obs}} = \tau(\mathbf{X}) + \epsilon$ where $\epsilon \sim \mathcal{N}[0, \sigma^2]$ and $\tau(\mathbf{X}) \sim \mathcal{N}[0, K(\mathbf{X}, \mathbf{X})]$ then the LPO is,

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

where $\mathbf{B} = K(\mathbf{X}, \mathbf{X}) + \sigma^2 \mathbf{I}$. 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 $\tau_{\text{obs}}^* = \tau(\mathbf{X}^*) + \epsilon^*$ with $\epsilon^* \sim \mathcal{N}[0, \sigma^2]$ then the LPH conditional on observed τ_{obs} is,

$$\begin{aligned} \log P_K(\tau_{\text{obs}}^* | \tau_{\text{obs}}) &= \log \mathcal{N}[K(\mathbf{X}^*, \mathbf{X}) \mathbf{B}^{-1} \tau_{\text{obs}}, \\ &\quad \mathbf{B}^* - K(\mathbf{X}^*, \mathbf{X}) \mathbf{B}^{-1} K(\mathbf{X}, \mathbf{X}^*)] \end{aligned} \quad (2.19)$$

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

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

$$\Delta\text{LPO}_C(\eta) \triangleq \frac{P_{\Delta\text{TEC}}(\tau_{\text{obs}} | \eta)}{P_C(\tau_{\text{obs}} | \eta)}, \quad (2.20)$$

$$\Delta\text{LPH}_C(\eta) \triangleq \frac{P_{\Delta\text{TEC}}(\tau_{\text{obs}}^* | \tau_{\text{obs}}, \eta)}{P_C(\tau_{\text{obs}}^* | \tau_{\text{obs}}, \eta)}, \quad (2.21)$$

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 $\Delta\text{LPO}_C(\eta)$ 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 and Williams, 2006a]. 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 and Williams, 2006a], 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 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 $\sigma_{n_e}^2$, and for the competitor kernels this corresponds to learning a variance and the length-scales for each subspace.

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)	$b\sigma_{n_e}$ (10^{11}km 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

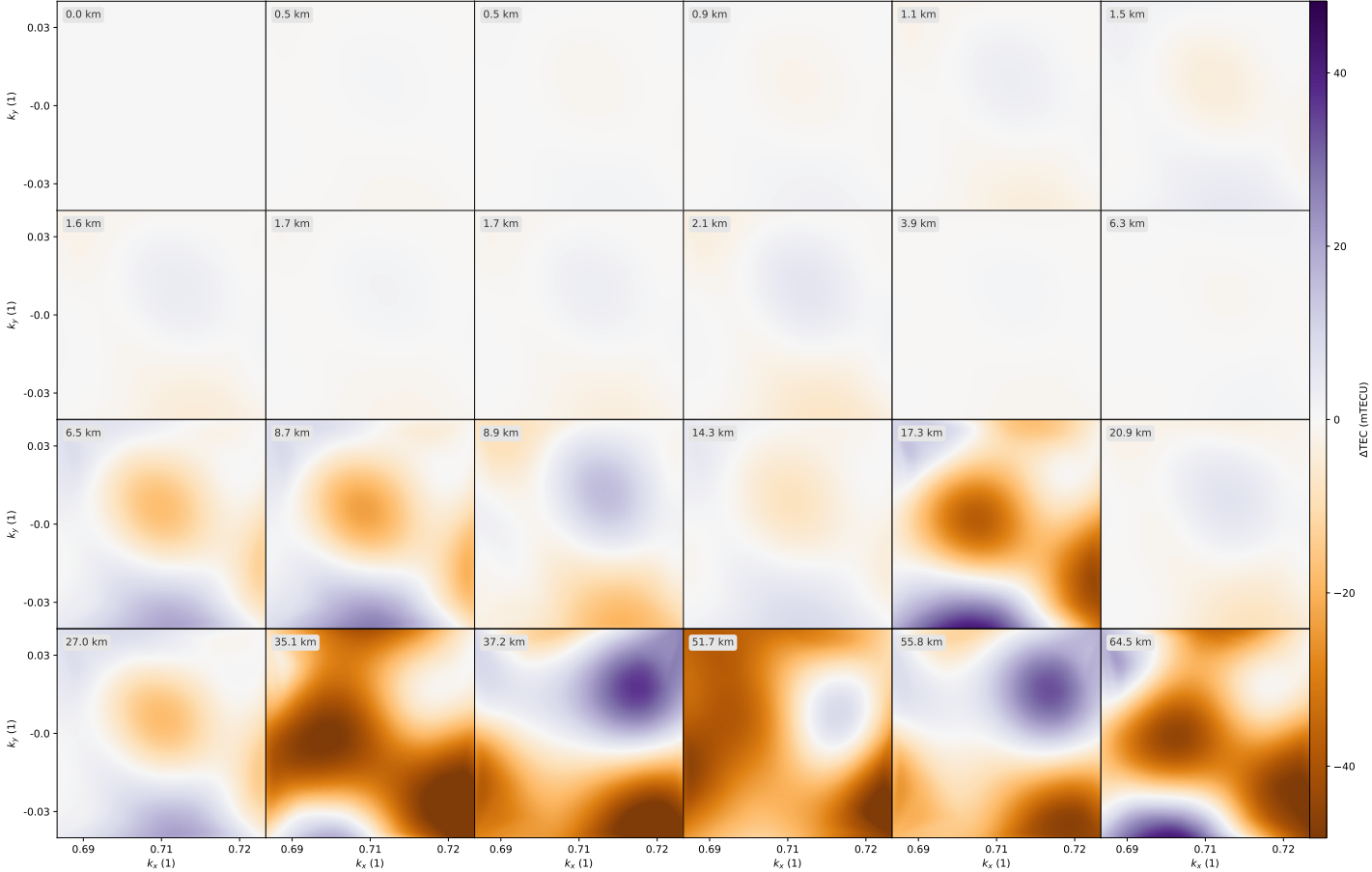


Figure 2.4: Caption continued on next page.

Figure 2.4: 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.

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 $\sim 10 \text{ km}$, 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 $b\sigma_{n_e}$, 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 $\Delta\text{LPO}_C(\eta)$, and $\Delta\text{LPH}_C(\eta)$ 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) $\langle\Delta\text{LPO}_C\rangle_\eta = 1.55$ and $\Delta\text{LPO}_C^{\eta_{\text{nom}}} = 1.46$, implying that the M52 kernel model is 55% and 46% 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, 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) $\langle\Delta\text{LPH}_C\rangle_\eta = 1.49$ and $\Delta\text{LPH}_C^{\eta_{\text{nom}}} = 1.31$, 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

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).

	$\langle \Delta \text{LPO}_C \rangle_\eta$	$\Delta \text{LPO}_C^{\eta_{\text{nom}}}$	$\langle \Delta \text{LPH}_C \rangle_\eta$	$\Delta \text{LPH}_C^{\eta_{\text{nom}}}$
	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

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 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 ,

$$\delta l \approx \frac{q^2}{\epsilon_0 m_e \nu^2 r} \delta \tau \quad (2.22)$$

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

Figure 2.5 shows the mean linear regression of the absolute equivalent source shift of the

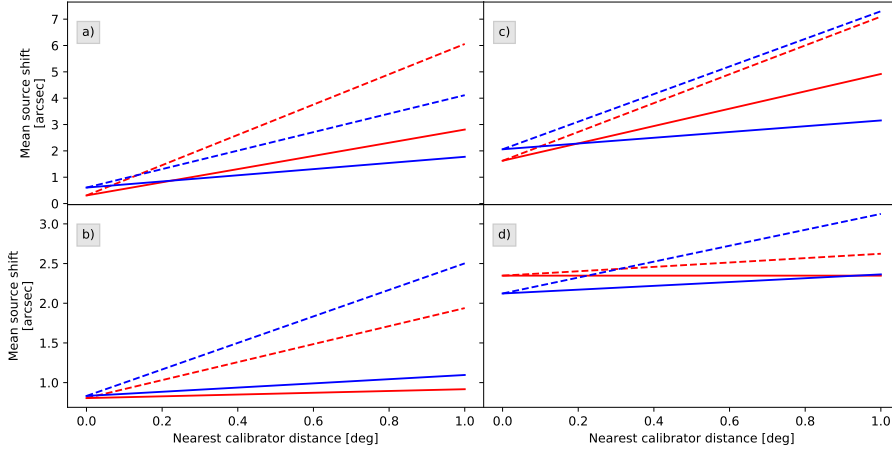


Figure 2.5: Mean equivalent source shift as a function of angular distance from the nearest calibrator caused by inference errors from the ground truth for **a)** remote stations (RS; > 3 km from the reference antenna) at nominal conditions (30 calibrators for 12.6deg^2 and 1 mTECU noise), **b)** core 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.

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 $\mathbf{B}^*$ 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 and 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 hyper parameters 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,

$$P(\boldsymbol{\tau} | \boldsymbol{\tau}_{\text{obs}}) = \int_{\mathbf{n}_e} P(\boldsymbol{\tau} | \mathbf{n}_e) P(\mathbf{n}_e | \boldsymbol{\tau}_{\text{obs}}) d\mathbf{n}_e, \quad (2.24)$$

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

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

$$P(\boldsymbol{\tau} | \boldsymbol{\tau}_{\text{obs}}) = \int_{\mathbf{n}_e} P(\mathbf{n}_e, \boldsymbol{\tau} | \boldsymbol{\tau}_{\text{obs}}) d\mathbf{n}_e \quad (2.25)$$

$$= \int_{\mathbf{n}_e} P(\mathbf{n}_e | \boldsymbol{\tau}_{\text{obs}}) P(\boldsymbol{\tau} | \mathbf{n}_e, \boldsymbol{\tau}_{\text{obs}}) d\mathbf{n}_e, \quad (2.26)$$

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 $P(\boldsymbol{\tau} | \mathbf{n}_e) = P(\boldsymbol{\tau} | \mathbf{n}_e, \boldsymbol{\tau}_{\text{obs}})$ is true, then our method is equivalent to first inferring FED and then using that distribution to

calculate ΔTEC . In Appendix A 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 $\mathbf{v}$, so that $n_e(\mathbf{x}, t) = n_e^0(\mathbf{x} - \mathbf{v}t)$. 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 structure function [van der Tol, 2009],

$$D(r) = \langle (\phi_{\mathbf{v}}(R) - \phi_{\mathbf{v}}(r + R))^2 \rangle_R \quad (2.27)$$

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

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

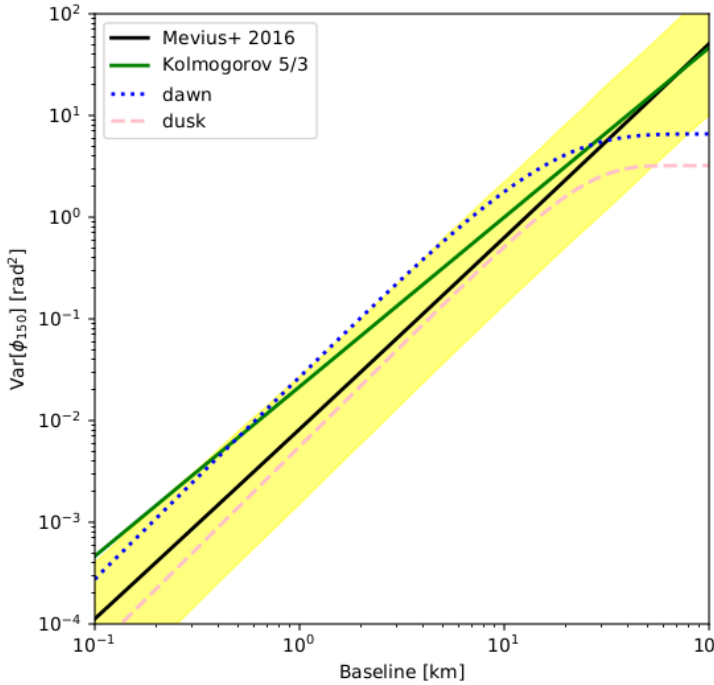


Figure 2.6: Structure functions predicted by our model compared with observations and theory. The dotted and dashed lines show the phase structure function corresponding to the physical kernel, with the dawn and dusk configurations, respectively (see **Figure 2.1**). Along side is the predicted structure 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] observes a turnover, but does not characterise it, and therefore we do not attempt to plot it here.

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 **Figure 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 B, 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 geometric 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

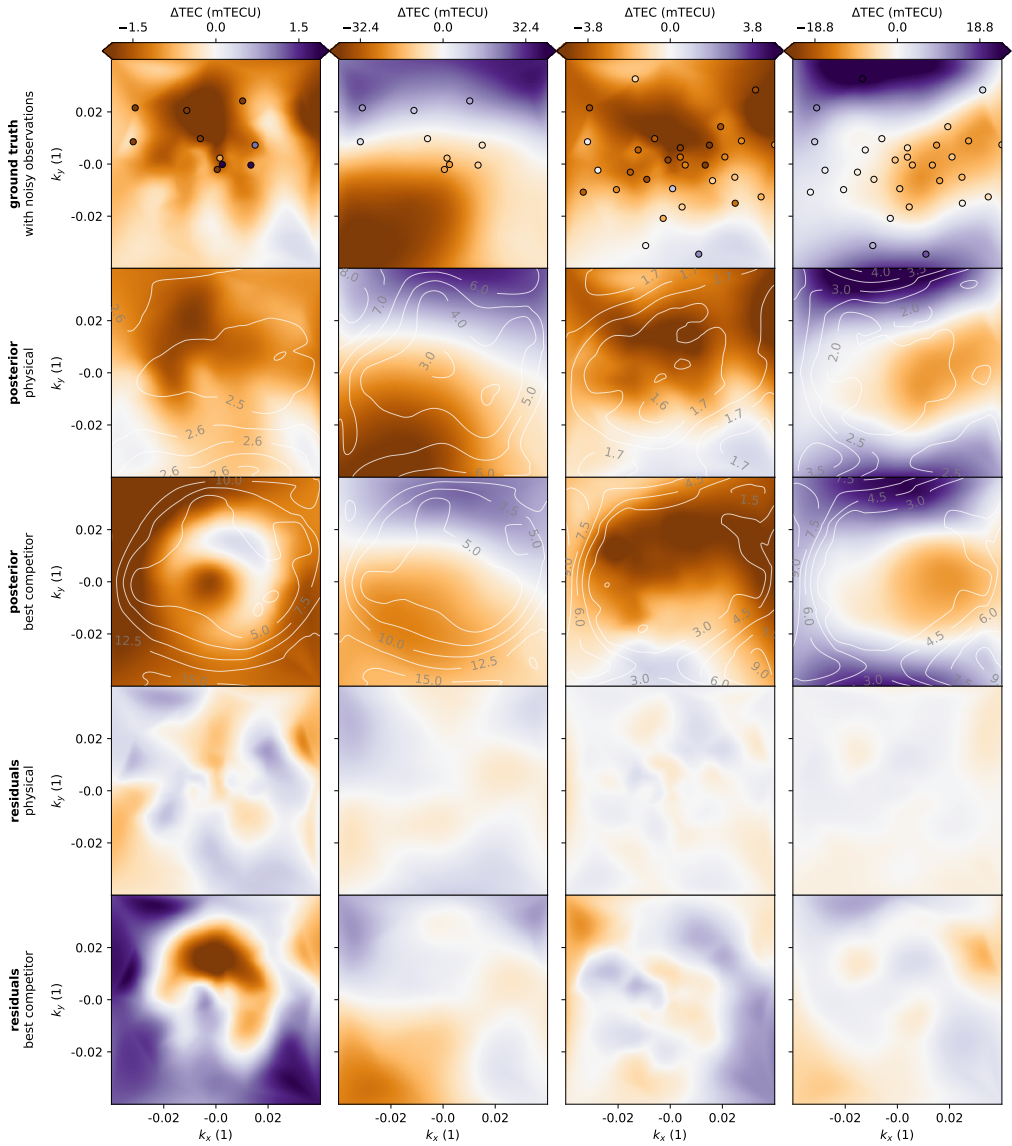


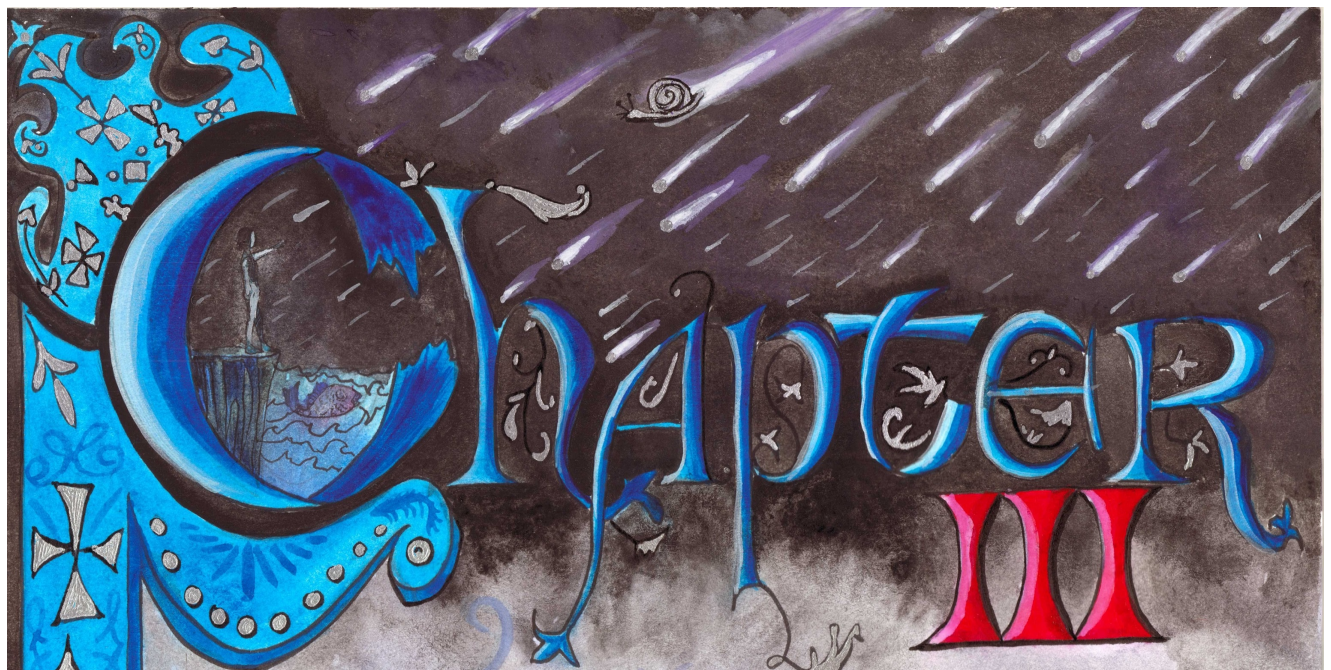
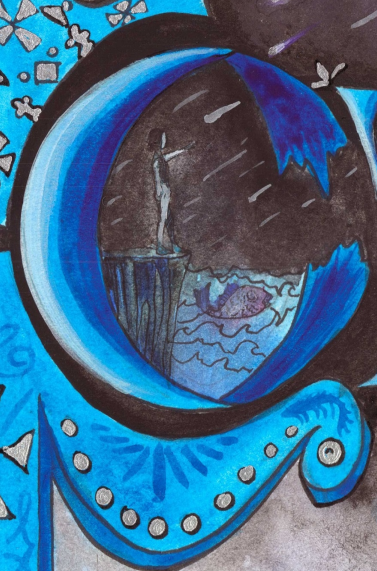
Figure 2.7: Caption continued on next page.

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 on noisy draws from the ground truth which are 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 experimental conditions, (10 directions, 2.5 mTECU noise), for a central antenna (near to reference antenna) and a remote station (far from reference antenna); *Last two columns*: Results for experimental conditions, (30 directions, 1.6 mTECU noise), for a central antenna and a remote station.

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 perform 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.



Chapter

III

Probabilistic Direction ~
Dependent Ionospheric
Calibration for LOSAR-KBA

J. G. Albert, R. J. Van Weeren,
H. T. Intema, and J. A. Röttgering

Probabilistic direction-dependent ionospheric calibration for LOFAR-HBA

J. G. ALBERT, R. J. VAN WEEREN, H. T. INTEMA, AND H. J. A. RÖTTGERING

Published in A&A

Received December 30, 2019 / Accepted January 31, 2020

Direction-dependent calibration and imaging is a vital part of producing radio images that are deep and have a high fidelity and highly dynamic range with a wide-field low-frequency array such as LOFAR. Currently, dedicated facet-based direction-dependent calibration algorithms rely on the assumption that the size of the isoplanatic patch is much larger than the separation between bright in-field calibrators. This assumption is often violated owing to the dynamic nature of the ionosphere, and as a result, direction-dependent errors affect image quality between calibrators. In this paper we propose a probabilistic physics-informed model for inferring ionospheric phase screens, providing a calibration for all sources in the field of view. We apply our method to a randomly selected observation from the LOFAR Two-Metre Sky Survey archive, and show that almost all direction-dependent effects between bright calibrators are corrected and that the root-mean-squared residuals around bright sources are reduced by 32% on average.

Acknowledgements. J. G. A. and H. T. I. acknowledge funding by NWO under ‘Nationale Roadmap Grootschalige Onderzoeksfaciliteiten’, as this research is part of the NL SKA roadmap project. J. G. A. and H. J. A. R. acknowledge support from the ERC Advanced Investigator programme NewClusters 321271. R. J. vW. acknowledges support of the VIDI research programme with project number 639.042.729, which is financed by the Netherlands Organisation for Scientific Research (NWO).

3.1 Introduction

In radio astronomy many of the greatest questions can only be answered by observing the faintest emission over a large fraction of the sky surface area. Such large puzzles include searching for the epoch of reionisation [e.g. Patil et al., 2017], potentially detecting the missing baryons in the synchrotron cosmic web [e.g. Vernstrom et al., 2017], understanding the dynamics of galaxy cluster mergers [e.g. van Weeren et al., 2019], and probing matter distribution with weak gravitational lensing at radio frequencies [e.g. Harrison et al., 2016]. In recent years, instruments such as the Low-Frequency Array [LOFAR; van Haarlem et al., 2013], the Murchison Widefield Array (MWA), and the future Square Kilometre Array (SKA) have been designed with the hope of discovering these key pieces in our understanding of the Universe. In particular, the prime objective of LOFAR, a mid-latitude array ($N52^\circ$), is to be a sensitive, wide-field, wide-band surveying instrument of the entire northern hemisphere at low frequencies, although it is important to note that LOFAR effectively enables a host of other scientific goals such as real-time monitoring of the northern sky at low frequencies [Prasad et al., 2016].

While these instruments take decades to plan and build, arguably most of the technological difficulty follows after the instrument comes online as scientists come to understand the instruments. Even though LOFAR came online in 2012, after six years of construction, it was not until 2016 that a suitable direction-dependent (DD) calibrated image was produced for its high-band antennae (HBA; 115–189MHz) [van Weeren et al., 2016], and although there has been much progress, there is currently still no DD calibrated image produced for its low-band antennae (LBA; 42–66MHz). Importantly, despite many years in developing these initial DD calibration and imaging algorithms, they were not suitable for the primary survey objective of LOFAR. Thus, there has been continuous work aimed at improving them.

The main cause of DD effects in radio interferometry is the ionosphere. The ionosphere is a turbulent, low-density, multi-species ion plasma encompassing the Earth, driven mainly by solar extreme ultra-violet radiation [Kivelson and Russell, 1995] that changes on the timescale of tens of seconds [Phillips, 1952]. The free electron density (FED) of the ionosphere gives rise to a spatially and temporally varying refractive index at radio wavelengths. This causes weak scattering of electric fields passing through the ionosphere [Ratcliffe, 1956], which becomes more severe at lower frequencies, and results in the scattering of radio radiation onto radio interferometers. Despite the weak-scattering conditions, the integral effect of repeated scattering through the thick ionosphere can cause dispersive phase variations of well over 1 radian on the ground [Hewish, 1951, 1952].

Wide-field radio arrays, such as LOFAR, are particularly susceptible to the DD effects of the ionosphere [Cohen and Röttgering, 2009] because the field of view is many isoplanatic patches [Fellgett and Linfoot, 1955] wherein the instantaneous scattering properties of the ionosphere change significantly over the field of view. In order to achieve the science goals in the wide-field regime, a DD calibration strategy must be adopted [Cornwell et al., 2005]. Many DD approaches have been developed, ranging from field-based calibration [Cotton et al., 2004] and expectation-maximisation [Kazemi et al., 2011] to facet-based approaches [Intema et al., 2009, van Weeren et al., 2016, Tasse et al., 2018].

Common to all of these DD techniques is the requirement that the signal-to-noise ratio needs to be high enough on short enough time intervals where the ionosphere can be considered fixed. Therefore, bright compact sources are required throughout the field of view,

which act as in-field calibrators. The fundamental assumption is that the distance between calibrators is less than the isoplanatic-patch size, which is about 1° for nominal ionospheric conditions. However, this quantity is dynamic over the course of an observation, and this assumption can be broken at times.

With the advent of the killMS DD calibration algorithm [Tasse, 2014b, Smirnov and Tasse, 2015], and the DDFacet DD imager [Tasse et al., 2018], the LOFAR Two-Metre Sky Survey [LoTSS; Shimwell et al., 2017] became possible, and the LoTSS first data release (DR1) has become available [Shimwell et al., 2019]. While LoTSS-DR1 provides an excellent median sensitivity of $S_{144\text{MHz}} = 71 \mu\text{Jy beam}^{-1}$ and point-source completeness of 90%, there are still significant DD effects in the images between the in-field calibrators. Improving the DD calibration for LoTSS is therefore a priority.

The second data release (DR2) is preliminarily described in Shimwell et al. [2019], and will be fully described in Tasse et al. (in prep.). Because we have internal access to this archive, we make reference to LoTSS-DR2.

The aim of this work is to present a solution to this limitation by probabilistically inferring the DD calibrations for all sources in a field of view based on information of the DD calibrations for a sparse set of bright calibrators. Importantly, this approach can be treated as an add-on to the existing LoTSS calibration and imaging pipeline.

We arrange this paper as follows. In Section 3.2 we formally define the problem of the DD calibration of ionospheric effects through inference of the doubly differential total electron content. In Section 3.3 we describe the procedure we used to perform the DD calibration and imaging of a randomly chosen LoTSS-DR2 observation using doubly differential total electron content screens. In Section 3.4 we compare our image with the LoTSS-DR2 archival image.

3.2 Doubly differential phase screens

In radio interferometry the observable quantity is the collection of spatial auto-correlations of the electric field, which are called the visibilities. The relation between the electric field intensity and the visibilities is given by the radio interferometry measurement equation [RIME; Hamaker et al., 1996],

$$V(\mathbf{x}, \mathbf{x}') = \sum_{\mathbf{k} \in \mathcal{K}} (\mathbf{B}(\mathbf{x}, \mathbf{k}) \mathbf{J}(\mathbf{x}, \mathbf{k}) \otimes \mathbf{J}^\dagger(\mathbf{x}', \mathbf{k}) \mathbf{B}^\dagger(\mathbf{x}', \mathbf{k})) (\mathbf{E}(\mathbf{k}) \otimes \mathbf{E}^\dagger(\mathbf{k})) \times h(\mathbf{x}, \mathbf{k}) h^*(\mathbf{x}', \mathbf{k}), \quad (3.1)$$

where we have left out frequency dependence for visual clarity, although for future reference, we let there be N_{freq} sub-bands of the bandwidth.

In order to understand Eq. 3.1, let the celestial radio sky be formed of a countably infinite number of discrete point sources with directions in $\mathcal{K} = \{\mathbf{k}_i \in \mathbb{S}^2 \mid i = 1..N_{\text{dir}}\}$. The propagation of a monochromatic polarised celestial electric field $\mathbf{E}(\mathbf{k})$ can be described by the phenomenological Jones algebra [Jones, 1941]. Letting the locations of antennae be $\mathcal{X} = \{\mathbf{x}_i \in \mathbb{R}^3 \mid i = 1..N_{\text{ant}}\}$, then $\mathbf{E}(\mathbf{x}, \mathbf{k}) = h(\mathbf{x}, \mathbf{k}) \mathbf{J}(\mathbf{x}, \mathbf{k}) \mathbf{E}(\mathbf{k})$ denotes propagation of $\mathbf{E}(\mathbf{k})$ to $\mathbf{x}$ through the optical pathway indexed by $(\mathbf{x}, \mathbf{k})$, where $h(\mathbf{x}, \mathbf{k})$ is the Huygens-Fresnel propagator in a vacuum. Each antenna has a known instrumental optical transfer function that gives rise to another Jones matrix known as the beam, $\mathbf{B}(\mathbf{x}, \mathbf{k})$. The total electric field measured by

an antenna at $\mathbf{x}$ is then $\mathbf{E}(\mathbf{x}) = \sum_{\mathbf{k} \in \mathcal{K}} h(\mathbf{x}, \mathbf{k}) \mathbf{B}(\mathbf{x}, \mathbf{k}) \mathbf{J}(\mathbf{x}, \mathbf{k}) \mathbf{E}(\mathbf{k})$, and Eq. 3.1 follows from considering the auto-correlation of the electric field between antennae and imposing the physical assumption that the celestial radiation is incoherent.

Some words on terminology: although h , $\mathbf{B}$, and $\mathbf{J}$ are all Jones operators, we only refer to $\mathbf{J}$ as the Jones matrices going forward, and we index them with a tuple of antenna location and direction $(\mathbf{x}, \mathbf{k})$. This corresponds to an optical pathway. When a Jones matrix is a scalar times identity, the Jones matrix commutes as a scalar, and we simply use scalar notation. This is the case, for example, when the Jones matrices are polarisation independent. It is also common terminology to call the sky-brightness distribution, $\mathbf{I}(\mathbf{k}) \triangleq \langle \mathbf{E}(\mathbf{k}) \otimes \mathbf{E}^\dagger(\mathbf{k}) \rangle$, the image or sky model, and values of the sky-brightness distribution are components.

In radio astronomy we measure the visibilities and invert the RIME iteratively for the sky model and Jones matrices. The process is broken into two steps called calibration and imaging. In the calibration step the current estimate of the sky model is held constant and the Jones matrices are inferred. In the imaging step the current estimate of the Jones matrices are held constant and the sky model is inferred. In both the imaging and calibration steps the algorithm performing the inference can be either direction independent (DI) or DD, which stems from the type of approximation made on the RIME when the visibilities are modelled. Namely, the DI assumption states that the Jones matrices are the same for all directions, and the DD assumption states that the Jones matrices have some form of directional dependence.

In the current work we focus entirely on the Stokes-I, that is, polarisation independent, calibration and imaging program used by the LoTSS pipeline. The LoTSS calibration and imaging pipeline is broken up into a DI calibration and imaging step followed by a DD calibration and imaging step.

Equation 3.1 shows that each optical pathway $(\mathbf{x}, \mathbf{k})$ has its own unique Jones matrix $\mathbf{J}(\mathbf{x}, \mathbf{k})$ describing the propagation of radiation. In total, this results in $2N_{\text{dir}}N_{\text{ant}}$ degrees of freedom for $\frac{1}{2}N_{\text{ant}}(N_{\text{ant}} - 1)$ observables. In practice, during DD calibration and imaging, this is never assumed for three main reasons. Firstly, the computational memory requirements of giving every direction a Jones matrix are prohibitive. Secondly, because $N_{\text{dir}} \gg N_{\text{ant}}$ it would allow too many degrees of freedom per observable, which leads to ill-conditioning. Thirdly, most sky-model components are too faint to infer a Jones matrix on the timescales that the ionosphere can be considered fixed.

The LoTSS calibration and imaging program manages to handle the computational memory requirement and ill-conditioning through a clever sparsification of the optimisation Jacobian and by invoking an extended Kalman filter, respectively; see Tasse [2014b,a], Smirnov and Tasse [2015] for details. However, there is no way to prevent the third issue of a too low signal on required timescales, and they are limited to tessellating the field of view into approximately $N_{\text{cal}} \approx 45$ facets, defined by a set of calibrator directions, $\mathcal{K}_{\text{cal}} \subset \mathcal{K}$. The result of DI and DD calibration and imaging for LoTSS is therefore a set of N_{ant} DI Jones matrices, and a set of $N_{\text{cal}}N_{\text{ant}}$ DD Jones matrices.

We consider the functional form of the final Jones matrices following this two-part calibration and imaging program. For each $(\mathbf{x}, \mathbf{k}) \in \mathcal{X} \times \mathcal{K}_{\text{cal}}$ we have the associated calibration Jones matrix,

$$\mathbf{J}^{\text{cal}}(\mathbf{x}, \mathbf{k}) \triangleq \mathbf{J}^{\text{DD}}(\mathbf{x}, \mathbf{k}) \mathbf{J}^{\text{DI}}(\mathbf{x}, \mathbf{k}') \quad (3.2)$$

$$= \mathbf{J}^{\text{DD}}(\mathbf{x}, \mathbf{k}) \mathbf{J}^{\text{DI}}(\mathbf{x}, \mathbf{k}') e^{i(\Delta_0 \phi^{\text{DD}}(\mathbf{x}, \mathbf{k}) + \Delta_0 \phi^{\text{DI}}(\mathbf{x}, \mathbf{k}'))}, \quad (3.3)$$

where notation $t \triangleq m$ means definition by equality, and can be read as ‘ t is equal by definition to m ’. An effective direction can intuitively be defined for the DI Jones matrices as the direction, $\mathbf{k}'$, that minimises the instantaneous image domain dispersive phase error effects in the DI image. Clearly, this leads to problems for a wide-field instrument, where the actual Jones matrices can vary considerably over the field of view. It is fairly simple to contrive examples where there is no well-defined effective direction. Nonetheless, we use this terminology and return to address the associated problems shortly.

From the Wiener-Khinchin theorem it follows that the visibilities are independent of the electric field phase, therefore a choice is made to spatially phase reference the Jones matrices to the location $\mathbf{x}_0$ of a reference antenna. Doing so defines the differential phase,

$$\Delta_0 \phi(\mathbf{x}, \mathbf{k}) \triangleq \phi(\mathbf{x}, \mathbf{k}) - \phi(\mathbf{x}_0, \mathbf{k}). \quad (3.4)$$

Now, we consider a set of observed visibilities, which are perfectly described by Eq. 3.1, where we neglect time, frequency, or baseline averaging. Then, there must be a set of true Jones matrices, $\mathcal{J}^{\text{true}} = \{\mathbf{J}^{\text{true}}(\mathbf{x}, \mathbf{k}) \mid (\mathbf{x}, \mathbf{k}) \in \mathcal{X} \times \mathcal{K}\}$, which give rise to the observed visibilities. We assume that this set of Jones matrices is unique. Because the Jones matrices are assumed to be scalars, we can write them as $\mathbf{J}^{\text{true}}(\mathbf{x}, \mathbf{k}) \triangleq J^{\text{true}}(\mathbf{x}, \mathbf{k}) e^{i\Delta_0 \phi^{\text{true}}(\mathbf{x}, \mathbf{k})}$.

Next, we consider the process of inferring the true Jones scalars from the observed visibilities using piece-wise constant calibration Jones scalars, $\mathbf{J}^{\text{cal}}(\mathbf{x}, \mathbf{k})$, as is done in killMS. To do so, we define the operation $[\mathbf{k}] = \text{argmax}_{\mathbf{k}'} \mathbf{k} \cdot \mathbf{k}'$ as the closest calibration direction to a given direction. In general, it is always possible to write the true Jones scalars, $\mathbf{J}^{\text{true}}(\mathbf{x}, \mathbf{k})$, in terms of the resulting calibration Jones scalars $\mathbf{J}^{\text{cal}}(\mathbf{x}, [\mathbf{k}])$ as

$$\mathbf{J}^{\text{true}}(\mathbf{x}, \mathbf{k}) \triangleq \alpha(\mathbf{x}, \mathbf{k}) e^{i\Delta_0 \beta(\mathbf{x}, \mathbf{k})} \mathbf{J}^{\text{cal}}(\mathbf{x}, [\mathbf{k}]), \quad (3.5)$$

where $\alpha(\mathbf{x}, \mathbf{k})$ and $\Delta_0 \beta(\mathbf{x}, \mathbf{k})$ correspond to the correction factors that would need to be applied to each calibration Jones scalar to make it equal to the true Jones scalar along that optical pathway. We assume the situation where there is enough information in the observed visibilities to strongly constrain the calibration Jones scalars, and that we are able to point-wise solve for a global optimum. This is unique by our above assumption. Then, if the distance between calibrators is much less than the isoplanatic-patch size [Fellgett and Linfoot, 1955], we have that $\alpha \approx 1$ and $\Delta_0 \beta \approx 0$.

The substance of this paper is aimed at inferring the function $\Delta_0 \beta$ for all optical pathways (we pay minor attention to α as well). Because of the dynamic nature of the ionosphere, the calibration facets are not always isoplanatic ($\mathbf{k} \not\approx [\mathbf{k}]$ in the appropriate sense defined in Fellgett and Linfoot [1955]), and consequently, $\Delta_0 \beta \not\approx 0$ sometimes.

We now focus our attention on the calibration Jones matrices in the direction of the calibrators and consider the value of $\Delta_0 \beta(\mathbf{x}, \mathbf{k})$ for $(\mathbf{x}, \mathbf{k}) \in \mathcal{X} \times \mathcal{K}_{\text{cal}}$. In the event that $\Delta_0 \beta(\mathbf{x}, \mathbf{k}) \approx 0$, then from Eqs. 3.5 and 3.3 it follows that the DD differential phases are

$$\Delta_0 \phi^{\text{DD}}(\mathbf{x}, \mathbf{k}) = \Delta_0 \phi^{\text{true}}(\mathbf{x}, \mathbf{k}) - \Delta_0 \phi^{\text{DI}}(\mathbf{x}, \mathbf{k}'). \quad (3.6)$$

This states that in the directions of the calibrators, the DD calibration Jones phases are equal to the difference between the true Jones phases and DI Jones phases. However, in a typical application of facet-based calibration, we can easily see that this statement is usually false. It basically means that the designation of DI and DD for the Jones matrices inferred in calibration and imaging is a misnomer.

Firstly, specifying that the DI Jones matrices have a single effective direction is an invalid statement. The effective direction, as defined above, minimises the instantaneous image domain dispersive phase error effects, and this depends largely on the ionosphere, geometry of the optical pathways, and the sky-brightness distribution. The ionosphere changes in time, acting as a time-varying scattering layer in the optical system, therefore the effective direction is also dynamic. The effective direction is heavily weighted towards the brightest sources in the field of view, and because longer baselines ‘see’ less flux, the effective direction is different for each antenna. To a lesser extent, the sensitivity drop-off at lower frequencies also introduces an effective direction that changes with sub-band frequency. Secondly, although DI calibration should remove all DI components such as station clocks, in practice, there are always remnant DI components in the DD Jones matrices.

It is desirable to assume that Eq. 3.6 holds in going forward because we interpret this to mean that we can make $\Delta_0\beta(\mathbf{x}, \mathbf{k}) \approx 0$ everywhere through an appropriate choice of $\Delta_0\phi^{\text{DD}}(\mathbf{x}, \mathbf{k})$. Therefore, we consider how the DD calibration can be modified such that Eq. 3.6 holds.

We address the first issue above, of ill-defined effective direction, by suitably forcing the observed visibilities to be well-approximated by isolated calibrators. This is applied and discussed further in Section 3.3.1.

To correct the second issue, to exactly separate all DI components from the DD Jones matrices, we use a trick. In Eq. 3.1 we can factor all commutative DI factors out in front of the summation. As we show in Appendix C, this is equivalent to directionally phase-referencing the Jones scalars, which we denote as doubly differential phase,

$$\Delta_0^2\phi \triangleq \Delta_0\phi(\mathbf{x}, \mathbf{k}) - \Delta_0\phi(\mathbf{x}, \mathbf{k}_0). \quad (3.7)$$

Moreover, we show that doubly differential phase is guaranteed to have no DI components. This trick is used and discussed further in Section 3.3.3.

After applying the above two fixes, Eq. 3.6 then takes the form of a doubly differential phase, $\Delta_0^2\phi^{\text{DD}}(\mathbf{x}, \mathbf{k})$. These not only have no DI components, but also have a well-defined reference direction. Both attributes are vital for our modelling method.

Finally, this paper is concerned with performing DD calibration by inferring $\Delta_0^2\phi^{\text{DD}}(\mathbf{x}, \mathbf{k})$ for a set of isolated calibrators followed by interpolation of these values to a multitude of optical pathways, which we call a screen. We assume that the DD Jones scalars are dominated by ionospheric dispersive phase effects and slowly changing beam errors. We neglect the ionospheric amplitude effects in this description because the phase effects are far more dominant. When the frequency of radiation is far above the ionospheric plasma frequency, then the ionospheric phase effects are given by a dispersive phase. The dispersive phase retards the wavefront without distorting it. The doubly differential ionospheric dispersive phase is given by

$$\Delta_0^2\phi^{\text{ion}}(\mathbf{x}, \mathbf{k}) = \frac{\kappa}{\nu} \Delta_0^2\tau(\mathbf{x}, \mathbf{k}), \quad (3.8)$$

where $\Delta_0^2\tau(\mathbf{x}, \mathbf{k})$ is doubly differential total electron content (DDTEC), and $\kappa = \frac{-q^2}{4\pi\epsilon_0 mc}$.

3.3 Methods

In this paper we push beyond the current state-of-the-art LoTSS DD calibration and imaging using a physics-informed probabilistic model to infer DDTEC screens. We perform our screen-based DD calibration and imaging on a randomly selected eight-hour observation from the LoTSS-DR2 archive. The selected observation, LoTSS-DR2 pointing P126+65, was observed between 23:12:14 on 8 December, 2016, and 07:28:34 on 9 December, 2016, during solar cycle 24¹. The ionosphere activity during this observation was mixed, with the first half of the observation ‘wild’ (DDTEC exceeding 100mTECU) and the last half ‘calm’.

We first provide an overview of the our DD calibration and imaging procedure, before providing a deeper explanation of each step. A summary of our DD calibration and imaging program can be seen in **Figure 3.1** and is as follows:

1. *Subtract and solve step.* Subtract a good model of the sky, except approximately 35 bright calibrators (peak flux $> 0.3 \text{ Jy beam}^{-1}$). Solve against isolated calibrators.
2. *Smooth and slow-resolve.* Smooth the Jones scalars, and resolve on a long timescale to simultaneously reduce the degrees of freedom and solve the hole problem [for a description of the problem, see Shimwell et al., 2019].
3. *Measure DDTEC.* Infer DDTEC from the Jones scalars in the directions of the bright calibrators.
4. *Infer DDTEC screen.* Apply our DDTEC model to infer DDTEC for all intermediate brightness optical pathways ($> 0.01 \text{ Jy beam}^{-1}$), forming a screen over the field of view.
5. *Image screen.* Concatenate the smoothed solutions with the DDTEC screen. Image the original visibilities with this set of solutions.

3.3.1 Subtract and solve

In this step we isolate a set of calibrators by subtracting the complementary sky model and then perform a solve. The process is visualised in the correspondingly labelled process box in **Figure 3.4**.

We begin by selecting a set of bright calibrators (peak flux $> 0.3 \text{ Jy beam}^{-1}$). The goal is to cover the field of view as uniformly as possible while not selecting too many calibrators, which can result in ill-conditioning of the system of equations that must be solved. In practice, we usually do not have enough bright calibrators to worry about ill-conditioning. Our selection criteria resulted in 36 calibrators whose layout is shown in **Figure 3.2** alongside the 45 calibrators selected by the LoTSS-DR2 pipeline.

As stated in Section 3.2, for our method to work, it is vital that the directions of the calibrators are well defined. The Jones matrix for each calibrator can be considered as a DI Jones matrix within the calibrator facet, therefore the notion of effective direction, as the direction that minimises instantaneous image domain artefacts, applies within a facet as well. Intuitively, the effective direction is located near the brightest sources in a facet. For example,

¹On 8 December, 2016, the total number of sunspots was 14 ± 1.2 and on 9 December, 2016, it was 20 ± 5.7 . The average daily sunspot count in December 2016 was 18.

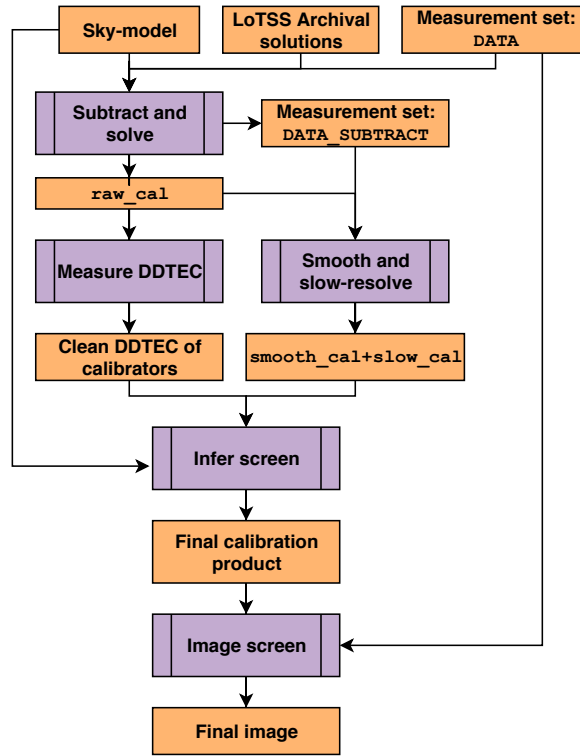


Figure 3.1: Flow diagram overview of our DD calibration and imaging program.

in **Figure 3.3** the central source is by far brighter than all nearby sources, consequently, the dispersive errors in that direction are minimised. This defines an average effective direction.

Because the effective direction changes over time, baseline, and frequency, it follows that DD modelling is systematically biased by the unknown direction of the Jones matrices if they are based on these ill-defined effective directions. We circumvent this issue by first subtracting all sources from the visibilities, except for the bright calibrators. This requires having a good enough initial sky model. In our case we use the sky model from the LoTSS-DR2 archive. We mask a $120''$ disc around each calibrator, ensuring that all artefacts for the calibrators are included in the mask. We predict the visibilities associated with the remainder of the sky model, pre-applying the solutions from the LoTSS-DR2 archive. We then subtract these predicted visibilities from the raw visibilities in the DATA column of the measurement set, placing the result in a DATA_SUBTRACT column of the measurement set. After they are subtracted, the isolated calibrator sources provide well-defined directions, which is crucial for Eq. 3.6 to hold.

We note that any unmodelled flux missing from the sky model will still reside in the visibilities, leading to a systematic calibration bias when the RIME is inverted. However, they will have been missing from the model due to their inherent faintness, therefore they should have minimal effect on the effective directions inside the facets of the isolated calibrators.

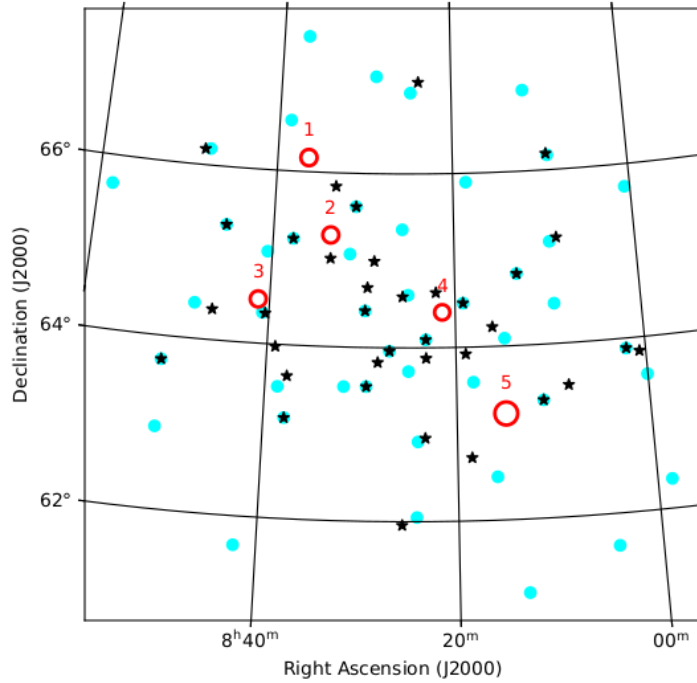


Figure 3.2: Layout of our selected calibrators (black stars) compared to the LoTSS-DR2 archival calibrator layout (cyan circles). The calibrators define a facet tessellation of the field of view. The red circles indicate the regions of the cutouts in **Figure 3.10**.

It is possible that a bright extended source be included in the set of calibrators. In this case, the model of the extended source is truncated at a radius of $120''$ from the peak. Because the sky model is necessarily incomplete, residual flux of the extended source is left over. As above, because these residual components are faint, they have negligible effects on the resulting effective direction of the calibrator. Another problem of choosing extended sources as calibrators is that of convergence due to their resolved nature. Extended calibrators can lead to divergent calibration Jones scalars, which is a problem that must also be handled by any calibration and imaging program. In our case, the sky model comes from the LoTSS-DR2 archive, which provides a carefully self-calibrated model at the same frequency and resolution, and we expect to suffer less strongly from divergent solutions.

We perform a solve on `DATA_SUBTRACT` against the sky-model components of the isolated bright calibrators. We call these solutions `raw_cal` for future reference.

Several non-ionospheric systematics will be in `raw_cal`. Firstly, beam errors are a systematic that come from performing calibration with an incomplete beam model, as well as with an aperture array that contains dead or disturbed array tiles. The beam model can be improved by better physical modelling of the aperture array, which requires a complex solutions to Maxwell's equations. The second issue is a transient effect because tiles can go dead over time and the dielectric properties of the antennae environment can change with temperature and air moisture content. It is important in the following to account for beam

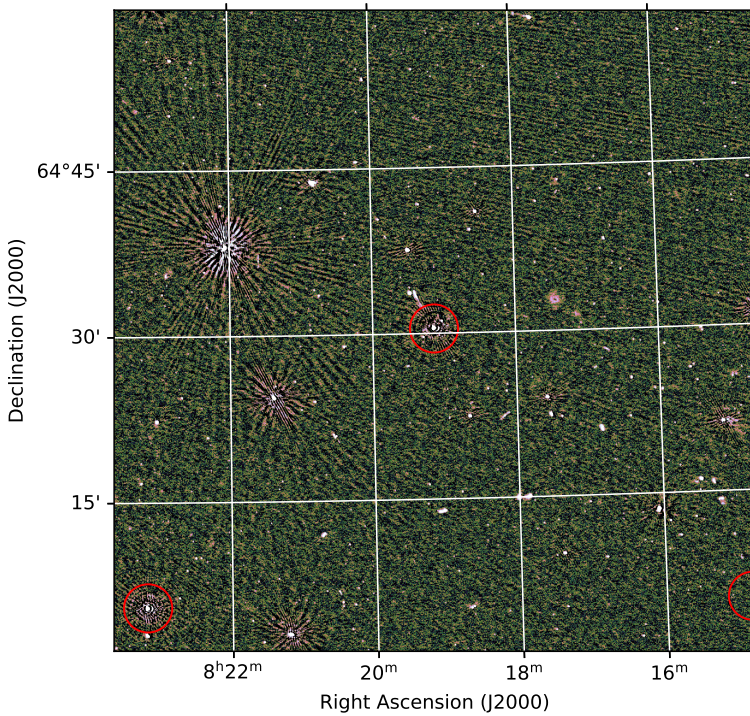


Figure 3.3: Example of an average effective direction in a facet. The red circles indicate calibrators. The central calibrator has a peak flux of $0.91 \text{ Jy beam}^{-1}$, while the next brightest non-calibrator source, approximately $16'$ to the left, is $0.12 \text{ Jy beam}^{-1}$. Within the facet of the calibrator, the Jones matrices are mostly optimised in a direction of the calibrator, but this effective direction changes over time, baseline, and frequency.

errors. Beam errors a priori are expected to change slowly over the course of an observation.

Another systematic that will find its way into the `raw_cal` are unmodelled sources outside the image, in LOFAR's relatively high-sensitivity side-lobes. The effect of such unmodelled sources on calibration is still being understood. As noted in Shimwell et al. [2019], such sources can be absorbed by additional degrees of freedom during calibration. Because they are not within the imaged region, this flux absorption can go unnoticed. It is therefore important in the following to limit the degrees of freedom of the calibration Jones scalars.

3.3.2 Smooth and slow-resolve

There are enough degrees of freedom in a DD solve for unmodelled flux to be absorbed [Shimwell et al., 2019], therefore naively applying `raw_cal` and imaging would be detrimental to source completeness. Shimwell et al. [2019] found that a post-processing step of Jones scalar phase and amplitude smoothing removed enough degrees of freedom to alleviate

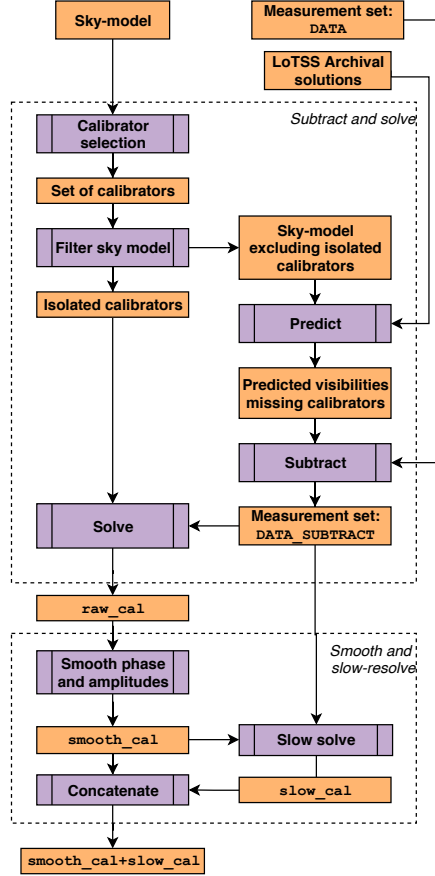


Figure 3.4: Flow chart of the *subtract and solve* and *smooth and slow-resolve* steps. Together, they represent a DD calibration program similar to the LoTSS-DR1 and DR2 pipelines.

this problem. However, they then discovered that such a smoothing operation introduces negative halos around sources, the systematic origin of which has yet to be determined. A second solve on a much longer timescale alleviates the negative halo problem. The process is visualised in the correspondingly labelled process box in **Figure 3.4**.

The LoTSS DD calibration and imaging pipeline performs Jones scalar phase smoothing by fitting a TEC-like and constant-in-frequency functional form to the Jones phases using maximum likelihood estimation. This results in two parameters per 24 observables (the number of frequency sub bands). We alter the smoothing method by fitting a three parameter model: a TEC-like term, a constant-in-frequency, and a clock-like functional form. Our choice results in three parameters per 24 observables. We add the additional term because a small ($\lesssim 1$ ns) remnant clock-like term appears to exist in the DD Jones phases.

Weak scattering in the ionosphere causes a well-known DD amplitude effect, however the amplitudes of the Jones scalars in `raw_cal` are more complicated, with possible contamination from an improperly modelled beam and unmodelled flux. Therefore, we do not

consider modelling amplitude in this work, but we do impose the prior belief that beam error amplitudes change slowly over time and frequency as in Shimwell et al. [2019]. To impose this prior, we simply smooth the log-amplitudes of `raw_cal` with a two-dimensional median filter in time and frequency with window size (15 min., 3.91 MHz).

Scintillation can also significantly effect the amplitudes of the Jones scalars. When the scintillation timescale is shorter than the integration timescale, the observed visibilities can decohere. In this case, the visibilities should be flagged. DDFacet contains an adaptive visibility weighting mechanism inspired by the optical method of "lucky imaging" that down-weights noisy visibilities [Bonnassieux et al., 2018], which effectively handles fast scintillation. When scintillation occurs on a longer timescale, the effect is characterised by amplitude jumps for groups of nearby antennae within the Fresnel zone. These effects are predominately removed during the DI calibration and are not considered further. This implies that we do not consider DD scintillation.

A surprising result of the smoothing procedure is that it introduces negative halos around bright sources, which destroys the integrity of the image. In the LoTSS pipeline, despite much effort to understand the fundamental cause of the problem, a temporary solution is to pre-apply the smoothed solutions and resolve against the same sky model on a long timescale. We perform this resolve with a solution interval of 45 minutes and call these solutions `slow_cal`. We form the concatenation of `smooth_cal` and `slow_cal` and call it `smooth_cal+slow_cal`.

3.3.3 Measure DDTEC

Next, we extract precise DDTEC measurements from `raw_cal`, which are later used to infer a DDTEC screen. While it would be possible to model a DDTEC screen directly on Jones scalars, the required inference algorithm would be extremely expensive to account for the high-dimensional, complicated, multi-modal posterior distribution that occurs due to phase wrapping. Therefore, in this section we describe how we use the ν^{-1} dependence of dispersive phase (Eq. 3.8) to measure DDTEC². The process is visualised in the correspondingly labelled box of **Figure 3.8**.

First, we directionally reference the Jones phases in `raw_cal` by subtracting the phase of the brightest calibrator, that is, the reference calibrator, from the phases of all other directions. We can assume that all remnant DI effects are removed from the resulting doubly differential phases (see Appendix C). We then assume that these Jones scalar phases are completely described by DDTEC and the smoothed amplitudes coming from beam errors.

The idea of fitting frequency ν^{-1} dependence to phase to extract a TEC-like term is not new [e.g. van Weeren et al., 2016]. However, fitting ionospheric components from Jones scalars is a notoriously difficult procedure because the Jones scalars are typically noise dominated, with heteroscedastic correlated noise profiles, and there are many local minima caused by phase-wrapping. **Figure 3.5** shows an example of simulated Jones scalars data for a single optical pathway with a typical observational noise profile corresponding to a remote antenna and calibrator peak flux of 0.3 Jy beam^{-1} . Typically, the problem is solved using maximum-likelihood estimators and the resulting TEC-like parameters are too biased or of high variance

²We use the term 'measure DDTEC' purely for distinction between the inferred DDTEC screen discussed in Section 3.3.4. Also, from a philosophical perspective, a measurement is a single realisation of inference.

to be used for any sort of inference. For our purposes we require uncertainties of the measured DDTEC, which should be of the order of a few mTECU.

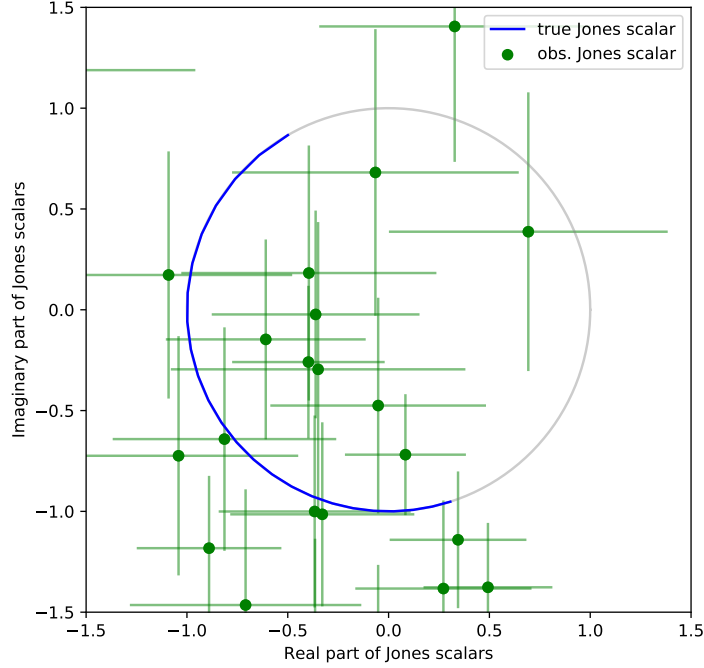


Figure 3.5: Simulated real and imaginary components of noise-dominated Jones scalars and observational uncertainties. The simulated ground truth DDTEC is 150 mTECU and the observational uncertainties are frequency dependent and correlated, corresponding to a remote station for a calibrator at the cutoff peak flux limit of 0.3 Jy beam^{-1} .

We formulate the problem as a hidden Markov model [HHM; Rabiner and Juang, 1986]. In this HMM the Jones scalars of an optical pathway form a sequence of observables indexed by time, and the DDTEC of the optical pathway forms a Markov chain of hidden variables. We then apply recursive Bayesian estimation to obtain the posterior distribution of the DDTEC at each point in time given the entire Jones scalar sequence. As shown in Appendix D, this can be done in two passes over the data, using the forward and backward recursions, Equations D.3 and D.5, respectively.

Let $\mathbf{g}_i \in \mathbb{C}^{N_{\text{freq}}}$ be the observed complex Jones scalar vector at time step i , and let $\Delta_0^2 \tau_i$ be the corresponding DDTEC. For visual clarity, we drop the optical pathway designation because this analysis happens per optical pathway.

We assume that the Jones scalars have Gaussian noise, described by the observational covariance matrix Σ . Thus, we define the likelihood of the Jones scalars as

$$p(\mathbf{g}_i | \Delta_0^2 \tau_i, \Sigma) = \mathcal{N}_{\mathbb{C}}[\mathbf{g}_i | |\tilde{\mathbf{g}}_i| e^{i \frac{\kappa}{\nu} \Delta_0^2 \tau_i}, \Sigma], \quad (3.9)$$

where $|\tilde{\mathbf{g}}_i|$ denotes the smoothed amplitudes, and $\mathcal{N}_{\mathbb{C}}$ denotes the complex Gaussian distri-

bution. The complex Gaussian distribution of a complex random variable is formulated as a multivariate Gaussian of the stacked real and imaginary parts.

Because we expect spatial and temporal continuity of FED, the DDTEC should also exhibit this continuity. Therefore, we assume that in equal time intervals the DDTEC is a Lévy process with Gaussian steps,

$$p(\Delta_0^2 \tau_{i+1} | \Delta_0^2 \tau_i) = \mathcal{N}[\Delta_0^2 \tau_{i+1} | \Delta_0^2 \tau_i, \omega^2 \delta t], \quad (3.10)$$

where ω^2 is the variance of the step size. This is equivalent to saying that the DDTEC prior is a Gaussian random walk.

Equations 3.9 and 3.10 define the HMM data likelihood and state transition distribution necessary to compute Equations D.3 and D.5. The relation between DDTEC and amplitude and the Jones scalars is non-linear, therefore Equation D.3 is analytically intractable. To evaluate the forward equation, we apply variational inference, which is an approximate Bayesian method. Variational inference proceeds by approximating a distribution with a tractable so-called variational distribution with variational parameters that must be learned from data [e.g. Hensman et al., 2013]. A lower bound on the Bayesian evidence is maximised for a point-wise estimate of the variational parameters, resulting in a variational distribution that closely resembles the actual distribution.

We assume that the variational forward distribution for DDTEC, at time step i given all past Jones scalars, $\mathbf{g}_{0:i}$, is closely approximated by a Gaussian,

$$q(\Delta_0^2 \tau_i) = \mathcal{N}[\Delta_0^2 \tau_i | m_i, \gamma_i^2], \quad (3.11)$$

where the mean and variance is m_i and γ_i^2 , respectively. When this choice is made, the state transition distribution is conjugate, and the forward recursion reduces to the well-known Kalman-filter equations, and the backward recursion reduces to the Rauch-smoother equations [Rauch, 1963].

We consider marginalisation of the state transition distribution in Equation D.2. Marginalising with respect to Equation 3.11 at time step $i-1$, we obtain

$$p(\Delta_0^2 \tau_i | \mathbf{g}_{0:i-1}, \Sigma) = \mathcal{N}[\Delta_0^2 \tau_i | m_{i-1}, \omega^2 \delta t + \gamma_{i-1}^2]. \quad (3.12)$$

This can be viewed as the propagated prior belief in DDTEC. The forward probability density is then given by Bayes equation,

$$p(\Delta_0^2 \tau_i | \mathbf{g}_{0:i}, \Sigma) = \frac{p(\mathbf{g}_i | \Delta_0^2 \tau_i, \Sigma) p(\Delta_0^2 \tau_i | \mathbf{g}_{0:i-1}, \Sigma)}{p(\mathbf{g}_i | \mathbf{g}_{0:i-1}, \Sigma)}. \quad (3.13)$$

This is the posterior probability density given all previous observations in the observable sequence.

Variational inference proceeds by minimising the Kullbeck-Leibler divergence $\text{KL}[q || p] = \int_{\mathbb{R}} \log q/p \, dq$ from the variational forward distribution q to the true forward distribution p with respect to the variational parameters. Equivalently, we may maximise the following evidence lower-bound objective (ELBO),

$$\mathcal{L}_i[q, p] = \mathbb{E}_{q(\Delta_0^2 \tau_{i-1})} \left[\log \frac{p(\Delta_0^2 \tau_i | \mathbf{g}_{0:i}, \Sigma) p(\Delta_0^2 \tau_i | \mathbf{g}_{0:i-1}, \Sigma)}{q(\Delta_0^2 \tau_{i-1})} \right] \quad (3.14)$$

$$\begin{aligned} &= \mathbb{E}_{q(\Delta_0^2 \tau_{i-1})} [\log p(\mathbf{g}_i | \Delta_0^2 \tau_i, \Sigma)] \\ &\quad - \text{KL}[q(\Delta_0^2 \tau_{i-1}) || p(\Delta_0^2 \tau_i | \mathbf{g}_{0:i-1}, \Sigma)]. \end{aligned} \quad (3.15)$$

Because the variational and prior distributions are both Gaussian, there is an analytic expression for the KL term³. The expectation of the likelihood term is called the variational expectation and is analytically derived in Appendix E.

One of the benefits of variational inference is that it turns a problem that requires computationally expensive Markov chain Monte Carlo methods into an inexpensive optimisation problem. It works well so long as the variational distribution adequately describes the true distribution.

We have assumed the variational forward distribution to be Gaussian, therefore the approximated system becomes a linear dynamical system (LDS). Several great properties follow. Firstly, the backward equation, Equation D.5, is reduced to the Rauch recurrence relations [Rauch, 1963], which are easily accessible [e.g. Shumway and Stoffer, 1982], so we do not write them down here. The Rauch recurrence relations are analytically tractable. As long as the variational forward distribution is a reasonable approximation, they therefore provide a computationally inexpensive improvement to DDTEC estimation. In contrast, performing the forward filtering problem alone uses only half of the available information in the observable sequence.

Secondly, the unknown step size variance ω and the observational uncertainty Σ can be estimated iteratively using the expectation-maximisation (EM) algorithm for LDS [Shumway and Stoffer, 1982]. For a good introduction to HHMs and parameter estimation see Shumway and Stoffer [1982], Rabiner and Juang [1986], Dean et al. [2014]. In summary, each iteration of the EM algorithm starts with the E-step, which calculates the forward and backward equations to obtain an estimate of the hidden variables given the entire sequence of observables. The M-step then consists of maximising the expected log-posterior probability of the hidden parameters found in the E-step. The M-step equations for LDS are given explicitly in Shumway and Stoffer [1982], therefore we do not provide them here. We call these solutions the LDS-EM solutions.

The complete DDTEC inference algorithm is given in Algorithm 1. For the initial prior DDTEC prior mean and prior uncertainty we use 0 mTECU and 300 mTECU, respectively. The top panel in **Figure 3.6** shows the ELBO basin for the Jones scalar data in **Figure 3.5** during the first forward pass iteration of the algorithm. During the first iteration, the HMM parameters are not known and the ELBO basin has many local optima. However, after the first iteration, the HMM parameters are estimated, and during the second pass, most of the local optima are gone, corresponding to better priors that were learned from the data. This happens because the constrained Lévy process variance regularises the inferred DDTEC. In situations with noisy data, but where a priori the DDTEC changes very little, this can significantly improve the estimate of posterior DDTEC. In situations where the DDTEC varies much between time steps, the time-coupling regularisation helps less, but it still reduces the variance of the estimates. **Figure 3.7** shows the resulting 90% confidence interval of the posterior phase for this example simulated DDTEC inference problem, and the method clearly correctly found the right phase wrap.

To globally optimise ELBO in the presence of many local optima, we use brute force to locate a good starting point, and then use BFGS quasi-Newton optimisation from that point to find the global optimum. Because there are only two variational parameters, the brute-force grid search is very quick.

³ $\text{KL}[\mathcal{N}[\mu_1, \sigma_1^2] \parallel \mathcal{N}[\mu_2, \sigma_2^2]] = \frac{1}{2} \left(\frac{\sigma_1^2}{\sigma_2^2} + \frac{(\mu_2 - \mu_1)^2}{\sigma_2^2} - 1 + \log \frac{\sigma_2^2}{\sigma_1^2} \right).$

```

/* Initialise */
Set  $N_{\max}$  to a max number of iterations;
Set  $\Sigma^{(0)} = \sigma^2 \mathbf{I}$  as an initial estimate of observational covariance of  $\mathbf{g}$  per frequency;
Set  $\omega^{(0)}$  to an initial estimate of the DDTEC Lévy variance per time-interval;
Set  $n = 0$ ;
while  $n < N_{\max}$  do
    /* E-step */
    Set  $i = 0$ ;
    while  $i < T$  ; // Forward pass
    do
        Use  $\omega^{(n)}$  and  $\Sigma^{(n)}$  to define  $\mathcal{L}_i[q \parallel p]$ ;
        Maximise  $\mathcal{L}_i[q \parallel p]$  for  $m_i^{(n)}$  and  $\gamma_i^{(n)}$ ;
        Set  $i = i + 1$ ;
    end
    Set  $i = T$ ;
    while  $i > 0$  ; // Backward pass
    do
        Revise estimate of  $m_i^{(n)}$  and  $\gamma_i^{(n)}$  with Rauch-recurrence relation;
        Set  $i = i - 1$ ;
    end
    /* M-step */
    Set  $\omega^{(n+1)}$  and  $\Sigma^{(n+1)}$  to the LDS-EM solutions;
    Set  $n = n + 1$ ;
end
/* Final variational mean and variance */
return  $\{m_i^{(n)}\}, \{\gamma_i^{(n)}\}$ 

```

Algorithm 1: Solving the recursive Bayesian estimation of DDTEC from Jones scalars using variational inference to approximate the analytically intractable forward distribution. It assumes that the DDTEC is a Lévy process with Gaussian steps. It uses the EM-algorithm to estimate the the observational covariance and Lévy process step variance.

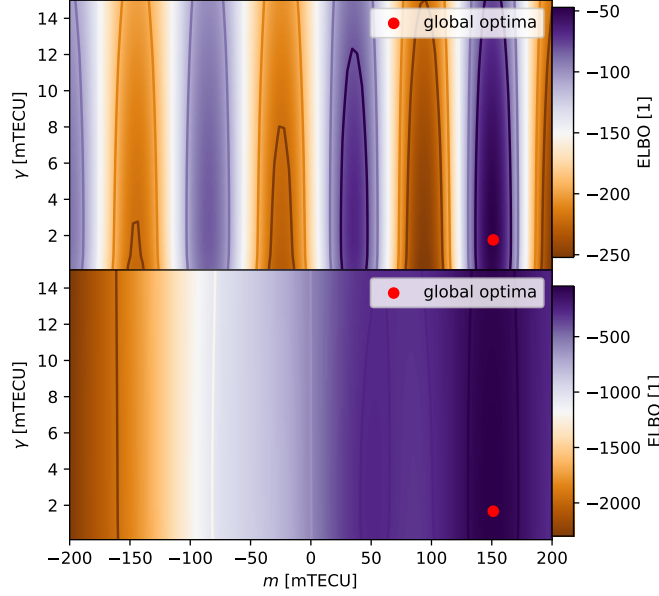


Figure 3.6: ELBO landscape of the simulated Jones scalars in **Figure 3.5** during the first and second iterations of Algorithm 1. The top panel shows the ELBO basin during the first iteration, and the lower panel shows the ELBO basin during the second iteration after one round of parameter estimation. The ground-truth DDTEC is 150 mTECU in this example.

In this randomly chosen data set 0.1% of the measured DDTEC are found to be outlier solutions, that is, approximately 1.6×10^3 of 2.6×10^6 optical pathways. Outliers are visually characterised by a posterior uncertainty that is too small to explain its deviation from its neighbouring (in time and direction) DDTEC. Outliers drastically affect the performance of the screen inference step because the uncertainties will be ‘trusted’.

We detect outliers using a two-part heuristic approach, where we value having more false positives over false negatives. First, we filter for large jumps in DDTEC. We use rolling-statistics sigma-clipping, where we flag DDTEC that deviate by more than two standard deviations from the median in a rolling temporal window of 15 min. Secondly, we filter for large jumps directionally. We fit a smoothed multiquadric radial basis function to the DDTEC over direction and flag outliers where the DDTEC deviates by more than 8 mTECU. Measured DDTEC that are flagged as outliers are given infinite uncertainty.

3.3.4 Infer DDTEC screen

We denote the DDTEC mean and uncertainties inferred with variational inference in Section 3.3.3 as the measured DDTEC with associated measurement uncertainties. In this section

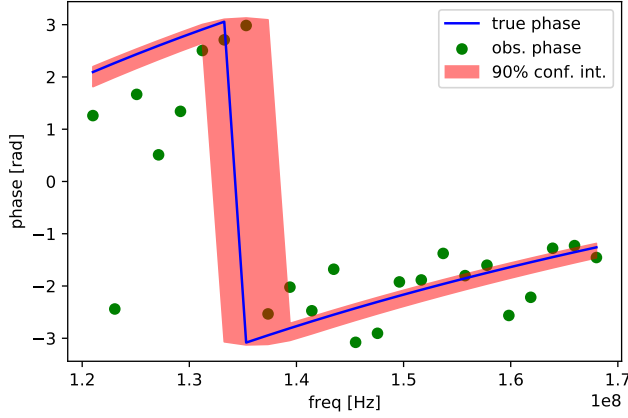


Figure 3.7: Posterior doubly differential phase solved with variational inference. The shaded region shows the central 90% confidence interval following two passes of Algorithm 1. The resulting variational estimate for DDTEC is 149.1 ± 2.3 mTECU. The ground-truth DDTEC is 150 mTECU.

we show how we perform probabilistic inference of a DDTEC screen. In order to distinguish the DDTEC screen from the measured DDTEC, we call them the inferred DDTEC. The process is visualised in the correspondingly labelled box in **Figure 3.8**.

We define the screen directions by selecting the brightest 250 directions with apparent brightness greater than $0.01 \text{ Jy beam}^{-1}$ and separated by at least $4'$. We exclude the original calibrator directions from the set of screen directions, which already have optimal solutions provided in `smooth_cal+slow_cal`.

We model the measured DDTEC over optical pathways, indexed by $(\mathbf{x}, \mathbf{k})$, following the ionospheric model proposed in Albert et al. [2020a], which geometrically models the ionosphere as a flat, thick, infinite-in-extent layer with free electron density (FED) realised from a Gaussian process. Whereas in Albert et al. [2020a] the modelled quantity was differential total electron content (DTEC), here we model DDTEC, which requires a modification to their model. Generalising the procedure outlined therein, we can arrive at similar expressions as their Eqs. 11 and 14. Because the results can be arrived at by following the same procedure, we do not show the derivation here and simply state the DDTEC prior mean and covariance,

$$m_{\Delta_0^2 \tau}(\mathbf{x}, \mathbf{k}) = \mathbb{E}[\Delta_0^2 \tau(\mathbf{x}, \mathbf{k})] = 0, \quad (3.16)$$

$$K_{\Delta_0^2 \tau}((\mathbf{x}_i, \mathbf{k}), (\mathbf{x}_j, \mathbf{k}')) = \mathbb{E}[\Delta_0^2 \tau(\mathbf{x}_i, \mathbf{k}) \Delta_0^2 \tau(\mathbf{x}_j, \mathbf{k}')] \quad (3.17)$$

$$\begin{aligned} &= I_{ij}^{\mathbf{k}\mathbf{k}'} - I_{i0}^{\mathbf{k}\mathbf{k}'} - I_{ij}^{\mathbf{k}\mathbf{k}_0} + I_{i0}^{\mathbf{k}\mathbf{k}_0} \\ &\quad - I_{0j}^{\mathbf{k}\mathbf{k}'} + I_{00}^{\mathbf{k}\mathbf{k}'} + I_{0j}^{\mathbf{k}\mathbf{k}_0} - I_{00}^{\mathbf{k}\mathbf{k}_0} \\ &\quad - I_{ij}^{\mathbf{k}_0\mathbf{k}'} + I_{i0}^{\mathbf{k}_0\mathbf{k}'} + I_{ij}^{\mathbf{k}_0\mathbf{k}_0} - I_{i0}^{\mathbf{k}_0\mathbf{k}_0} \\ &\quad + I_{0j}^{\mathbf{k}_0\mathbf{k}'} - I_{00}^{\mathbf{k}_0\mathbf{k}'} - I_{0j}^{\mathbf{k}_0\mathbf{k}_0} + I_{00}^{\mathbf{k}_0\mathbf{k}_0}, \end{aligned} \quad (3.18)$$

where we have used the same notation with Latin subscripts denoting antenna indices. Each term in Eq. 3.18 is a double integral

$$I_{ij}^{\mathbf{k}\mathbf{k}'} = \int_{s_i^{\mathbf{k}-}}^{s_i^{\mathbf{k}+}} \int_{s_j^{\mathbf{k}'-}}^{s_j^{\mathbf{k}'+}} K(\mathbf{x}_i + s\mathbf{k}, \mathbf{x}_j + s'\mathbf{k}') ds ds' \quad (3.19)$$

of the FED covariance function K , and the integration limits are given by

$$s_i^{\mathbf{k}\pm} = \left(a \pm \frac{b}{2} - (\mathbf{x}_i - \mathbf{x}_0) \cdot \hat{\mathbf{z}} \right) (\mathbf{k} \cdot \hat{\mathbf{z}})^{-1}, \quad (3.20)$$

where a is the height of the centre of the ionosphere layer above the reference antenna, and b is the thickness of the ionosphere layer. The coordinate frame is one where the $\hat{\mathbf{z}}$ is normal to the ionosphere layer, for instance, the east-north-up frame.

The FED kernel, K , that is chosen determines the characteristics of the resulting DDTEC. In Section 3.3.4 we discuss our choice of FED kernels. For a comprehensive explanation of the effect of different FED kernels on DTEC and examples of inference with this model, see Albert et al. [2020a]. Importantly, the only difference between inference of DDTEC and inference of DTEC is in the number of terms in Eq. 3.18, where there are 16 for DDTEC and only 4 for DTEC.

Using Eqs. 3.16 and 3.18 to define the DDTEC prior distribution, we infer the posterior distribution for the DDTEC screen given our measured DDTEC following standard Gaussian process regression formulae [e.g. Rasmussen and Williams, 2006a]. Specifically, given a particular choice for the FED kernel, K , if the set of optical pathways of the measured DDTEC are given by $\mathbf{Z}$ and the set of all optical pathways that we wish to infer DDTEC over are given by $\mathbf{Z}^*$, then the posterior distribution is [cf. Eq. 19 in Albert et al., 2020a]

$$P(\Delta_0^2 \tau(\mathbf{Z}^*) | \Delta_0^2 \tau(\mathbf{Z}), K) = \mathcal{N}[\Delta_0^2 \tau(\mathbf{Z}^*) | \mathbf{m}(\mathbf{Z}^*), \mathbf{K}(\mathbf{Z}^*, \mathbf{Z}^*)], \quad (3.21)$$

where the conditional mean is $\mathbf{m}(\mathbf{Z}^*) = K_{\Delta_0^2 \tau}(\mathbf{Z}^*, \mathbf{Z}) \mathbf{B}^{-1} \Delta_0^2 \tau(\mathbf{Z})$, the conditional covariance is $\mathbf{K}(\mathbf{Z}^*, \mathbf{Z}^*) = K_{\Delta_0^2 \tau}(\mathbf{Z}^*, \mathbf{Z}^*) - K_{\Delta_0^2 \tau}(\mathbf{Z}^*, \mathbf{Z}) \mathbf{B}^{-1} K_{\Delta_0^2 \tau}(\mathbf{Z}, \mathbf{Z}^*)$, and $\mathbf{B} = K_{\Delta_0^2 \tau}(\mathbf{Z}, \mathbf{Z}) + \Sigma_{\Delta_0^2 \tau}(\mathbf{Z}, \mathbf{Z})$. The measured DDTEC mean and uncertainties are $\Delta_0^2 \tau(\mathbf{Z})$ and $\Sigma_{\Delta_0^2 \tau}(\mathbf{Z}, \mathbf{Z})$, respectively. To handle infinite measurement uncertainties, we factor $\Sigma_{\Delta_0^2 \tau}(\mathbf{Z}, \mathbf{Z})$ into the product of two diagonal matrices, then symmetrically factor the diagonals out of $\mathbf{B}$. The inverted infinities then act to zero-out corresponding columns on right multiplication and rows on left multiplication.

For a given choice of FED kernel, K , the geometric parameters of the ionosphere and FED kernel hyper parameters are determined by maximising the log-marginal likelihood of the measured DDTEC [cf. Eq. 18 in Albert et al., 2020a],

$$\log P(\Delta_0^2 \tau(\mathbf{Z}) | K) = \log \mathcal{N}[\Delta_0^2 \tau(\mathbf{Z}) | \mathbf{0}, \mathbf{B}]. \quad (3.22)$$

For all our Gaussian process computations we make use of GPFlow [Matthews et al., 2017], which uses auto-differentiation of Tensorflow [Abadi et al., 2015] to expedite complex optimisation procedures on both CPUs and GPUs.

Accounting for a dynamic ionosphere

The ionosphere is very dynamic, with a variety of influencing factors. It can change its overall behaviour in a matter of minutes [e.g. Mevius et al., 2016, Jordan et al., 2017]. In our

experimentation, we found that choosing the incorrect FED kernel results in very poor quality screens because of a systematic modelling bias.

We delimit two cases of incorrectly specifying the FED kernel. The first is when the true FED covariance is stationary but the spectral properties are incorrectly specified. This can happen, for example, when the character of the ionosphere changes from rough to smooth, or vice versa. The tomographic nature of $K_{\Delta_0^2\tau}$ then incorrectly infers DDTEC structure that is not there. In the second case, the true FED covariance is non-stationary and a stationary FEED kernel is assumed. This happens, for example, when a travelling ionospheric disturbance (TID) passes over the field of view [van der Tol, 2009]. In this case, the FED has a locally non-stationary component and the tomographic nature of the DDTEC kernel will incorrectly condition on the TID structures, causing erratic predictive distributions.

Both types of bias can be handled by a dynamic marginalisation over model hypotheses. To do this, we denote each choice of FED kernel as a hypothesis. For each choice of FED kernel, we optimise the FED and ionosphere hyperparameters by maximising the log-marginal likelihood of the measured DDTEC, Eq. 3.22. We then define the hypothesis-marginalised posterior as

$$P(\Delta_0^2\tau(\mathbf{Z}^*) | \Delta_0^2\tau(\mathbf{Z})) = \sum_i w_i P(\Delta_0^2\tau(\mathbf{Z}^*) | \Delta_0^2\tau(\mathbf{Z}), K_i), \quad (3.23)$$

where $w_i = P(\Delta_0^2\tau(\mathbf{Z}) | K_i) / \sum_i P(\Delta_0^2\tau(\mathbf{Z}) | K_i)$. This marginalisation procedure can be viewed as a three-level hierarchical Bayesian model [e.g. Rasmussen and Williams, 2006a]. The hypothesis-marginalised distribution is a mixture of Gaussian distributions, which is a compound model and is no longer Gaussian in general. We approximate the mixture with a single Gaussian by analytically minimising the Kullbeck-Leibler divergence from the mixture to the single Gaussian. For reference, the mean of the single Gaussian approximation is the convex sum of means weighted by the hypothesis weights.

For our FED hypothesis kernels we select the Matérn- $p/2$ class of kernels for $p \in \{1, 3, 5, \infty\}$ and the rational-quadratic kernel [Rasmussen and Williams, 2006a]. All selected kernels are stationary. While we could incorporate non-stationary FED kernel hypotheses to handle the second type of bias, we do not explore such kernels in this work. See Albert et al. [2020a] for examples of the Matérn- $p/2$ class of FED kernels.

Computational considerations

In order for this approach to be practically feasible, the computation should only require a few hours. The main bottleneck of the computation is computing the 16 double integral terms in Eq. 3.18 for every pair of optical pathways, and inverting that matrix. A single covariance matrix requires evaluating $16(N_{\text{res}}N_{\text{ant}}N_{\text{cal}})^2$ double-precision floating point numbers that are stored in a $N_{\text{ant}}N_{\text{cal}} \times N_{\text{ant}}N_{\text{cal}}$ matrix. In the case of P126+65, using an abscissa resolution of $N_{\text{res}} = 5$ as suggested by [Albert et al., 2020a], the covariance matrix is approximately 23GB of memory, with approximately 600GB of memory used in intermediate products. The large intermediate products mean that the covariance matrix evaluation is memory bus speed bound. Furthermore, the inversion scales with $N_{\text{ant}}^3 N_{\text{cal}}^3$. Hyperparameter optimisation requires approximately 200 iterations of hessian-based gradient descent, with each iteration computing the matrix and inverse at least once. Therefore, a lower limit on the number of matrix inversions required for a typical observation with five FED kernel hypotheses is 10^6 .

On a single computer with 32 cores and 512 GB of RAM this takes several weeks. Chunking data in time produces modest savings. Using GPUs would make this calculation feasible in just a few hours. However, while GPUs are intrinsically enabled by Tensorflow, we do not currently have access to GPUs with enough memory.

To perform this scalably with enough FED kernel hypotheses to handle the dynamic nature of the ionosphere would require too much computing power for a single measurement set calibration, and unfortunately, a trade-off must be made. Therefore, we first make the approximation that the ionosphere is thin, that is, we let $b \rightarrow 0$. When we make this approximation, the double integral terms become

$$\tilde{I}_{ij}^{\mathbf{k}\mathbf{k}'} = K(\mathbf{x}_i + s_i^{\mathbf{k}0} \mathbf{k}, \mathbf{x}_j + s_j^{\mathbf{k}'0} \mathbf{k}'), \quad (3.24)$$

where $s_i^{\mathbf{k}0} = (a - (\mathbf{x}_i - \mathbf{x}_0) \cdot \hat{\mathbf{z}})(\mathbf{k} \cdot \hat{\mathbf{z}})^{-1}$. This reduces the memory size of intermediate products by N_{res}^2 approximately from 600GB to 30GB, but the size of the matrix is still the same, and the same problem of inverting a large matrix appears many times.

Therefore we introduce a second approximation, that there is no a priori coupling between antennae. When we do this, Eq. 3.18 becomes

$$\begin{aligned} K_{\Delta_{0\tau}^2}^{\mathbf{x}_i}(\mathbf{k}, \mathbf{k}') = & \tilde{I}_{ii}^{\mathbf{k}\mathbf{k}'} - \tilde{I}_{i0}^{\mathbf{k}\mathbf{k}'} - \tilde{I}_{ii}^{\mathbf{k}\mathbf{k}_0} + \tilde{I}_{i0}^{\mathbf{k}\mathbf{k}_0} \\ & - \tilde{I}_{0i}^{\mathbf{k}\mathbf{k}'} + \tilde{I}_{00}^{\mathbf{k}\mathbf{k}'} + \tilde{I}_{0i}^{\mathbf{k}\mathbf{k}_0} - \tilde{I}_{00}^{\mathbf{k}\mathbf{k}_0} \\ & - \tilde{I}_{ii}^{\mathbf{k}_0\mathbf{k}'} + \tilde{I}_{i0}^{\mathbf{k}_0\mathbf{k}'} + \tilde{I}_{ii}^{\mathbf{k}_0\mathbf{k}_0} - \tilde{I}_{i0}^{\mathbf{k}_0\mathbf{k}_0} \\ & + \tilde{I}_{0i}^{\mathbf{k}_0\mathbf{k}'} - \tilde{I}_{00}^{\mathbf{k}_0\mathbf{k}'} - \tilde{I}_{0i}^{\mathbf{k}_0\mathbf{k}_0} + \tilde{I}_{00}^{\mathbf{k}_0\mathbf{k}_0}. \end{aligned} \quad (3.25)$$

This covariance matrix requires only computing $16(N_{\text{cal}})^2$ double-precision floating point numbers and requires inverting an $N_{\text{cal}} \times N_{\text{cal}}$ matrix. This is extremely fast, even with CPUs, and the whole P126+65 measurement set can be performed in 50 minutes with five FED kernel hypotheses.

After making these approximations with a stationary FED kernel, it becomes difficult to infer the height of the ionosphere, a . Rather, if the FED kernel is isotropic and parametrised by a length scale, for instance, l , then the marginal likelihood becomes sensitive to the ratio a/l . Therefore it is still possible to learn something of the ionosphere with this approximation.

3.3.5 Image screen

DDFacet has internal capability for applying solutions during imaging in arbitrary directions [Tasse et al., 2018]. DDFacet typically works with its own proprietary solution storage format. To facilitate working with our solutions, we extended DDFacet to work with and receive the directions layout from H5Parm files whose specifications are defined by LoSoTo [de Gasperin et al., 2019]. We convert the `smooth_cal+slow_cal` solutions to H5Parm format and then convert DDTEC to DTEC by adding back the reference direction phase from `smooth_cal`.

We use the DDFacet hybrid matching pursuit clean algorithm, with five major cycles, 10^6 minor cycles, a Briggs robust weight of -0.5 [Briggs, 1995], a peak gain factor of 0.001, and clean threshold of $100 \mu\text{Jybeam}^{-1}$. We initialise the clean mask with the final mask from LoTSS-DR2. All other settings are the same as in the LoTSS-DR2 pipeline and can be found in Shimwell et al. [2019]. We also apply the same settings to image the solutions in `smooth_cal+slow_cal` for comparison purposes later.

3.4 Results

Figure 3.9 shows several dimensional slices of the resulting DDTEC screens for LOFAR remote antenna RS508HBA, which is 37 km from the reference antenna. Remote antennae are typically more difficult to calibrate because flux on smaller angular scales (longer baselines) drops off, therefore they are a good choice for showing the performance of the DDTEC measurement and screen inference. The top panel shows the temporal evolution of a single direction. We observe two distinct ionospheric behaviours over the course of the observation. From 0 to 3.5 hours the ionosphere is ‘wild’, with DDTEC variations greater than 150 mTECU, and the variation between time steps are more noise-like. From 3.5 hours until the end of the observation, the DDTEC rarely exceeds 10 mTECU, and the variations between time steps are smoother. For comparison, a central antenna within 500 m of the reference antenna typically has DDTEC variations of 10 mTECU. In order to have such a small DDTEC so far from the reference antenna, the FED of the ionosphere should be highly spatially correlated on length scales longer than that baseline. This implies that a structure larger than 50 km, potentially at low altitude, is passing over the array during that time.

Because we applied FED model hypothesis marginalisation, we can confirm that the spatial power spectrum indeed changed during these two intervals. During the first half of the observation, coinciding with the ‘wild’ ionosphere, the most highly weighted FED kernel hypothesis was the Matérn-1/2 kernel, followed by the Matérn-3/2 kernel. During the last half of the observation, coinciding with the ‘calm’ ionosphere, the most highly weighted FED kernel hypothesis was the Matérn-3/2 kernel, followed by the Matérn-5/2 kernel.

The change in temporal variation roughness implies that the temporal power spectrum of the FED changed in shape, becoming more centrally concentrated in the last half of the observation. If we apply the frozen-flow assumption, then, over short enough time intervals, the temporal covariance becomes the FED spatial covariance and therefore we expect to see rougher temporal correlations when the spatial correlations become rougher. The frozen-flow assumption is thus supported by the fact that the first and last half of the observations are better described by rough and smooth FED kernels, respectively, matching the temporal behaviour.

The middle and bottom rows of **Figure 3.9** show the antenna-based DDTEC-screens (showing what the ionosphere looks like from the perspective the antenna) of our inferred DDTEC and measured DDTEC, respectively. The left columns are a slice during the ‘wild’ interval and the right column is a slice during the ‘calm’ interval (points indicated in the temporal profile). We note that during the ‘calm’ time the DDTEC-screens are mostly negative. From Equation 3.16 we expect that the DDTEC should be zero on average, which seems at odds with the overall negative DDTEC screen at this time slice. As pointed out in Albert et al. [2020a], the mean DTEC, and therefore mean DDTEC, should be zero only in the case of a flat infinite-in-extent ionosphere layer, and then in a model where the ionosphere follows the curvature of the Earth, the mean is not zero. The remote antennae of LOFAR are indeed far enough from the reference antenna to have non-zero mean DDTEC as a result of curvature.

Because we applied the two approximations in Section 3.3.4, our model is no longer tomographic, therefore we should not expect super-resolution DDTEC inference as observed in Albert et al. [2020a]. However, our model still retains the directional coupling of a thin-layer ionosphere model. Therefore, we observe that the reference direction is constrained to have zero DDTEC in the inferred DDTEC screens.

Figure 3.10 shows, side-by-side, several sources in the LoTSS-DR2 archival image, the `smooth_cal+slow_cal` image, and the inferred DDTEC screen image. The column numbers correspond to the region indicated in **Figure 3.2** with red circles, which shows that all selected sources are between calibrators. The LoTSS-DR2 and `smooth_cal+slow_cal` sources display similar dispersive phase errors, which implies that the isoplanatic assumption was violated in that region of the image at some point during the observation. In the inferred DDTEC screen image the dispersive phase error effects appear less pronounced, indicating that the inferred DDTEC screen provided a better calibration than nearest-neighbour interpolation of the calibrator solutions.

In order to quantify the improvement to the image quality caused by applying the inferred DDTEC screen, we systematically compare the scale of artefacts in the vicinity around bright sources within the primary beam. We consider an annulus from $12''$ to $90''$ around all sources brighter than 0.1 Jy beam^{-1} , and calculate the root-mean-squared residuals for the LoTSS-DR2 image, σ_{DR2} , and inferred DDTEC screen image, σ_{screen} , within that annulus. **Figure 3.11** shows the histogram of $\log \sigma_{\text{screen}} / \sigma_{\text{DR2}}$. In total, 76% of the sources have $\sigma_{\text{screen}} < \sigma_{\text{DR2}}$ with a mean reduction in σ_{screen} of 32% with respect to σ_{DR2} .

3.5 Discussion

In P126+65 there is a clear improvement over LoTSS-DR2 resulting from applying our inferred DDTEC screens. Remarkably, we see this improvement using only 35 in-field calibrators compared to the 46 used in LoTSS-DR2. This suggests that the method may have applications to other regimes where calibrators are very sparse, such as in very long baseline interferometry or at ultra-low frequencies.

However, it is unclear whether this observation is typical and if the same method applied to other observations would achieve similar results. One obvious difference between observations is the different layout of sources, and therefore potential calibrators. The distribution of bright sources in P126+65 is relatively uniform over the field of view, which plays to the advantage of our screen inference method. As a result of the approximations in Section 3.3.4, the method is no longer tomographic, and sparsity of calibrators could have a significant effect on screen quality. One of the next steps would therefore be to test the performance on a large number of observations. Furthermore, there are potential ways in which to make the tomographic approach computationally efficient and therefore computationally viable, but these require more research.

There is also the question of how robust this method is to various systematics. For example, it is not yet known if unsubtracted compact or diffuse emission could be absorbed during the calibration, as is known to occur if there are too many degrees of freedom during calibration. Another source of uncertainty is the robustness of this method to the behaviour of the ionosphere. In this observation two distinct characteristics of the ionosphere were observed: one ‘wild’ and one ‘calm’. There are perhaps many more types of ionospheres that could be encountered [e.g. Mevius et al., 2016, Jordan et al., 2017]. Because of the FED hypothesis marginalisation, we expect that any new ionosphere should be manageable as long as the calibrators are not too sparse by incorporating the right FED kernel hypotheses.

Furthermore, while the changing ionosphere character may be handled with model weighting, the quality of the inferred DDTEC screens ultimately depends upon the quality of

the measured DDTEC. If this method were to be applied to LOFAR-LBA or long-baseline data, the robustness of the DDTEC measurements must be guaranteed. For P126+65, 0.1% of all measured DDTEC were found to be outliers, but this could easily be much larger when there is less flux in the field against which to calibrate, as is the case at lower frequencies and on very long baselines, or when the ionosphere coherence time is shorter than the time interval of calibration (tens of seconds).

The improvements to root-mean-squared residuals near bright sources would be more appreciable with lower thermal noise. An interesting avenue on which to test this method would therefore be multi-epoch observations with tens to hundreds of hours of data. In a deep observation, scattering of emission off of the ionosphere can cause low level structure in the image, mimicking astrophysical sources such as cosmic web accretion shocks that are hidden below the thermal noise. Any wide-field statistical study of faint emission should ensure that such effects are properly calibrated.

3.6 Conclusion

We have put forward a method for DD calibration and imaging of wide-field low-frequency interferometric data by probabilistically inferring DDTEC for all bright sources ($>0.01 \text{ Jy beam}^{-1}$) in the field of view, using a physics-informed model. In order to do this, we proposed an HMM with variational inference of the forward distribution for measuring DDTEC with uncertainties from Jones scalars solved against isolated calibrators. Isolating the calibrators was found to be important because this gives the Jones scalars well-defined effective directions. We handled the dynamic nature of the ionosphere by marginalising the probabilistic model over a number of FED kernels hypotheses. We only explored stationary hypothesis kernels, therefore this method may not perform well on ionospheres with TIDs.

We tested the method on a randomly selected observation taken from the LoTSS-DR2 archive. The resulting image had fewer direction-dependent effects, and the root-mean-squared residuals around bright sources were reduced by about 32% with respect to the LoTSS-DR2 image. Remarkably, we achieved this improvement using only 35 calibrator directions compared to 46 used in LoTSS-DR2. While these results are promising, the robustness of the method must be verified on more observations.

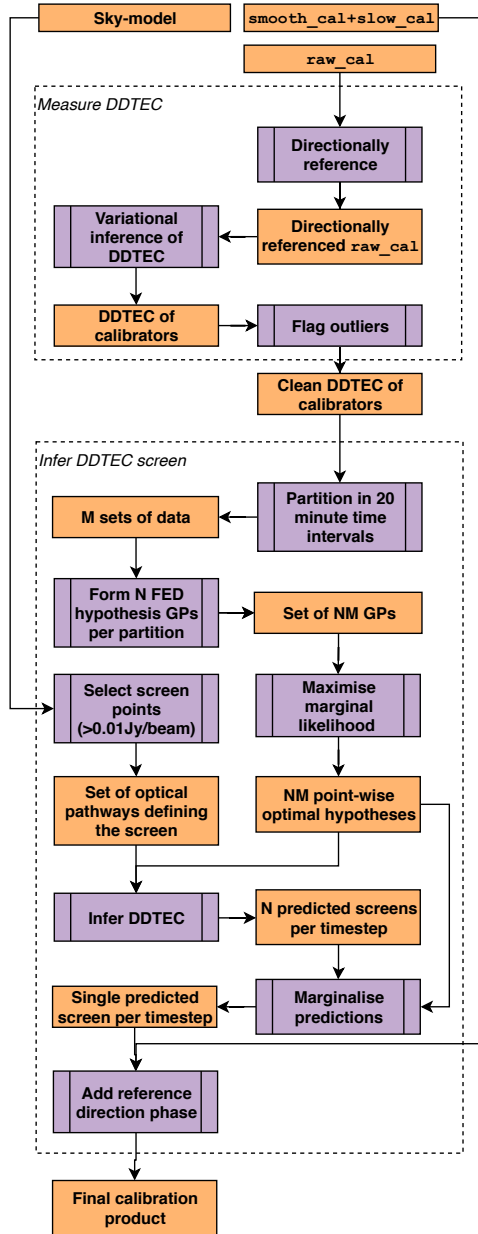


Figure 3.8: Flow diagram of the *measure DDTEC* and *infer DDTEC screen* steps. These steps can be considered the main difference between our method and the method used in the LoTSS-DR1 and DR2 pipelines.

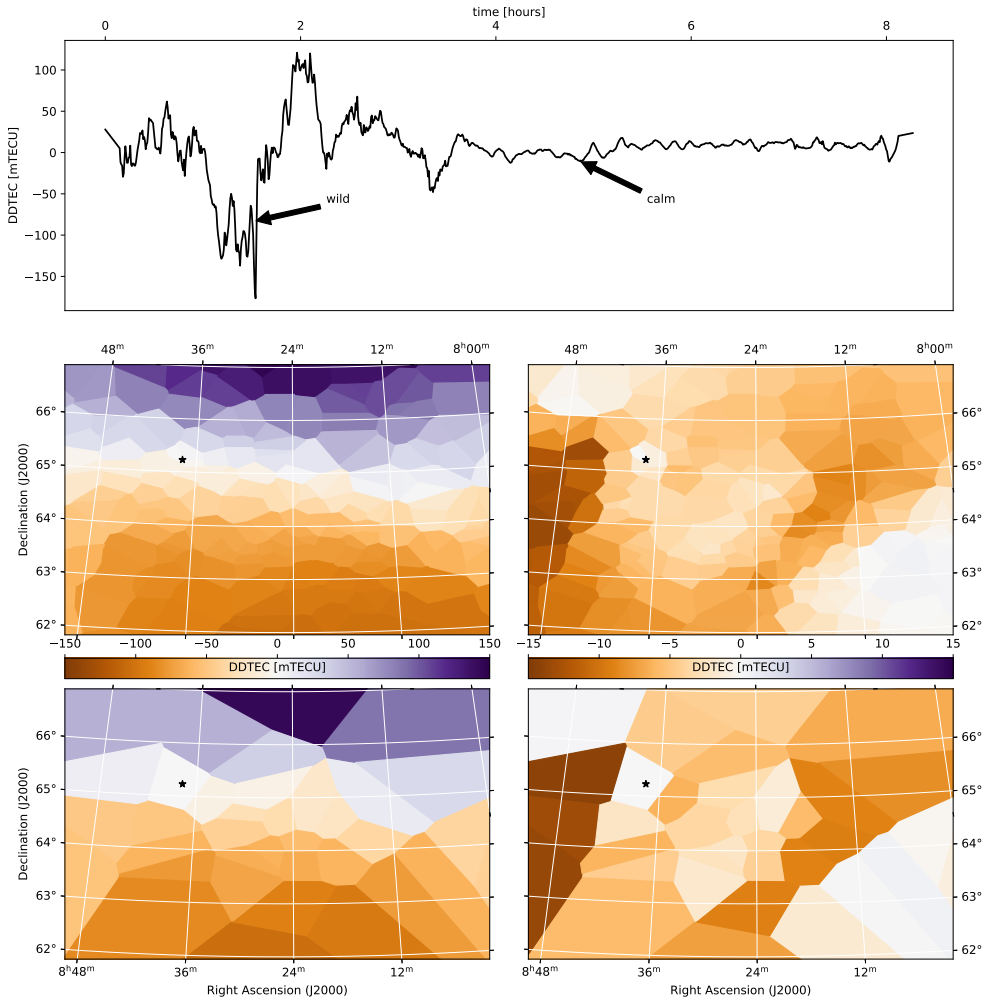


Figure 3.9: DDTEC screens and temporal profile of LOFAR antenna RS508HBA. The lower four plots show the path-length distortions due to the ionosphere from the perspective of antenna RS508HBA during wild and calm periods. The middle row shows the inferred DDTEC screen, and the bottom row shows the measured DDTEC screen. The black star indicates the reference direction, hence the DDTEC is always zero in that direction. The top plot shows the DDTEC for a single direction over the course of the observation, and it clearly shows that the temporal behaviour changed around 3.5 hours into the observation.

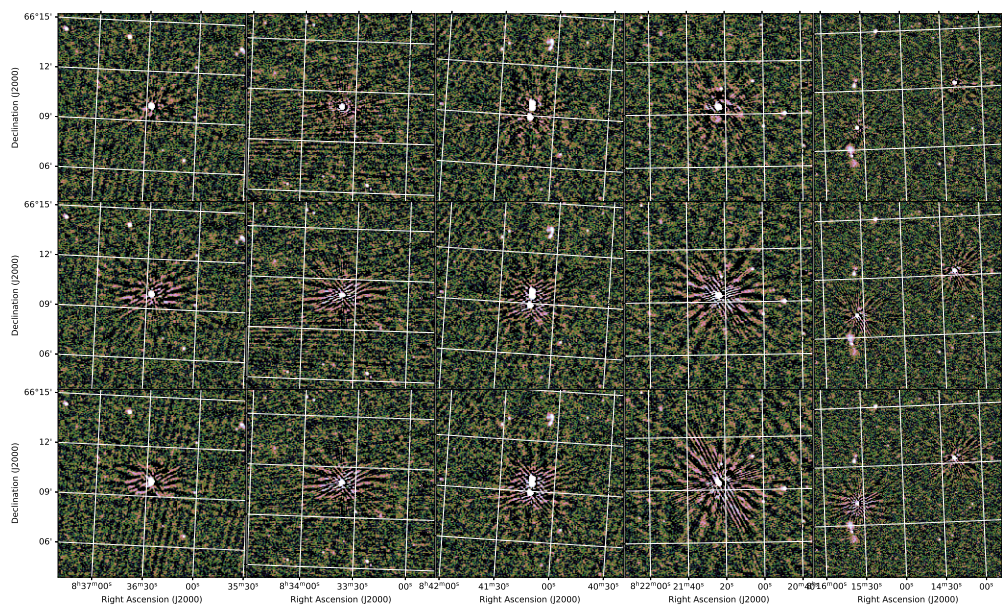


Figure 3.10: Comparison of direction dependent effects in regions far from a calibrator. The bottom row shows the archival LoTSS-DR2 image, the middle row shows the image when `smooth_cal+slow_cal` is applied, and the top row shows the final image with our inferred DDTEC screen applied. The column numbers correspond to the regions labelled in **Figure 3.2**.

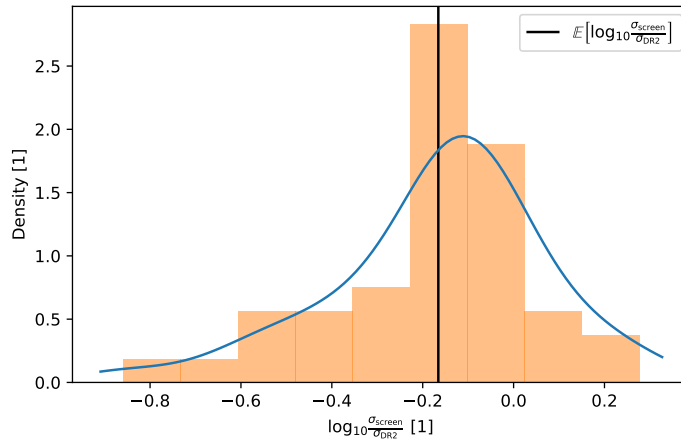
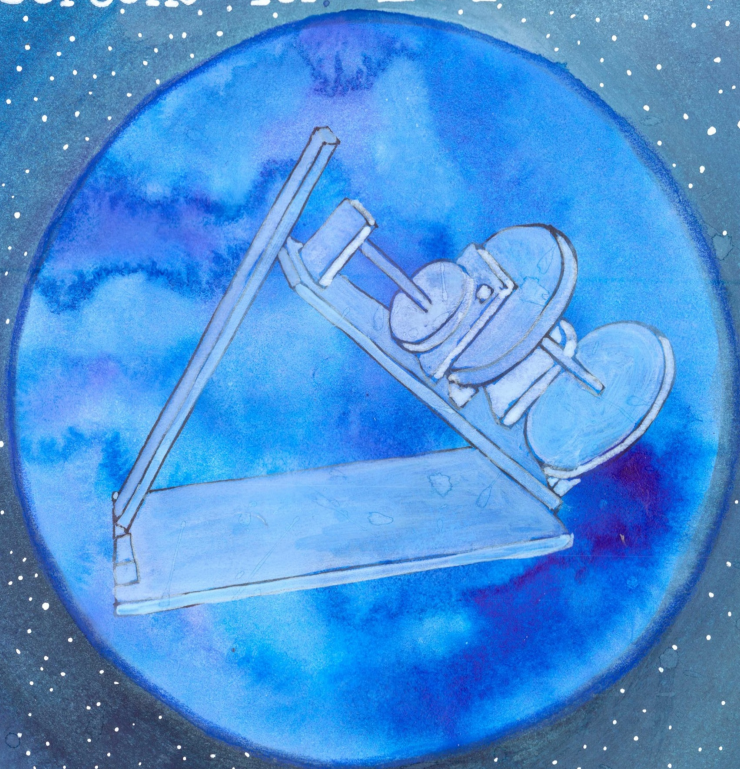


Figure 3.11: Histogram and Gaussian kernel density estimation of the difference in magnitude of root-mean-square residuals in annuli around bright source, screen-corrected, σ_{screen} , over LoTSS-DR2 archival, σ_{DR2} . The root-mean-square residuals are calculated within an annulus between $12''$ to $90''$ centred on sources brighter than 0.1 Jy beam^{-1} . The residuals are primary-beam uncorrected. The black line denotes the mean ratio, and corresponds to $\sigma_{\text{screen}} = 0.68\sigma_{\text{DR2}}$.

MASTER 4

On the Feasibility of Probabilistic Ionospheric
Screens for LoTSS



J.G. Albert
R.J. Van Weeren.
H.T. Intema
AND
H.J.A. Röttgering

On the feasibility of probabilistic ionospheric screens for LoTSS

J. G. ALBERT, R. J. VAN WEEREN, H. T. INTEMA, AND
H. J. A. RÖTTGERING

In preparation for submission

Producing direction dependent calibrated radio images is vital for low-frequency wide-field radio surveys such as the LOFAR Two-Metre Sky Survey (LoTSS). The current facet-based direction dependent calibration of LoTSS is limited by two main factors: the sparsity of suitable in-field calibrators, and ill-conditioning when many in-field calibrators are used. A screen-based method of direction dependent ionospheric calibration was recently proposed in Albert et al. [2020b] which was found to surpass the state-of-the-art direction dependent calibration on a single randomly selected LoTSS data set, using fewer calibrators than used by the LoTSS calibration. The proposed method potentially poses a logical next LoTSS survey improvement if it can be shown to be feasible and robust on a larger sample set. In this paper we apply the method to the deep Lockman Hole data set (12 observations and 100 hours in total) which spans varying times-of-day, season, and varieties of ionosphere. We present a neural-network based outlier detection method that significantly improves over the original outlier detection method. We quantify the improvement to image artefacts around inter-calibrator sources by computing equivalent integration-time gain. We find that for observations with low and high ionospheric activity that the method provides, respectively, an average equivalent integration-time gain near bright sources of 1.3, and 2.0. In two of the observations the improvements are less significant. We suggest that during these two observations small-scale ionospheric structure limits the screen-based method. In no cases is the resulting image quality worse than LoTSS. We apply the same method to calibrate and image the full data set and find that the same improvements scale to deep observations. Our method produces a zoo of ionospheric doubly differential total electron content screens supporting that the ionosphere has many different distinct behaviours. We discover an unmodelled systematic in the Jones scalars, which results in residual artefacts that limit the effectiveness of method. Using a combination of real and simulated Jones scalar data we propose and implement a preliminary correction for it. The proposed correction is found to partially correct for the systematic, and indicates that the general approach is correct and can be improved to completely account for the systematic. Using these results we determine that the method is capable of extending the LoTSS calibration and imaging pipeline, and is also promising for < 100 MHz data.

Acknowledgements. J. G. A. and H. T. I. acknowledge funding by NWO under ‘Nationale Roadmap Grootchalige Onderzoeksfaciliteiten’, as this research is part of the NL SKA roadmap project. J. G. A. and H. J. A. R. acknowledge support from the ERC Advanced Investigator programme NewClusters 321271. R. J. vW. acknowledges support of the VIDI research programme with project number 639.042.729, which is financed by the Netherlands Organisation for Scientific Research (NWO). This research has made use of the University of Hertfordshire high-performance computing facility and the LOFAR-UK compute facility, located at the University of Hertfordshire and supported by STFC [ST/P000096/1]. J. G. A. thanks Aleksandrina Skvortsova for spending many hours clicking through outlier data.

4.1 Introduction

One of the primary objectives of the Low-frequency array [LOFAR; van Haarlem et al., 2013] is to be a wide-field low-frequency surveying instrument of the entire northern hemisphere with unprecedented high sensitivity and resolution. There are numerous challenges involved in realising this goal, which requires efficiently storing and processing petabytes of data. First among these, any good wide-field low-frequency survey requires good direction dependent (DD) calibration and imaging [Cohen and Röttgering, 2009]. Great progress has been made to meet the DD calibration and imaging challenge. The calibration program *killMS* [Tasse, 2014b, Smirnov and Tasse, 2015] applies a clever sparsification of the optimisation problem to overcome computational limitations of DD calibration, and uses an extended Kalman filter to regularise the Jones matrices which improves the conditioning of the problem. The imaging program *DDFacet* [Tasse et al., 2018] is a wide-field wide-band non-coplanar deconvolution algorithm with spatially varying point-spread-function, and allows externally defined Jones matrices to be applied. Together these two programs supply the tools that make up the state-of-the-art DD calibration and imaging solution used by the LOFAR Two-Metre Sky-Survey [LoTSS; Shimwell et al., 2019].

The LoTSS is planned in a tiered program of successive image quality, sky-coverage, and depth. The first data release [DR1; Shimwell et al., 2019] provided the first tier of the survey, presenting images of 5% of the northern sky, with an excellent median sensitivity of $S_{144\text{MHz}} = 71 \mu\text{Jy beam}^{-1}$ and point-source completeness of 90% at integrated flux density of 0.45 Jy. The coming second data release (DR2) will present improved DD calibrated images of a much large fraction of the northern sky, and should improve the median sensitivity as well as improve the detection of diffuse emission. Future data releases will go deeper, and use long baselines to achieve resolutions of $0.3''$. Since we have internal access to the unreleased LoTSS-DR2 archive, we'll make reference to DR2.

Despite enabling one of the most ambitious radio surveys ever conducted, the state-of-the-art DD calibration and imaging method is fundamentally limited by 1) the sparsity of available calibrators in the field of view, and 2) the ill-conditioning of the system. When the angular separation between calibrators is too great, the ionospheric distortions are directionally under-sampled, leading to non-isoplanaticity [Fellgett and Linfoot, 1955], and when there are too many degrees of freedom this leads to 'self-cal bias'¹ [Grobler et al., 2014]. These two limitations are not mutually exclusive; even when the field of view is densely populated with bright calibrators, selecting too many potentially leads over-parametrisation and flux absorption. Indeed, the LoTSS-DR2 choice of 45 calibration directions was chosen to balance these two aspects.

The scattering effects of the ionosphere can be seen as a diffuse halo around sources, which are not necessarily visible above the noise [Vedantham and Koopmans, 2015] and can significantly impact studies sensitive to the power-spectrum of faint diffuse emission [e.g. Harrison et al., 2016, Patil et al., 2017, Vernstrom et al., 2017, van Weeren et al., 2019]. Some part of this scattering can be corrected via DD calibration, however there is always a remaining component which adds noise to the image, which depends on the spatio-temporal power-spectrum of ionospheric FED and the availability of bright enough calibrators. For LOFAR this noise-like component is expected to be of a level comparable to thermal noise [Vedantham and Koopmans, 2016]. The DD correction performed by *killMS* is a facet-based

¹The term 'self-cal bias' was coined by Ger de Bruyn.

algorithm that assumes each facet is an isoplanatic patch. When this is true, then all sources in a facet can be calibrated with a single number per antenna, to good approximation, removing most ionospheric effects and leaving only a thermal noise-like halo residual. However, since the ionosphere is highly dynamic often the isoplanatic assumption is violated and a facet fails to be well calibrated.

Recently, Albert et al. [2020b] proposed a probabilistic DD calibration and imaging approach for ionospheric calibration of LOFAR high-band antennae (HBA; 115–189MHz) radio interferometric data that alleviates the issue of the sparsity of calibrators. The method is based on inferring doubly differential total electron content [DDTEC; described in Albert et al., 2020b] using a probabilistic non-diffractive tomographic technique [Albert et al., 2020a]. The method was shown to significantly improve DD dispersive phase errors between calibrators on a single randomly selected observation taken from the LoTSS-DR2 archive. In particular, the method was shown to reduce the root-mean-squared residuals by 32% within $90''$ of inter-calibrator, bright (peak flux $> 100 \text{ mJy beam}^{-1}$) sources in comparison with the archival LoTSS-DR2 image. For comparison, the reduction in DD scattering artefacts, around these inter-calibrator sources, is equivalent to observing for approximately twice as long. If this result can be expected when the method is applied to most observations, then it prompts the possibility of reprocessing the entire LoTSS-DR2 archive (13 000 hours of observing time) with these inter-calibrator ionospheric artefacts suppressed. However, before proposing such a computationally expensive reprocessing the feasibility of the method on multiple nights needs to be assessed, as well as an understanding and amelioration of the systematic biases of the method.

Understanding the systematic biases of the method is vital for releasing a product that the scientific community can trust. Such systematic biases include flux absorption due to an incomplete sky model and over-parametrisation, dispersive phase and amplitude errors due to an incomplete beam model, and incorrect DDTEC inference due to model misspecification of the Jones scalars. Since this method introduces no extra degrees of freedom and maintains the same calibrator sparsity as LoTSS, we expect the same source completeness properties as the current LoTSS-DR2 processing, and we do not investigate it here.

In this paper we examine 1) the method effectiveness in a larger sample of observations as well as on a multi-epoch (100 hours) imaging application, 2) the robustness to ionospheric conditions, and 3) systematic biases due to model misspecification. In Section 4.2 we describe the data and calibration procedure. In Section 4.3 we assess the method individually on 12 data sets with identical points taken from the LoTSS archive. We also combine all data sets and produce a deep image, which we compare to LoTSS. We then we quantify the robustness to differences in ionospheric conditions. Finally, in Section 4.4 we investigate a systematic due to model misspecification using a combination of real and simulated Jones scalar data, and then present a correction for this systematic.

4.2 Method

In this paper, we are primarily focused on assessing our method for robustness to ionospheric conditions. To that end, we limit the variation due to calibrator distribution, sky model completeness, field declination, and proximity to bright sources. The optimal experimental setup is therefore to compare the method on the same field over multiple observations. We

select the deep Lockman Hole ($162^\circ, 58^\circ$) data set (12 observations and 100 hours in total) for this purpose, whose observation dates and times are shown in **Table 4.1** along with total daily Sun-spot count. The selected data set can be divided into two sample sets with six samples in each category: high Sun-spot count at dawn (Summer of 2018), and low Sun-spot count at dusk (Spring of 2015).

Sun-spot counts are known to correlate strongly with ionospheric free electron density (FED) due to extreme ultra violet ionising radiation from the Sun [Kiepenheuer, 1946], and therefore are an approximate measure of the expected ionospheric conditions. As can be seen from the Sun-spot counts, the observations from Summer of 2018 are expected to have a lower FED compared to the observations from Spring of 2015. We note that high Sun-spot count, and thus FED, does not necessarily imply temporal or spatial characteristics of the ionosphere, as these qualities are largely influenced by the bulk motion properties of the Earth's atmosphere. In fact, the local time-of-day has the largest impact on the temporal and spatial properties of the ionosphere, for example it has been suggested that scintillation is more pronounced near sun-rise due to increased FED variation [e.g. Spoelstra, 1983]. Given the different season and time of day of the data sets, they probe two vastly different ionosphere varieties.

Table 4.1: Date and time of 12 Lockman Hole 8 hour observations selected for method analysis

Obs. ID	Start time (UTC)	Total daily Sun-spots
667218	2018-09-13 07:05:31	0 ± 0
667204	2018-09-12 07:06:28	14 ± 1.1
664480	2018-08-19 08:39:05	14 ± 0.9
664320	2018-08-15 08:49:19	13 ± 0.9
659948	2018-07-12 11:08:29	0 ± 0
659554	2018-07-10 11:11:19	0 ± 0
342938	2015-05-08 14:50:43	150 ± 10.2
340794	2015-04-25 17:08:19	69 ± 6.2
299961	2015-03-24 17:47:39	108 ± 5.9
294287	2015-03-21 19:11:19	26 ± 2
281008	2015-03-14 18:26:58	55 ± 4.7
274099	2015-03-08 20:11:19	28 ± 1.9

Note: Sun-spot uncertainties are standard deviations of measurements from solar observatories around the world.

For the purpose of discussion in this current work we provide a summary of the steps of the method from Albert et al. [2020b]:

1. *Subtract and solve step.* Subtract a good model of the sky from the visibilities, except for a set of bright calibrators. Solve against the isolated calibrators.
2. *Smooth and slow-resolve.* Smooth the Jones scalars, and resolve on a long time scale to simultaneously improve the conditioning and solve the holes problem. For description of the hole problem see Shimwell et al. [2019]. The phases are smoothed in frequency by fitting a phase model, and amplitudes in time using a median filter with 15 min. filter window.

3. *Measure DDTEC.* Infer the DDTEC of the Jones scalars with a variational hidden Markov model. Perform outlier detection and flagging on the inferred DDTEC.
4. *Infer DDTEC screen.* Apply model marginalised Gaussian process regression on the measured DDTEC, with a physically motivated DDTEC covariance function, to infer DDTEC for a screen of directions covering the field of view.
5. *Image.* Image the original visibilities with a concatenation of the smoothed and screen solutions.

With the exception of isolating the calibrators before solving, the first two steps can be seen as equivalent to the LoTSS DD calibration, and steps 3 and 4 can be seen as an extension of the LoTSS DD calibration pipeline. A good quality of an extension to an existing pipeline is that it does not worsen the data product with respect to the preceding part of the pipeline. The calibrator selection criteria in step 1 of the original method was peak flux $> 0.3 \text{ Jy beam}^{-1}$ separated by at least $6'$. Depending on the field, this selection could result in less than 20 to more than 50 in-field calibrators, which significantly impacts the inter-observational DDTEC screen consistency. To that end, in this paper we choose to standardise step 1 by choosing the same number of calibrators as used by LoTSS – the brightest 45 directions separated by at least $6'$. Our calibrator selection is not not exactly the same as in LoTSS. In LoTSS, the direction of calibrators is not necessarily centred on a source, whereas our calibrator directions must precisely encircle a bright source. This can be seen in **Figure 4.6a**, where the LoTSS calibrators do not centre on any particular source. This is because LoTSS calibrates against all the sources within a facet, and we calibrate against isolated bright sources. However, by choosing the same number of calibrators, and ensuring they are sufficiently separated, the average distance between our calibrators is the same as in LoTSS.

The Lockman Hole data sets have been previously jointly calibrated and imaged as part of LoTSS-DR2. Therefore, we have a deep sky model that is a factor of approximately 3.5 deeper than the sky model for an individual eight-hour observation. We use this deep sky model to perform calibrator isolation in step 1. In DDFacet, deconvolution is performed using hybrid matching pursuit with a genetic algorithm that locally optimises deconvolution around each clean component [Tasse et al., 2018]. The sky model resulting from this process may contain scattered negative components, especially around bright sources, which are called sky model artefacts. Failure to include these sky model components in the subtraction mask can cause undefined distortion of the Jones scalars solved for. We manually ensure that the subtraction mask in step 1 encompasses all sky model components of the calibrator source. In particular, radio galaxy 3C244.1 is the brightest source in the field, and has scattered sky model components that would fall outside the default subtraction mask, a circle with radius $120''$ around each calibrator. We used a $240''$ radius mask for this calibrator.

4.2.1 An improved outlier detection

As stated in Albert et al. [2020b], outlier detection and flagging of the DDTEC inferred in step 3 is extremely important for the performance of the method, since the Gaussian process regression (step 4) depends on accurate estimates of the DDTEC uncertainty. That is, the method is robust to outlier DDTEC values so long as the corresponding uncertainties are large enough to reflect this. In the original method outlier detection was done with a heuristic

method using radial basis functions that detects outliers based on spatial similarity, however this method resulted in too many false negatives when applied to the Lockman Hole data sets. To this end, we manually identified 4531 outliers and 74460 non-outliers then designed and trained a neural network to perform classification.

The input to the neural network is the time series of DDTEC mean and log-uncertainty of the K nearest neighbours of an optical pathway under consideration, as shown in **Figure 4.1**. We choose $K = 15$ (1/3 calibrator directions in the field of view), i.e. a DDTEC value is classified as an outlier based on the nearest 15 calibrators in the antenna's field of view. This choice of input reflects how humans visually determine what is an outlier. In particular, in order for a DDTEC value to be classified as an outlier by a human we require 1) that it visually 'looks out of place' in the context of the DDTEC screen, or 2) the temporal evolution of that optical path 'looks out of place' when compared with neighbouring calibrators. These two criteria correspond to being spatially and temporally 'out of place'. It is important that we consider these two criteria, because sometimes when there are many outliers in an antenna's field of view the spatial information alone cannot discriminate between outlier and non-outlier.

The neural network architecture is an ensemble of eight 10-layer residual temporal à trous convolutional neural networks [CNN; e.g. He et al., 2015]. Each layer has a kernel of shape (*kernel*, *filters*) corresponding respectively to how wide the convolution window is, and the number of features, similar to channels in an image. We choose *kernel*=3, and a *filter*=48. The non-linearities of each CNN is the function $f(x) = \max(0, x)$. After we apply the non-linearity we down sample the data according to a rule known as a pooling operation. Each layer is a residual layer, i.e. the input to the layer is added to the output after applying the non-linearity and pooling operation. Each layer is an à trous convolution [Yu and Koltun, 2016] which has inspiration from wavelet decomposition. In à trous convolution the kernel window is expanded by a parameter called the dilation rate. Roughly speaking this controls how sensitive the layer is to diffuse signal. Each member of the ensemble has a different combination of pooling operation and dilation rate which gives each member a unique structure that is sensitive to different types and scales of structure in the data. The set of eight combinations of pooling and dilation rate is $\{\text{MaxPool}, \text{AvgPool}\} \times \{1, 2, 3, 4\}$. The idea of using an ensemble of classifiers is that it averages away the bias of a single classifier, which is the motivation behind many ensemble-based methods such as random forests.

A final layer projects the input to *filter*=1, which is interpreted as the logit of a binary classification problem. To facilitate learning we add a constant to the output of each CNN equal to the log prior probability of the training set, which is -3.6. This means that when the CNN outputs zero the logits will correspond to the prior logits and the neural network only needs to learn deviations from the prior. The loss function of each CNN is weighted binary cross-entropy. We weight the positive classifications by the ratio of non-outliers to outliers, which is approximately 36, so that the average gradient magnitudes of all outlier examples is equal to that of the non-outlier examples. This ensures that we do not preferentially learn to classify non-outliers.

We then jointly train the whole ensemble with Adam stochastic gradient descent [Kingma and Ba, 2014] using a mini-batch size of 32, time-segments of 50 minutes, and learning rate of 10^{-4} . We mask out the optical paths without ground truth labels. We train for 20 epochs which takes approximately 6 hours on a 32 core CPU. Following training we set the classification threshold of each member of the ensemble to the optimal value using the receiver operating characteristic (ROC) curve. The ROC curve is a plot of the FPR against FNR as a function of

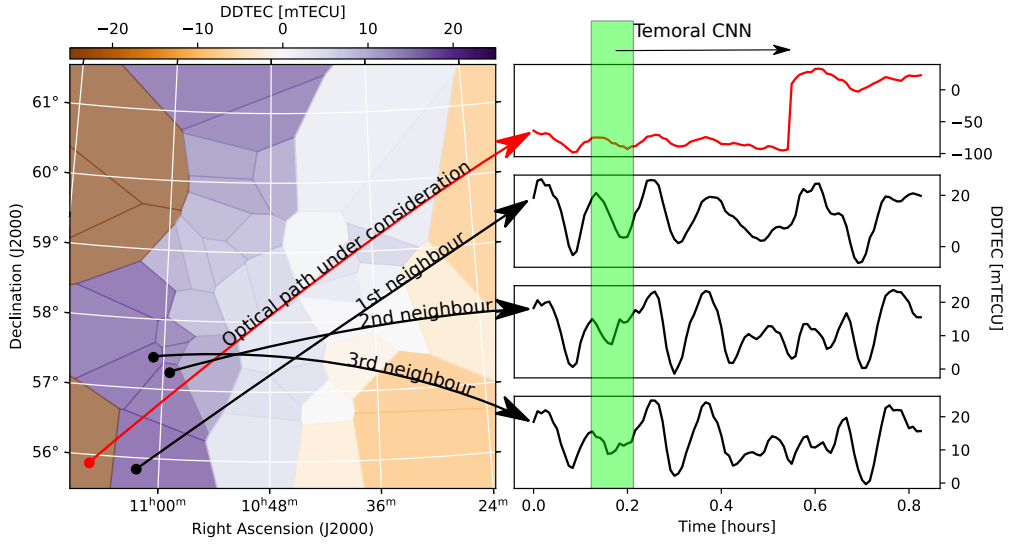


Figure 4.1: Pictorial description of the input to the neural network. For each optical pathway we take the K nearest neighbours in the field of view, measured by great circle separation, and stack the time series of measured DDTEC and log-uncertainty. An ensemble of temporal CNNs are applied to the input and the outputs are trained for logistic regression per time step. The ensemble median classification is used to identify outliers.

Table 4.2: Comparison of outlier detection methods.

Obs. ID	Neural network method				Heuristic method			
	FNR	FPR	Est. FN	Est. FP	FNR	FPR	Est. FN	Est. FP
667218	1.5%	11.2%	728	274 474	45.2%	5.0%	21 999	122 109
667204	0.4%	4.4%	125	113 679	54.8%	6.1%	19 387	155 669
664480	0.1%	10.4%	49	275 810	58.7%	11.4%	22 768	300 110
664320	1.4%	18.3%	1 500	461 773	49.1%	8.2%	51 741	206 502
659948	1.5%	6.0%	1 247	150 594	45.8%	5.1%	38 239	128 304
659554	0.6%	7.5%	452	195 082	39.3%	7.2%	28 220	187 883
342938	2.2%	11.2%	1 056	320 300	42.3%	5.7%	20 422	161 647
340794	5.9%	5.4%	6 148	152 023	43.7%	4.2%	45 270	118 631
299961	2.1%	2.8%	806	76 993	43.6%	3.2%	17 068	88 911
294287	12.4%	6.9%	25 327	185 568	43.7%	6.9%	89 487	186 491
281008	1.6%	4.2%	532	119 309	42.4%	3.7%	13 705	107 378
274099	4.7%	9.9%	17 329	250 393	48.8%	5.9%	181 955	150 008
Total	2.6%	7.0%	55 298	2 575 997	48.0%	5.5%	550 261	1 913 642

Note: On average there are 2 749 800 optical paths per observation.

Note: FNR is false-negative rate, FPR is false-positive rate, FN is false negatives, and FP is false positives.

the threshold. We choose the threshold for each member of the ensemble that minimises FNR plus FPR. At deployment time we take the median classification from all members of the ensemble as the classification.

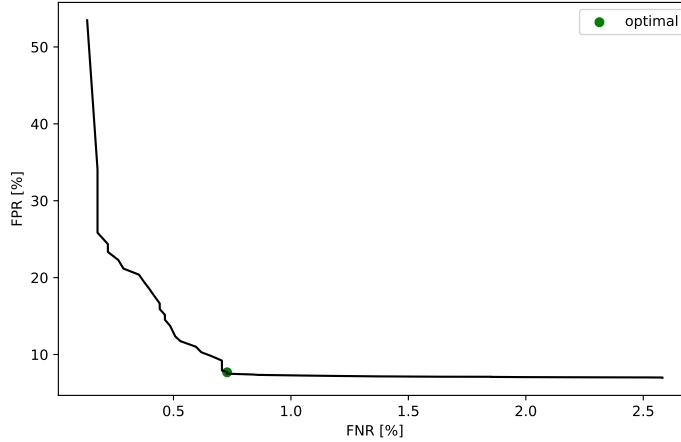


Figure 4.2: Receiver operating characteristic curve as a function of the threshold for the number of neural-network identified outliers in an antenna’s field of view. We flag an antenna’s entire field of view if the number of neural-network identified outliers is above a threshold. We value low FNR without significantly raising FPR. The optimal threshold of 30 directions gives a FNR of 0.7% and FPR of 7.7%.

A comparison of performance of the heuristic method to the neural network is shown in **Table 4.2**. The average false-negative rate (FNR) of the heuristic and neural network methods is 48% and 2.6% respectively. False negatives (FN), being outliers that are missed, have a much larger impact on performance than false positives (FP), which are accidentally flagged non-outliers. Therefore, the neural network method provides an order of magnitude improvement.

Using the FNR, FPR and the number of neural network classified outliers we can estimate the number of FN and FP in the neural network classifications. Despite the significant improvement of the outlier detection, we estimate that approximately 55 298 outliers were missed. We observe that most of these FN occur when there are a large number of outliers detected in the same field of view. This suggests that when the field of view has too many outliers it is difficult for the neural network to decide which data points are not outliers.

We use this intuition to motivate a secondary rejection step. We reject an antenna’s entire field of view if the number of neural-network identified outliers in the field of view is above a certain threshold. When an antenna’s entire field of view is flagged we flag all corresponding baselines in the visibilities involving the antenna under consideration. This rejection mechanism requires no extra computation since it uses the results of the neural network classifications. **Figure 4.2** shows the receiver ROC curve for this rejection mechanism. The ROC curve plots the FPR against FNR as a function of the threshold. We identify a

threshold that meets our requirements of low FNR without raising the FPR significantly. From the ROC curve we choose an optimal threshold of 30 directions. This results in a FNR of 0.7%, which is a factor of almost four in reduction, and a FPR of 7.7%, which is unchanged.

4.3 Image improvements and ionospheric robustness

The method significantly improved the DD errors in 10/12 observations, while 2/12 showed only moderate improvements. **Figure 4.3** shows typical examples of how inter-calibrator DD errors are corrected by the method. In observations 274099 and 297287 many inter-calibrator DD effects still remained after applying the method, despite verifying that the screens looked sensible. Note that the resulting image quality in these observations is still at least as good, or better, than the LoTSS-DR2, i.e. DD errors did not get worse.

There are several possible explanations why these two observations did not show significant improvements. One is that the calibrator layout, with an average inter-facet separation of 38', may under-sample small scale ionospheric structure during these observations. This separation is equivalent to ionospheric scales of 2 km to 5 km, which occasionally occur in the ionosphere [Yeh and Swenson, 1959, Mevius et al., 2016]. Another possibility is that non-stationary structure due to travelling ionosphere disturbances (TIDs) [e.g. van der Tol, 2009] may have been prevalent in these observations. Our method currently does not support non-stationary disturbances, though extensions are possible to account for such behaviour. Alternatively, the thin-screen approximation described in Albert et al. [2020b] may not be valid in these observations. Koopmans [2010] shows that wide-field low frequency arrays should take into account the 3D ionosphere. In our application we take a full 3D non-diffractive tomographic model in the limit of a thin-ionosphere for computational reasons, although there are possible optimisations that might make 3D modelling feasible. Finally, we observe in **Table 4.2** that these two observations have a much higher number of estimated FN, suggesting that the Jones scalar data may have been of sub-standard quality resulting more missed outliers.

In **Figure 4.3** we observe that the spoke-like patterns, which originate from uncorrected ionospheric scattering, are corrected by the method. We observe that around all corrected sources, especially the right-most one, there are radially asymmetric arc-like patterns. These arc-like patterns appear in varying degrees in all observations. We analyse and propose a correction for these arc-like patterns in Section 4.4.

To quantify improvement to DD errors we measure the root-mean-squared residuals (RMSR) around inter-calibrator bright sources before and after applying the screen. To measure RMSR, we select the brightest 66 non-calibrator compact sources across the field of view and mask an annulus with inner radius set to avoid the source flux and outer radius set to 45". If there are nearby compact or diffuse structures these are excluded from the mask. The RMSR is then computed as the square-root of the mean of the square of the flux density in the mask. We denote RMSR in the screen-corrected image as σ_{screen} , and $\sigma_{\text{no-screen}}$ as the RMSR in the image with same calibration procedure as LoTSS-DR2. We refer to $\sigma_{\text{background}}$ as the background noise.

From the RMSR we compute the equivalent integration-time gain (EITG), as the relative extra amount of integration time that would be required to achieve the RMSR of the screen, σ_{screen} , when using the LoTSS-DR2 calibration method. Assuming that scattering effects occur

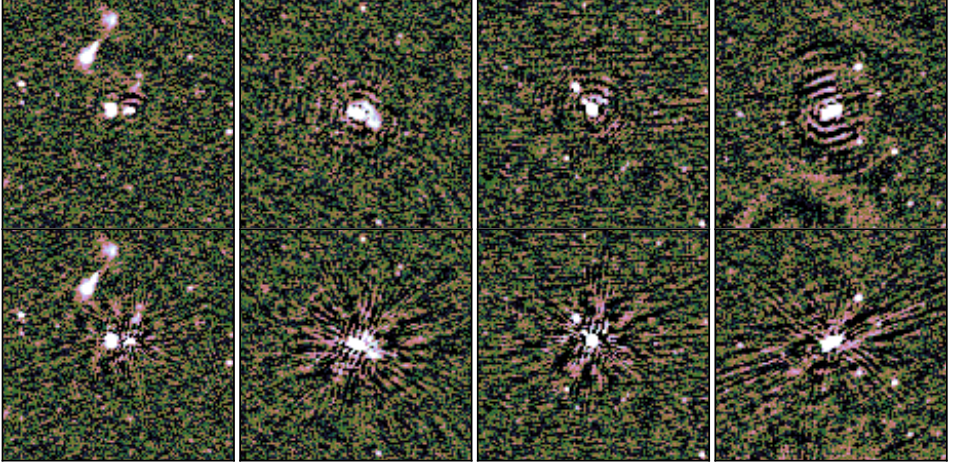


Figure 4.3: Visual example of how the method corrects inter-calibrator DD errors. These sources are taken from observation 342928. The lower panels are images with 45-direction calibration using the same method as LoTSS-DR2, and the upper panels are images with DDTEC screens applied. The mean background noise is $70 \mu\text{Jybeam}^{-1}$.

uniformly random throughout an observation then the EITG is given by $(\sigma_{\text{no-screen}}/\sigma_{\text{screen}})^2$.

Figure 4.4 shows a histogram of the EITG for the two sample sets as well and also plots the EITG of the individual observations against the total daily Sun-spot counts. We observe that the EITG in both sample sets are peaked above one and have long tails extending to higher values. There is an apparent difference in the histograms of the two sample sets. The high Sun-spot (Spring 2015) sample set has a mean EITG of 2.0 while the low Sun-spot (Summer 2018) sample set has a mean EITG of 1.3, and the tail of the prior one at larger EITG is heavier. Given that the Spring 2015 is closer to Solar maximum, has high Sun-spot counts, and is at dusk, the ionospheric FED is likely much larger than in the Summer 2018 observations, therefore scattering is likely stronger. Indeed, the scattering artefacts in the Summer 2018 images are far weaker. These results suggests that the method corrections are more significant during times of high ionospheric FED and less significant when the ionosphere FED is lower, which are equivalent to stronger and weaker scattering respectively. This is also compatible with the result of Albert et al. [2020b].

This result is partially explained by the fact that observations with less scattering have fewer DD errors to correct, therefore the maximal achievable EITG is lower. This necessarily constrains the EITG in the Summer 2018 sample set to lower numbers than the EITG in the Spring 2015 sample set, regardless of the model accuracy. Thus, we cannot rule out the possibility that the model is not accurately modelling the underlying DDTEC in the Summer 2018 sample set. For example, there is likely more turbulence near dawn [e.g. Spoelstra, 1983] which produces small scale structure that the average calibrator spacing of $38'$ does not resolve. Since the scattering effects are very weak in the Summer 2018 observations,

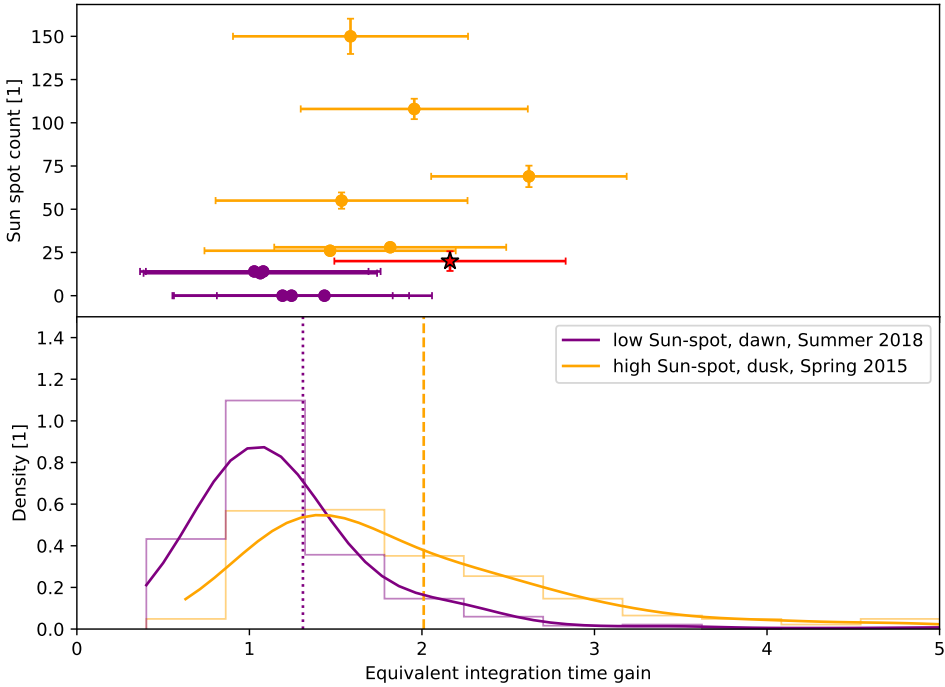


Figure 4.4: Upper panel: Total daily Sun-spot count individual observations plotted with respect to the EITG. The red star is the data point from Albert et al. [2020b]. Lower panel: Histograms and kernel density estimation of ETIG aggregated per sample set. The dashed line indicates the mean of 2.0 for the histogram for the Spring 2015 sample set. The dotted line indicates the mean of 1.3 for the histogram for the Summer 2018 sample set.

sometimes with DDTEC values only a few times the uncertainty, it is not possible to investigate this possibility.

We can effectively characterise the model efficacy independent of the ionospheric conditions by considering the EITG as a function of the strength of ionospheric scattering. As a proxy for ionospheric scattering strength we use the ratio $\sigma_{\text{no-screen}}/\sigma_{\text{background}}$. **Figure 4.5** plots EITG against this proxy of ionospheric strength. We observe that when the $\sigma_{\text{no-screen}}$ is close to the background noise the EITG is approximately unity, indicating that there is no improvement from the model. Conversely, when the $\sigma_{\text{no-screen}}$ is an order of magnitude larger than the background noise the EITG is in the vicinity of three. There is a large scatter in the relation, however a power-law fit indicates that the $\text{EITG} \propto \sqrt{\sigma_{\text{no-screen}}/\sigma_{\text{background}}}$. The ideal model that completely removes all scattering effects would produce a power-law index of two. The deviation from this ideal relation is likely explained by at least two things. Firstly, there is always a low-level halo of speckles around sources, proportional to the brightness of the source, due to ionospheric scattering on sub-solution interval timescales. Secondly, as mentioned already, in Section 4.4 we investigate a systematic which contributes to the residuals around sources.

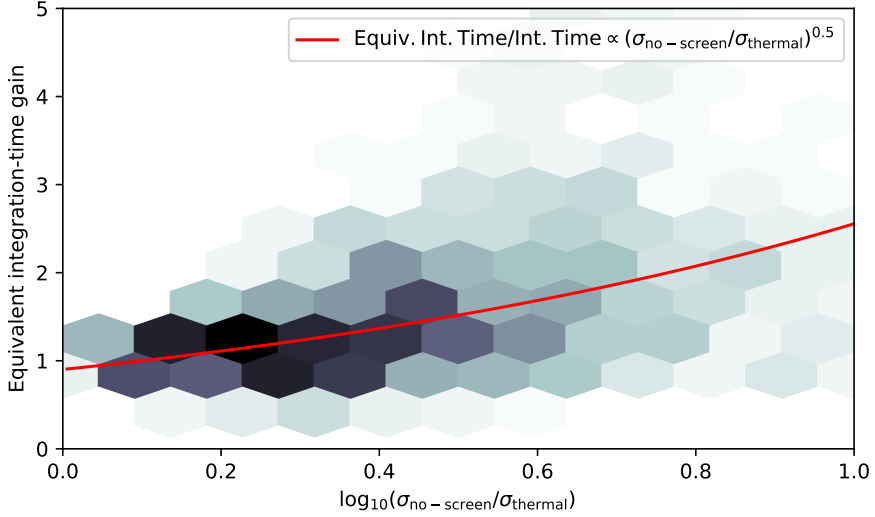


Figure 4.5: 2D Histogram of the EITG and ionospheric scattering strength proxy, $\sigma_{\text{no-screen}}/\sigma_{\text{background}}$, aggregated over all observations. The red line shows a power-law fit.

The scatter also slightly goes below one for $\sigma_{\text{no-screen}} \lesssim 3\sigma_{\text{background}}$. This can be explained by the enhanced effect of small errors or systematics at low noise levels. Specifically, suppose the flux of the pixels in the masked regions where RMSR is measured is composed of the background flux and the artefact flux, with root-variances $\sigma_{\text{background}}$ and $\sigma_{\text{artefact}} = \alpha\sigma_{\text{background}}$ respectively. The artefact flux can be due scattering as well as modelling systematics. Then, since the background and artefacts are independent we have that the EITG is,

$$\text{EITG} = \frac{1 + \alpha_{\text{no-screen}}^2}{1 + \alpha_{\text{screen}}^2}, \quad (4.1)$$

where $\alpha_{\text{no-screen}}$ and α_{screen} correspond to before and after applying the screen model, respectively. Note that α_{screen} contains systematics imposed by modelling, which $\alpha_{\text{no-screen}}$ is free of. When $\alpha_{\text{no-screen}}$ and α_{screen} are small then a modelling error can lead to EITG less than one. For example, suppose $\sigma_{\text{no-screen}} = 3\sigma_{\text{background}}$, then a modelling error of $2\sigma_{\text{background}}$ implies EITG less than one.

We jointly imaged all observations producing a deep image. **Figure 4.6** shows the comparison between our deep image and the LoTSS-DR2 archival deep image. Since the calibrator layout is different between the LoTSS-DR2 image and our image, there are two comparisons we make.

The first is a comparison of scattered flux near calibrators. Since LoTSS calibrates against all flux in a facet, the solution does not perfectly correct any particular source. Instead, the solution partially corrects all sources in the facet, with boundaries indicated by cyan lines in **Figure 4.6a**. Compare this with our calibrator sources, inside red circles, which are all well corrected. This reflects the fact that we isolate the calibrators before solving.

One of the benefits of not isolating calibrators is that all solutions in a facet will be partially improved, and this is a good strategy when a screen-based method is not available. If we were to image with only the isolated calibrator solution then we would see very clean calibrators but worse inter-calibrator scattering than in the LoTSS image. Therefore, when no screen-based method is available calibrating against a facet is appropriate. However, calibrating against a whole facet can fail entirely when the isoplanatic patch size becomes much smaller than the facet.

The second comparison we make is that of the scattered flux around inter-calibrator sources. **Figure 4.6b** shows some examples of the improvements made by our model. We observe that the spoke-like artefacts are mostly gone, indicating that the effects of the ionosphere have been mostly removed. The majority of the remaining artefacts are from the modelling systematic mentioned earlier, which are also visible in the individual observations, in **Figure 4.3**. These artefacts are typically lower level than the scattering related artefacts, moreover we expect to be able to characterise and remove this modelling systematic.

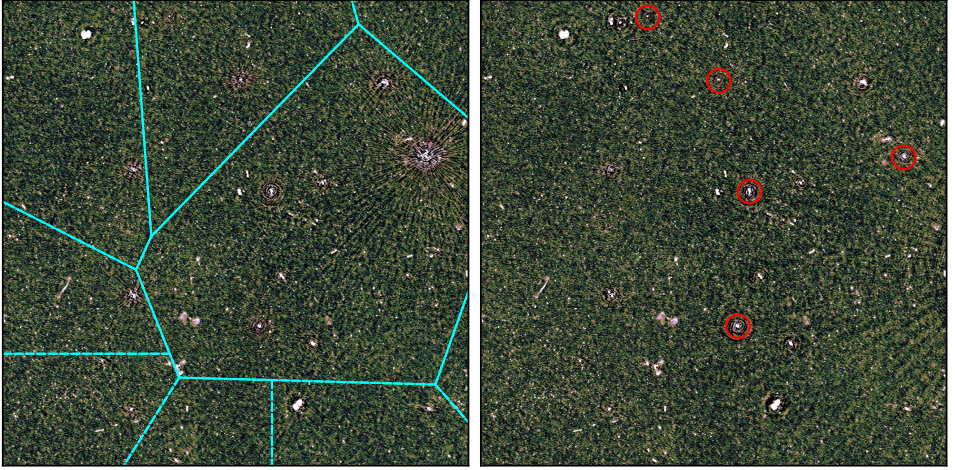
4.4 A non-ionospheric systematic and correction

In **Figures 4.3** and **4.6b** we observe asymmetric arc-like artefacts around inter-calibrator sources after applying our screen-based calibration. These artefacts do not appear uniformly across the field of view, and are less severe in some observations. For example, they were not noticeable in Albert et al. [2020b]. The asymmetric pattern suggests that this is a phase-effect, and the arc-like pattern suggests that it varies on the scale of hours. If it varied on shorter timescales then the patterns would be more radial as ionospheric scattering artefacts are. Therefore, this suggests that the phase component of the Jones scalar model is missing a slowly varying component.

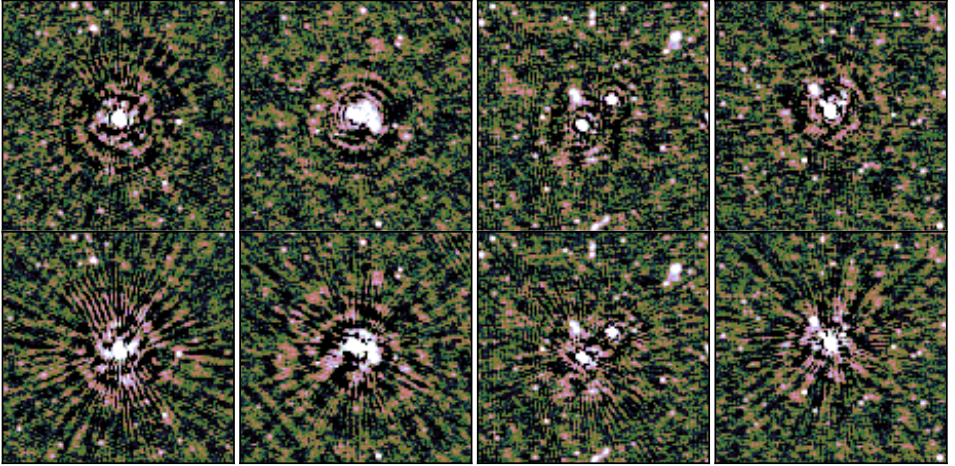
We observe that both in LoTSS-DR2 archival images as well as our images without screens applied, e.g. the LoTSS-DR2 deep image in **Figure 4.6** and the lower panels in **Figure 4.3**, that there are no arc-like artefacts. This implies that the Jones scalar smoothing performed in LoTSS calibration and step 2 of our method can account for this systematic. In LoTSS calibration the phase model is a two-parameter linear model, a TEC-like $\propto \nu^{-1}$ term and a constant term. In step 2 of our method the phase model is a three-parameter linear model, a TEC-like $\propto \nu^{-1}$ term, a constant term, and a clock-like $\propto \nu$ term. Both LoTSS and step 2 of our method assume slowly changing amplitudes. Note, we introduced a clock-like term upon noticing residual direction independent components in the Jones scalars [see Albert et al., 2020b]. This is fundamentally why we perform directional referencing before inferring DDTEC.

The performance of the method depends on the Jones scalars being modelled correctly, since a DDTEC screen inferred from biased DDTEC will lead to biased screens. The model assumes that the only DD contributions to the Jones scalars are slowly varying amplitudes due to beam errors, and phases $\propto \nu^{-1}$ due to non-diffractive weak scattering in the ionosphere. The temporal aspect of the Jones scalars is modelled with a hidden Markov model using variational inference assuming a linear phase model containing only DDTEC [Albert et al., 2020b].

Global optimisation is guaranteed via a clever basin-hopping routine. This is based on the analytic form of the variational expectation in [Albert et al., 2020b]. The variational



(a) Central region of deep image. Left: LoTSS-DR2 deep archival. Right: our deep image.



(b) Cut-outs of inter-calibrator sources. Bottom row: LoTSS-DR2 deep archival. Top row: our deep image.

Figure 4.6: Comparison between the LoTSS-DR2 archival deep Lockman Hole image and the same data calibrated with our screen-based method. Our deep image does not apply the systematic correction proposed in Section 4.4. Cyan lines and red circles correspond to the LoTSS calibration regions and to our calibrators, respectively. The mean background noise in both the LoTSS-DR2 archival deep image and our deep image is $27 \mu\text{Jy beam}^{-1}$.

expectation has many local maxima due to phase wrapping. Since we have the analytic form we know precisely how far apart the local minima are, therefore we can simply hop between them to find the global maximum (which is unique).

We search for a slowly varying unmodelled phase systematic, by plotting the residual between the posterior mean phase and the phase of the data. The upper panel of **Figure 4.7** shows the residual phase over time of a single antenna and direction following Jones scalar modelling. We observe a low-order (in the sense of polynomials), symmetric in frequency, slowly varying underlying structure in the residuals. The phase residuals in this example can be quite large, reaching approximately 0.4 radians.

We perform linear regression for each time-slice and plot the first regression coefficient (the slope) in the lower panel. The regression coefficient has units of radians/MHz and can be converted to an effective phase residual by multiplying by half of the bandwidth. The effective phase residual scale is indicated on the right axis, and reflects the scale of the residuals in the upper panel. From the effective phase residual we can see that this residual changes slowly over the observation, on the scale of hours. This suggests that it is related to the arc-like artefacts.

The regression coefficient is insensitive to phase residuals originating from noise and therefore is a good measure of the systematic. **Table 4.3** summarises the effective phase residuals for each observation in terms of percentiles of the effective phase residuals. We observe that the distribution of effective phase residuals is heavily tailed. While half of the effective phase residuals are below approximately 0.07 radians, 20% of effective phase residuals are above approximately 0.15 radians and 10% are above approximately 0.20 radians.

Table 4.3: Effective phase residuals per observation.

Obs. ID	Effective phase residual (rad)		
	50-%ile	80-%ile	90-%ile
667218	0.06	0.13	0.20
667204	0.06	0.13	0.19
664480	0.06	0.13	0.19
664320	0.07	0.15	0.23
659948	0.06	0.14	0.21
659554	0.06	0.13	0.20
342938	0.07	0.14	0.19
340794	0.07	0.16	0.23
299961	0.06	0.12	0.18
294287	0.06	0.15	0.23
281008	0.05	0.11	0.16
274099	0.07	0.16	0.25

The phase residuals must originate from one or more DD systematics, since we have directionally referenced the Jones scalars. This is why we infer doubly differential total electron content instead of differential total electron content. **Figure 4.8** shows an example of the effective phase residual plotted over a field of view. We observe that the effective phase residuals are clearly correlated over direction.

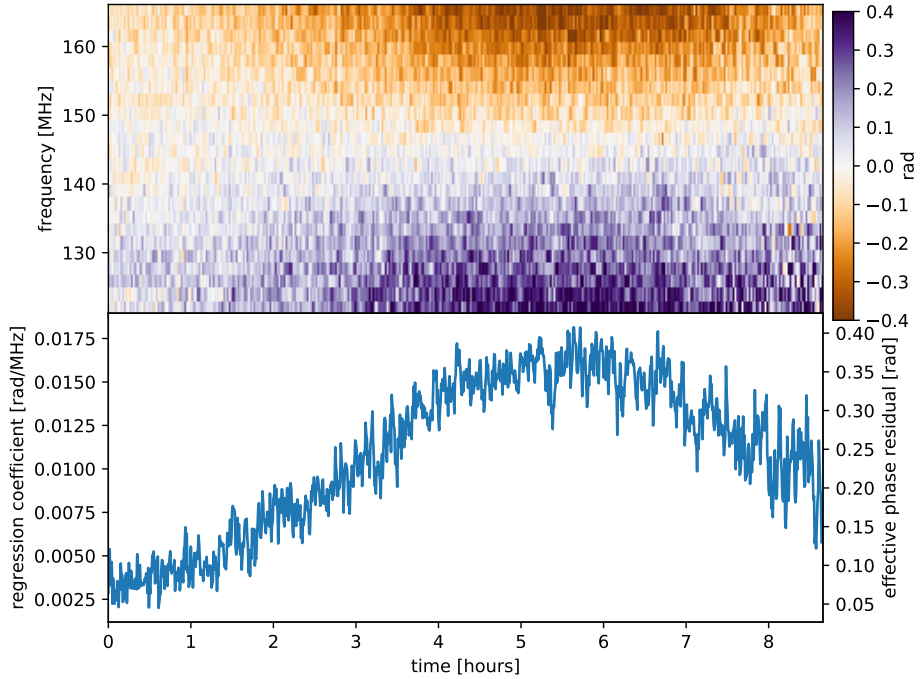


Figure 4.7: Example of unmodelled systematics in the phase residuals. The upper panel shows the phase residuals for a single antenna and direction. The lower panel shows the first regression coefficient (slope) of the phase residuals and the corresponding effective phase residual on the right axis.

Given that there is a correlation over direction and that the effect changes on the time scale of hours there are only a few known candidates for this systematic. The most logical cause is that the beam model is inaccurate. The beam of a LOFAR station is quite complicated since the tiles of the array exhibit electromagnetic coupling, which are not perfectly modelled. Since the beam shape is a function of phase tracking centre, an inaccurate beam model would introduce an effect that changes slowly over the course of an observation.

We propose that the phase residuals can be heuristically modelled by the addition of a constant-in-frequency term in the Jones phase model which changes slowly in time. This proposal is based on the fact that both LoTSS calibration and step 2 of our method produce images without arc-like patterns and both include a constant-in-frequency term in the phase model. Albert et al. [2020b] show how to perform variational Bayesian inference with a hidden Markov model [HMM; Rabiner and Juang, 1986] with any arbitrary linear phase model. In step 3 of the method we use a single linear term, i.e. DDTEC. Here we consider the effects of introducing a constant-in-frequency term to the linear phase model. In particular, we consider the bias-variance trade-off of introducing an additional degree of freedom.

We simulate 100 independent Jones scalar sequences using a two-component linear phase model with DDTEC plus a constant-in-frequency term. We use the HMM to simulate

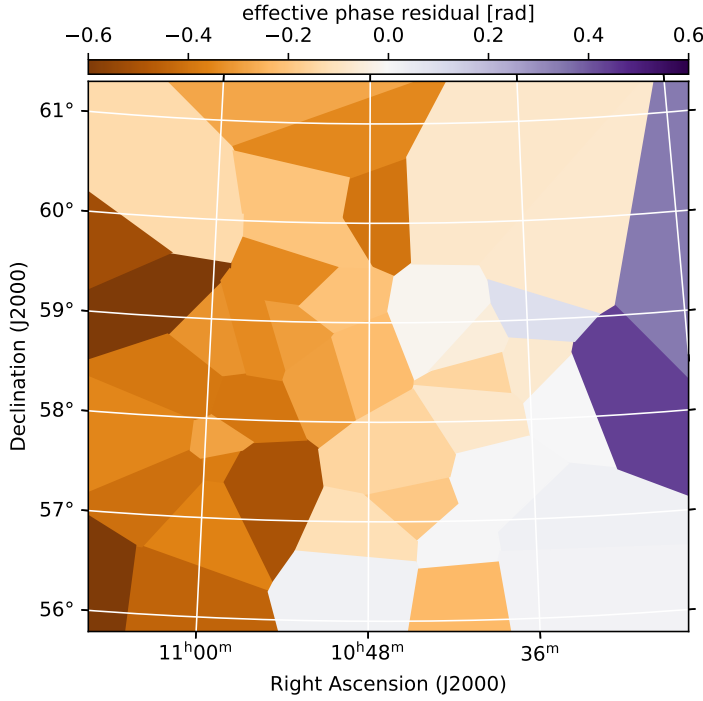


Figure 4.8: Example of the effective phase residual from observation 659948 exhibiting correlation among directions.

the Jones scalars, drawing from the empirical HMM parameter distribution for DDTEC, i.e. the variance of the Gaussian steps and observational uncertainty. We choose a Gaussian step variance of $(0.04 \text{ rad})^2$ for the constant-in-frequency term, which was selected to produce effective phase residuals that resemble the lower panel of **Figure 4.7**.

Given these simulated Jones scalars we perform HMM variational inference with two different linear phase models: 1) a DDTEC-only linear phase model, and 2) a DDTEC plus constant-in-frequency linear phase model. The first model is exactly the same one used in step 3, and the second model is the exact same one that the Jones scalars are simulated with. We then compute the residual between the posterior mean DDTEC and the ground truth.

Figure 4.9 plots the DDTEC residual for both linear phase models as a function of the ground truth constant-in-frequency term. We also compute the effective phase residual for the simulated data and find that there is a scaling relation of 0.157 times the ground truth constant-in-frequency term. The effective phase residual scale is shown on the top axis. We observe that the single component linear phase model (DDTEC-only) results in a residual that depends linearly on the ground truth constant-in-frequency term, with a slope of approximately $-16.7 \text{ mTECU rad}^{-1}$. The variance of the scatter is very small, $(0.75 \text{ mTECU})^2$. We observe that the two component linear phase model results in residuals with no significant dependence on the ground truth constant-in-frequency term. The variance of the scatter is an order of magnitude larger than the single component model, $(15 \text{ mTECU})^2$.

The first model results in high-bias low-variance residuals, while the second model results in low-bias high-variance residuals. The high-bias of the first model follows because the missing component in the phase model shifts the global optimum away from the ground truth, whereas the high-variance of the second model follows because we have introduced twice as many degrees of freedom given the same amount of information in the data. The high-bias of the first model implies that for antennas with phase residuals such as the one in **Figure 4.8**, that the inferred DDTEC is offset by as much as 60 mTECU, and this offset changes very slowly over the course of an observation. We suspect that this gives rise to the asymmetric arc-like patterns. Furthermore, while the effective phase residuals might seem deceptively small, $\ll 1$ rad, the effect on the inferred DDTEC is quite large.

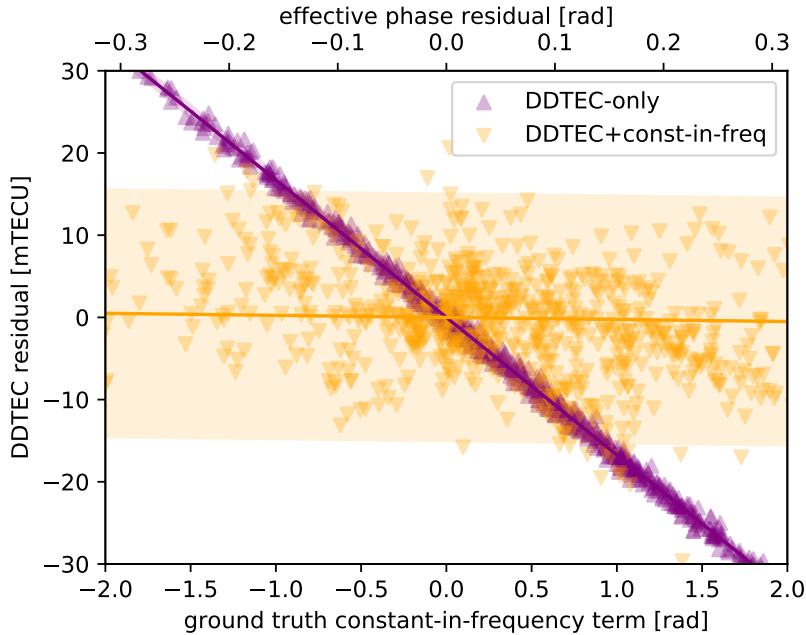


Figure 4.9: Deviations of the inferred DDTEC from the ground truth for two linear phase models. The single component model (purple) is the DDTEC-only model, and the two component model (orange) is the DDTEC and constant-in-frequency model. The shaded regions are 95% confidence regions.

The high-variance of the second model implies that we cannot simply add the constant-in-frequency term to the HMM in step 3, since the resulting DDTEC would be useless for screen-based modelling. We propose an alternative method of debiasing based on the effective phase residual and a strong prior on the smoothness. We use the scaling relation between the effective phase residual and the ground truth constant-in-frequency term to get a noisy estimate of the constant-in-frequency term. We then apply a median filter over a time window of 2.5 hours with reflecting boundaries to produce a slowly varying smooth estimate of the constant-in-frequency term. We then subtract this estimate of the constant-in-frequency term and resolve again with the DDTEC-only model. This method requires running

the HMM inference twice, once to get the effective phase residuals, and a second time once the phases have been debiased. Since this systematic is DD, we also interpolate the constant-in-frequency term to the screen directions. We apply nearest-neighbour interpolation as a first-order approximation.

Figure 4.10 shows a comparison of inter-calibrator sources significant arc-like artefact before and after debiasing. We observe the arc-like artefacts are suppressed, however a low-level diffuse artefact appeared. This confirms that the arc-like artefacts were indeed due to these phase residuals. The fact that the arc-like pattern disappears, despite the resulting diffuse artefacts suggests that our general approach to debiasing is correct, however there is still work to be done to perfect the method. For example, good outlier flagging proved critical for the DDTEC before inferring the screen, therefore perhaps outlier detection and flagging with these constant-in-frequency terms using spatial information would improve the results. Also, the timescale of 2.5 hours may be too long. Given that the resulting diffuse artefact is large scale, they likely stem from the treatment on short baselines, i.e. central antennae.

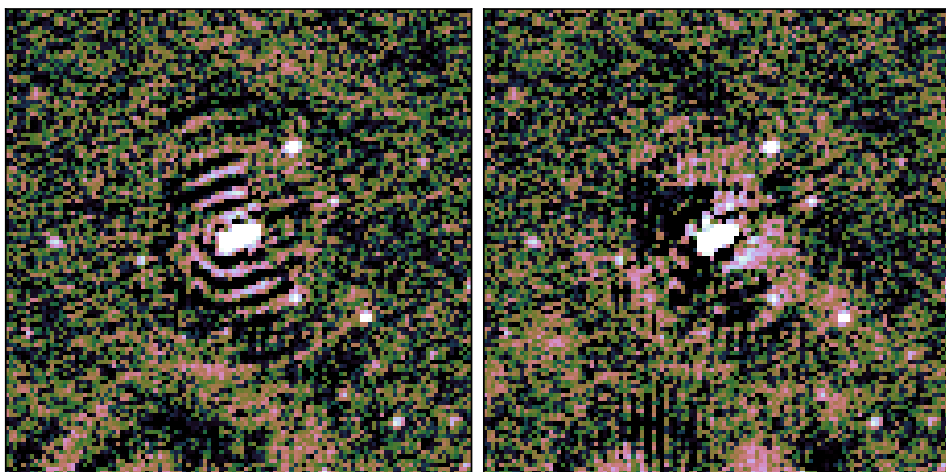
4.5 Information in DDTEC screens

An interesting product of this method is the rich view of the ionosphere's behaviour. **Figure 4.11** shows a random selection of DDTEC screens across the Lockman Hole data set, organised according to similarity of features. The posterior uncertainty in these DDTEC screens is $\lesssim 1$ mTECU. The screens were randomly selected from antennae at least 1 km from the reference antenna. Because they are randomly sampled they provide a cross section of the common types of ionosphere varieties. We see that features ranges from simple gradients across the field of view, to wave-like patterns, to higher-order aperiodic features. This supports the fact that there are many distinct behaviours in the ionosphere [e.g. Mevius et al., 2016, Jordan et al., 2017]. This is one of the most detailed views of the ionosphere ever made in terms of angular (sub-arcminute) and temporal (sub-minute) resolution, as well as precision of the DDTEC measurements (sub-mTECU). This zoo of DDTEC screens also allows us to explore the limits of our screen-based method. In particular, it enables us to explore in the future how sparse the calibrators can be selected such that we can still recover these screens.

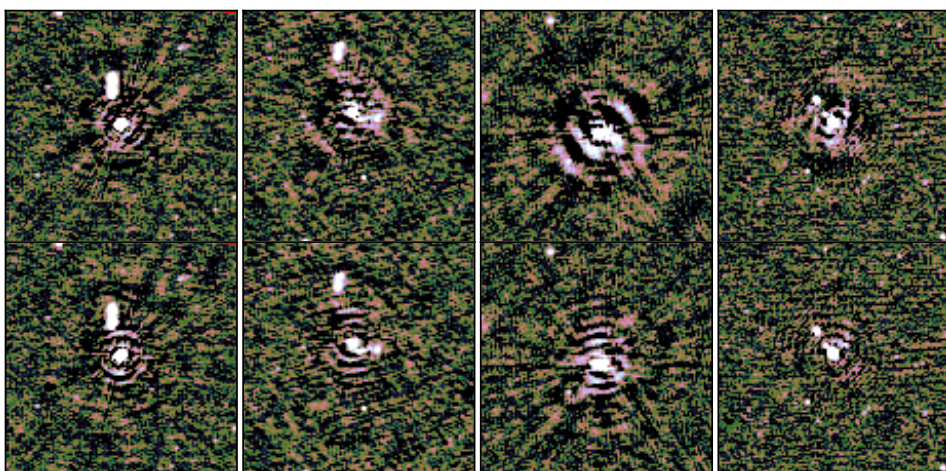
One of the tantalising future applications of this method is DD calibration of LOFAR-LBA. While it is beyond the scope of this paper to give an in-depth prediction for how feasible this will be, we can already draw conclusions from the variety of structures seen in **Figure 4.11**. Specifically, we note that roughly 20% of the DDTEC screens have wave-like patterns and another 20% have rough small-scale structure. Such prevalence of small-scale structure suggests that LOFAR-LBA, with typically less than ten bright calibrators with good quality Jones scalars per field of view², may struggle to perform DD calibration. This suggests DD calibration of LOFAR-LBA may benefit from the planned LOFAR expansion which would allow dual HBA-LBA observations, where the DDTEC screens from HBA can be applied to LBA.

As observed in Albert et al. [2020a], the full tomographic application is able to perform inference on sub-calibrator spacing. This suggests that a valuable future direction would be to make the full tomographic method computationally feasible for application to LOFAR-LBA.

²Private communications with W. Williams.



(a) Left: before applying debiasing. Right: After applying debiasing. Source comes from observation 342938.



(b) Cut-outs of inter-calibrator sources. Bottom row: before applying debiasing. Top row: After applying debiasing. Sources come from observations (left to right): 281008, 340794, 299961, 342938.

Figure 4.10: Comparison of inter-calibrator sources with significant arc-like patterns before and after applying our proposed debiasing correction.

On top of this, we would also like to understand what further improvements to the screen model would be beneficial at better handling some of the more difficult ionospheres, e.g. non-stationary FED covariances, or curvature of Earth.

Since the ionospheric model has undergone several approximations [see Albert et al., 2020b] for computational efficiency, we no longer perform tomography. Therefore, it is not possible to learn about most of the ionosphere's physical parameters. In principle, it is still possible to constrain the ratio of irregularity correlation scale to height of the ionosphere. However, since we have performed a FED hypothesis marginalisation it is not straight forward to interpret these parameters. Furthermore, as shown in Section 4.4, the modelling systematic can result in DDTEC that differ from the ground truth as much as 60 mTECU. It would therefore be premature to interpret the DDTEC screens for ionospheric science.

4.6 Conclusion

We have tested the probabilistic screen-based method of Albert et al. [2020b] on 12 observations (100 hours) of Lockman Hole LOFAR HBA data, which covers a wide range of ionospheric conditions, and compared the results to the calibration procedure of LoTSS-DR2. We find that 10/12 images showed significant improvement beyond the method of LoTSS-DR2, while 2/10 observations showed as-good or only slightly better improvements. We suggest that these moderate improvements are due to unresolved small scale structure in the ionosphere, or non-stationary structure. This implies that the method robust to most ionospheric conditions.

We have quantified the improvements from screen-based modelling in terms of equivalent integration-time gain around bright sources, which is the extra relative observation time that would be needed using the LoTSS-DR2 calibration method to reduce the scattering artefacts to the level that the screen achieves. We find that when the free electron density is high, due to increased Sun activity, measured in terms of Sun-spots, the equivalent integration-time gain is on average two. That is, the method achieves scattering effects equivalent to an observation twice as long. We also jointly imaged all observations, and compared this the to deep LoTSS-DR2 image. We find that similar improvements also scale to deep images, though modelling systematics build up and increase the noise around bright sources.

We discovered a modelling systematic that results in asymmetric arc-like artefacts around inter-calibrator sources. The origin of the phase component is not yet known, though we suggest it could be related to beam model inaccuracies. We propose that it can be accounted for as an additional DD constant-in-frequency phase component in the Jones phase model. Using a combination of real and simulated data we show that including a constant-in-frequency term in the HMM results in high-variance DDTEC inference unsuitable for screen inference. Therefore, an alternative method was proposed to account for the bias. We propose a simple filtering method, based on the slowly varying nature of the systematic, to remove it from the data. We show that this simple procedure removes the arc-like patterns, but introduces low-level diffuse artefacts. This suggest the general approach to accounting for the systematic is correct, but more work is needed to perfect it.

Once this systematic is robustly accounted for, the results of this paper suggest our screen-based method is feasible for reprocessing the entire LoTSS archive. Future improvements to the screen-based method, that would improve the robustness and also make it applicable to

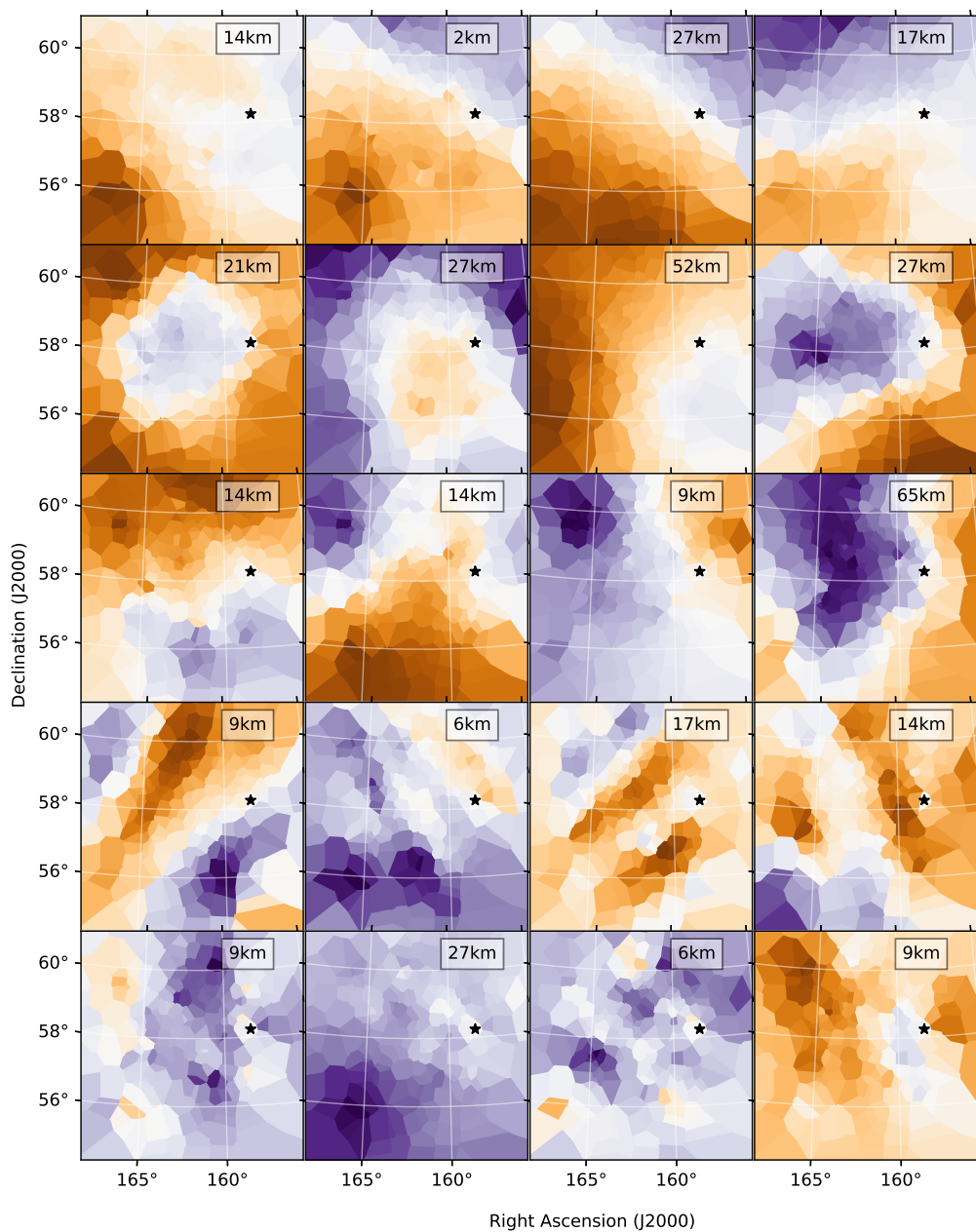
Figure 4.11: *Caption continued on next page.*

Figure 4.11: A zoo of DDTEC screens. Each panel is a DDTEC screen with normalised scale. The posterior uncertainty in these DDTEC screens is $\lesssim 1$ mTECU. The screens are randomly taken throughout the Lockman Hole data set from antennae at least 1 km from the reference antenna. They have been organised according to similar features ranging from simple gradients, to wave-like structures, to rough nearly uncorrelated structures. The black star is the reference direction. The distance from the reference antenna is shown in the box.

LOFAR-LBA, would be to relax the thin-layer approximation so that the model becomes fully (non-diffractive) tomographic as originally proposed in Albert et al. [2020a].

Bibliography

- Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dan Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. URL <https://www.tensorflow.org/>. Software available from tensorflow.org.
- J. G. Albert, M. S. S. L. Oei, R. J. van Weeren, H. T. Intema, and H. J. A. Röttgering. A probabilistic approach to direction-dependent ionospheric calibration. *A&A*, 633:A77, Jan 2020a. doi: 10.1051/0004-6361/201935668.
- J. G. Albert, R. J. van Weeren, H. T. Intema, and H. J. A. Röttgering. Probabilistic direction-dependent ionospheric calibration for LOFAR-HBA. *A&A*, 635:A147, March 2020b. doi: 10.1051/0004-6361/201937424.
- V. A. Ambartsumian. On the Question of the Nature of Radiosources. In *Proceedings of the Fifth Meeting on Cosmogony. Trudy Pjatogo soveshanija po voprosam kosmogonii. 9-12 marta 1955 g. Radioastronomija*, pages 413–416, January 1956.
- Miguel A. Aragón-Calvo, Rien van de Weygaert, and Bernard J. T. Jones. Multiscale phenomenology of the cosmic web. *MNRAS*, 408(4):2163–2187, November 2010. doi: 10.1111/j.1365-2966.2010.17263.x.
- B. S. Arora, J. Morgan, S. M. Ord, S. J. Tingay, M. Bell, J. R. Callingham, K. S. Dwarakanath, B. Q. For, P. Hancock, L. Hindson, N. Hurley-Walker, M. Johnston-Hollitt, A. D. Kapińska, E. Lenc, B. McKinley, A. R. Offringa, P. Procopio, L. Staveley-Smith, R. B. Wayth, C. Wu, and Q. Zheng. Ionospheric Modelling using GPS to Calibrate the MWA. II: Regional Ionospheric Modelling using GPS and GLONASS to Estimate Ionospheric Gradients. *PASA*, 33:e031, Jul 2016. doi: 10.1017/pasa.2016.22.
- J. A. Ball, D. Cesarsky, A. K. Dupree, L. Goldberg, and A. E. Lilley. Detection and Identification of Recombination Lines from an H i Region. *ApJ*, 162:L25, October 1970. doi: 10.1086/180616.

-
- J. A. Barcelo, A. Ruiz, and L. Vega. Weighted estimates for the helmholtz equation and some applications. *Journal of Functional Analysis*, 150(2):356 – 382, 1997. ISSN 0022-1236. doi: <https://doi.org/10.1006/jfan.1997.3131>. URL <http://www.sciencedirect.com/science/article/pii/S0022123697931311>.
- Mitchell C. Begelman, Roger D. Blandford, and Martin J. Rees. Theory of extragalactic radio sources. *Reviews of Modern Physics*, 56(2):255–351, April 1984. doi: 10.1103/RevModPhys.56.255.
- B. Bhattacharyya, S. Cooper, M. Malenta, J. Roy, J. Chengalur, M. Keith, S. Kudale, M. McLaughlin, S. M. Ransom, P. S. Ray, and B. W. Stappers. The GMRT High Resolution Southern Sky Survey for Pulsars and Transients. I. Survey Description and Initial Discoveries. *ApJ*, 817(2):130, February 2016. doi: 10.3847/0004-637X/817/2/130.
- D. Bilitza and B. W. Reinisch. International Reference Ionosphere 2007: Improvements and new parameters. *Advances in Space Research*, 42:599–609, August 2008. doi: 10.1016/j.asr.2007.07.048.
- J. Blackledge. *Digital Image processing*. Horwood Publishing, 2005. ISBN 1-898563-49-7.
- Roger Blandford and David Eichler. Particle acceleration at astrophysical shocks: A theory of cosmic ray origin. *Phys. Rep.*, 154(1):1–75, October 1987. doi: 10.1016/0370-1573(87)90134-7.
- A. Bonafede, L. Feretti, G. Giovannini, F. Govoni, M. Murgia, G. B. Taylor, H. Ebeling, S. Allen, G. Gentile, and Y. Pihlström. Revealing the magnetic field in a distant galaxy cluster: discovery of the complex radio emission from MACS J0717.5 +3745. *A&A*, 503:707–720, September 2009. doi: 10.1051/0004-6361/200912520.
- A. Bonafede, L. Feretti, M. Murgia, F. Govoni, G. Giovannini, D. Dallacasa, K. Dolag, and G. B. Taylor. The Coma cluster magnetic field from Faraday rotation measures. *A&A*, 513:A30, April 2010. doi: 10.1051/0004-6361/200913696.
- A. Bonafede, M. Brüggen, R. van Weeren, F. Vazza, G. Giovannini, H. Ebeling, A. C. Edge, M. Hoeft, and U. Klein. Discovery of radio haloes and double relics in distant MACS galaxy clusters: clues to the efficiency of particle acceleration. *MNRAS*, 426:40–56, October 2012. doi: 10.1111/j.1365-2966.2012.21570.x.
- Etienne Bonnassieux, Cyril Tasse, Oleg Smirnov, and Philippe Zarka. The variance of radio interferometric calibration solutions. Quality-based weighting schemes. *A&A*, 615:A66, Jul 2018. doi: 10.1051/0004-6361/201732190.
- Max Born and Emil Wolf. *Principles of Optics*. 1999.
- A. Botteon, F. Gastaldello, G. Brunetti, and R. Kale. A $M \gtrsim 3$ shock in ‘El Gordo’ cluster and the origin of the radio relic. *MNRAS*, 463(2):1534–1542, December 2016. doi: 10.1093/mnras/stw2089.
- K. L. Bouman, M. D. Johnson, D. Zoran, V. L. Fish, S. S. Doeleman, and W. T. Freeman. Computational Imaging for VLBI Image Reconstruction. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, March 2016.

-
- S. Y. Braude, B. A. Dubinskii, N. L. Kaidanovskii, N. S. Kardashev, M. M. Kobrin, A. D. Kuzmin, A. P. Molchanov, Y. N. Pariiskii, O. N. Rzhiga, A. E. Salomonovich, V. A. Samanian, I. S. Shklovskii, R. L. Sorochenko, V. S. Troitskii, and K. I. Kellermann. *A Brief History of Radio Astronomy in the USSR*. 2012.
- Robert Braun. *The Square Kilometer Array Interferometer*, volume 208 of *Astrophysics and Space Science Library*, page 167. 1996. doi: 10.1007/978-94-009-1734-7_9.
- T. K. Breus. Istoriya prioritetov sinkhrotronnoj kontseptsii v astronomii %t Historical problems of the priority questions of the synchrotron concept in astrophysics. *Istoriko-Astronomicheskie Issledovaniya*, 26:88–97, January 2001.
- D. Briggs. *High Fidelity Deconvolution of Moderately Resolved Sources*. PhD thesis, New Mexico Institute of Mining Technology, 1995. URL <http://www.aoc.nrao.edu/dissertations/dbriggs>.
- M. Brocklehurst and M. J. Seaton. On the interpretation of radio recombination line observations. *MNRAS*, 157:179, January 1972. doi: 10.1093/mnras/157.2.179.
- M. Brüggén and F. Vazza. Analytical model for cluster radio relics. *MNRAS*, 493(2):2306–2317, April 2020. doi: 10.1093/mnras/staa418.
- G. Brunetti. The challenge of turbulent acceleration of relativistic particles in the intra-cluster medium. *Plasma Physics and Controlled Fusion*, 58(1):014011, January 2016. doi: 10.1088/0741-3335/58/1/014011.
- G. Brunetti and T. W. Jones. Cosmic Rays in Galaxy Clusters and Their Nonthermal Emission. *International Journal of Modern Physics D*, 23:1430007–98, March 2014. doi: 10.1142/S0218271814300079.
- G. Brunetti, S. Giacintucci, R. Cassano, W. Lane, D. Dallacasa, T. Venturi, N. E. Kassim, G. Setti, W. D. Cotton, and M. Markevitch. A low-frequency radio halo associated with a cluster of galaxies. *Nature*, 455:944–947, October 2008. doi: 10.1038/nature07379.
- G. R. Burbidge. On Synchrotron Radiation from Messier 87. *ApJ*, 124:416, September 1956. doi: 10.1086/146237.
- R. Byrd, P. Lu, J. Nocedal, and C. Zhu. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, 16(5):1190–1208, 1995. doi: 10.1137/0916069.
- Xiaohao Cai, Marcelo Pereyra, and Jason D. McEwen. Uncertainty quantification for radio interferometric imaging: II. MAP estimation. *MNRAS*, 480(3):4170–4182, November 2018. doi: 10.1093/mnras/sty2015.
- Xiaohao Cai, Luke Pratley, and Jason D. McEwen. Online radio interferometric imaging: assimilating and discarding visibilities on arrival. *MNRAS*, 485(4):4559–4572, June 2019. doi: 10.1093/mnras/stz704.

-
- J. Callingham, H. Vedantham, T. Shimwell, B. J. Pope, and LoTSS Team. Stellar systems at low radio frequencies: The discovery of radio exoplanets. In *American Astronomical Society Meeting Abstracts*, American Astronomical Society Meeting Abstracts, page 342.02, January 2020.
- P J Cargill. Fundamentals of plasma physics. *Plasma Physics and Controlled Fusion*, 49(2): 197, 2007. URL <http://stacks.iop.org/0741-3335/49/i=2/a=B01>.
- R. Cassano, G. Brunetti, and G. Setti. Statistics of giant radio haloes from electron reacceleration models. *MNRAS*, 369:1577–1595, July 2006. doi: 10.1111/j.1365-2966.2006.10423.x.
- R. Cassano, G. Brunetti, H. J. A. Röttgering, and M. Brüggen. Unveiling radio halos in galaxy clusters in the LOFAR era. *A&A*, 509:A68, January 2010a. doi: 10.1051/0004-6361/200913063.
- R. Cassano, S. Ettori, S. Giacintucci, G. Brunetti, M. Markevitch, T. Venturi, and M. Gitti. On the Connection Between Giant Radio Halos and Cluster Mergers. *ApJ*, 721:L82–L85, October 2010b. doi: 10.1088/2041-8205/721/2/L82.
- Lawrie Challis and Fred Sheard. The Green of Green Functions. *Physics Today*, 56(12):41–46, December 2003. doi: 10.1063/1.1650227.
- A. S. Cohen and H. J. A. Röttgering. Probing Fine-Scale Ionospheric Structure with the Very Large Array Radio Telescope. *The Astronomical Journal*, 138(2):439–447, Aug 2009. doi: 10.1088/0004-6256/138/2/439.
- M. H. Cohen. Introduction to very-long-baseline interferometry. *IEEE Proceedings*, 61: 1192–1197, 1973. doi: 10.1109/PROC.1973.9244.
- J. J. Condon, W. D. Cotton, E. W. Greisen, Q. F. Yin, R. A. Perley, G. B. Taylor, and J. J. Broderick. The NRAO VLA Sky Survey. *AJ*, 115(5):1693–1716, May 1998. doi: 10.1086/300337.
- T. J. Cornwell and R. A. Perley. Radio-interferometric imaging of very large fields. The problem of non-coplanar arrays. *A&A*, 261:353–364, July 1992.
- T. J. Cornwell, K. Golap, and S. Bhatnagar. Wide Field Imaging: Fourier and Fresnel. In N. Kassim, M. Perez, W. Junor, and P. Henning, editors, *From Clark Lake to the Long Wavelength Array: Bill Erickson's Radio Science*, volume 345 of *Astronomical Society of the Pacific Conference Series*, page 350, Dec 2005.
- T. J. Cornwell, K. Golap, and S. Bhatnagar. The Noncoplanar Baselines Effect in Radio Interferometry: The W-Projection Algorithm. *IEEE Journal of Selected Topics in Signal Processing*, 2:647–657, November 2008. doi: 10.1109/JSTSP.2008.2005290.
- William D. Cotton, James J. Condon, Richard A. Perley, Namir Kassim, Joseph Lazio, Aaron Cohen, Wendy Lane, and William C. Erickson. Beyond the isoplanatic patch in the VLA Low-frequency Sky Survey. In Jr. Oschmann, Jacobus M., editor, *Proc. SPIE*, volume 5489 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 180–189, Oct 2004. doi: 10.1117/12.551298.

-
- F. de Gasperin, T. J. Dijkema, A. Drabent, M. Mevius, D. Rafferty, R. van Weeren, M. Brüggen, J. R. Callingham, K. L. Emig, G. Heald, H. T. Intema, L. K. Morabito, A. R. Offringa, R. Oonk, E. Orrù, H. Röttgering, J. Sabater, T. Shimwell, A. Shulevski, and W. Williams. Systematic effects in LOFAR data: A unified calibration strategy. *A&A*, 622:A5, Feb 2019. doi: 10.1051/0004-6361/201833867.
- Francesco de Gasperin, Huib T. Intema, Timothy W. Shimwell, Gianfranco Brunetti, Marcus Brüggen, Torsten A. Enßlin, Reinout J. van Weeren, Annalisa Bonafede, and Huub J. A. Röttgering. Gentle reenergization of electrons in merging galaxy clusters. *Science Advances*, 3(10):e1701634, October 2017. doi: 10.1126/sciadv.1701634.
- Francesco de Gasperin, M. Mevius, D. Rafferty, Huib Intema, and Richard Fallows. The effect of the ionosphere on ultra-low frequency radio-interferometric observations. *A&A*, 615, 04 2018. doi: 10.1051/0004-6361/201833012.
- A. De Luca and S. Molendi. The 2-8 keV cosmic X-ray background spectrum as observed with XMM-Newton. *A&A*, 419:837–848, June 2004. doi: 10.1051/0004-6361:20034421.
- Thomas A. Dean, Sumeetpal S. Singh, Ajay Jasra, and Gareth W. Peters. Parameter estimation for hidden markov models with intractable likelihoods. *Scandinavian Journal of Statistics*, 41(4):970–987, 2014. doi: 10.1111/sjos.12077. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/sjos.12077>.
- P. E. Dewdney, P. J. Hall, R. T. Schilizzi, and T. J. L. W. Lazio. The Square Kilometre Array. *IEEE Proceedings*, 97(8):1482–1496, August 2009. doi: 10.1109/JPROC.2009.2021005.
- J. Donnert, K. Dolag, G. Brunetti, and R. Cassano. Rise and fall of radio haloes in simulated merging galaxy clusters. *MNRAS*, 429:3564–3569, March 2013. doi: 10.1093/mnras/sts628.
- J. R. Driscoll and D. M. Healy. Computing fourier transforms and convolutions on the 2-sphere. *Advances in Applied Mathematics*, 15(2):202 – 250, 1994. ISSN 0196-8858. doi: <https://doi.org/10.1006/aama.1994.1008>. URL <http://www.sciencedirect.com/science/article/pii/S0196885884710086>.
- L. O. Drury. An introduction to the theory of diffusive shock acceleration of energetic particles in tenuous plasmas. *Reports on Progress in Physics*, 46:973–1027, August 1983. doi: 10.1088/0034-4885/46/8/002.
- Michael W. Eastwood, Marin M. Anderson, Ryan M. Monroe, Gregg Hallinan, Benjamin R. Barsdell, Stephen A. Bourke, M. A. Clark, Steven W. Ellingson, Jayce Dowell, Hugh Garsden, Lincoln J. Greenhill, Jacob M. Hartman, Jonathon Kocz, T. Joseph W. Lazio, Danny C. Price, Frank K. Schinzel, Gregory B. Taylor, Harish K. Vedantham, Yuankun Wang, and David P. Woody. The Radio Sky at Meter Wavelengths: m-mode Analysis Imaging with the OVRO-LWA. *AJ*, 156(1):32, July 2018. doi: 10.3847/1538-3881/aac721.
- K. L. Emig, P. Salas, F. de Gasperin, J. B. R. Oonk, M. C. Toribio, H. J. A. Röttgering, and A. G. G. M. Tielens. The first detection of radio recombination lines at cosmological distances. *A&A*, 622:A7, February 2019. doi: 10.1051/0004-6361/201834052.

-
- T. A. Enßlin and Gopal-Krishna. Reviving fossil radio plasma in clusters of galaxies by adiabatic compression in environmental shock waves. *A&A*, 366:26–34, January 2001. doi: 10.1051/0004-6361:20000198.
- T. A. Ensslin, P. L. Biermann, U. Klein, and S. Kohle. Cluster radio relics as a tracer of shock waves of the large-scale structure formation. *A&A*, 332:395–409, April 1998.
- G. Evans, J. Blackledge, and P. Yardley. *Analytic Methods for Partial Differential Equations: Green's Functions*. Springer Undergraduate Mathematics Series, 1999.
- H. I. Ewen and E. M. Purcell. Observation of a Line in the Galactic Radio Spectrum: Radiation from Galactic Hydrogen at 1,420 Mc./sec. *Nature*, 168(4270):356, September 1951. doi: 10.1038/168356a0.
- P. B. Fellgett and E. H. Linfoot. On the Assessment of Optical Images. *Philosophical Transactions of the Royal Society of London Series A*, 247(931):369–407, Feb 1955. doi: 10.1098/rsta.1955.0001.
- W. Freudling, M. Romaniello, D. M. Bramich, P. Ballester, V. Forchi, C. E. García-Dabó, S. Moehler, and M. J. Neeser. Automated data reduction workflows for astronomy. The ESO Reflex environment. *A&A*, 559:A96, November 2013. doi: 10.1051/0004-6361/201322494.
- D. Gabor. Microscopy by Reconstructed Wave-Fronts. *Proceedings of the Royal Society of London Series A*, 197(1051):454–487, July 1949. doi: 10.1098/rspa.1949.0075.
- Jorge Garcia-Sucerquia, Román Castañeda, and Francisco F Medina. Fresnel-Fraunhofer diffraction and spatial coherence. *Optics Communications*, 205(4-6):239–245, May 2002. doi: 10.1016/S0030-4018(02)01378-0.
- John V. Gillespie, Dina A. Zinnes, G. S. Tahim, Philip A. Schrodtt, and R. Michael Rubison. An optimal control model of arms races. *The American Political Science Review*, 71(1):226–244, 1977. ISSN 00030554, 15375943. URL <http://www.jstor.org/stable/1956964>.
- F. Govoni, M. Murgia, L. Feretti, G. Giovannini, D. Dallacasa, and G. B. Taylor. A2255: The first detection of filamentary polarized emission in a radio halo. *A&A*, 430:L5–L8, January 2005. doi: 10.1051/0004-6361:200400113.
- T. L. Grobler, C. D. Nunhokee, O. M. Smirnov, A. J. van Zyl, and A. G. de Bruyn. Calibration artefacts in radio interferometry - I. Ghost sources in Westerbork Synthesis Radio Telescope data. *MNRAS*, 439(4):4030–4047, April 2014. doi: 10.1093/mnras/stu268.
- Y. Gupta, B. Ajithkumar, H. S. Kale, S. Nayak, S. Sabhapathy, S. Sureshkumar, R. V. Swami, J. N. Chengalur, S. K. Ghosh, C. H. Ishwara-Chandra, B. C. Joshi, N. Kanekar, D. V. Lal, and S. Roy. The upgraded GMRT: opening new windows on the radio Universe. *Current Science*, 113(4):707–714, August 2017.
- J. P. Hamaker, J. D. Bregman, and R. J. Sault. Understanding radio polarimetry. I. Mathematical foundations. *A&AS*, 117:137–147, May 1996.
- I. Harrison, S. Camera, J. Zuntz, and M. L. Brown. SKA weak lensing - I. Cosmological forecasts and the power of radio-optical cross-correlations. *MNRAS*, 463:3674–3685, December 2016. doi: 10.1093/mnras/stw2082.

-
- A. H. Hassan, C. J. Fluke, D. G. Barnes, and V. A. Kilborn. Tera-scale astronomical data analysis and visualization. *MNRAS*, 429(3):2442–2455, March 2013. doi: 10.1093/mnras/sts513.
- Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep Residual Learning for Image Recognition. *arXiv e-prints*, art. arXiv:1512.03385, December 2015.
- S. E. Healey, R. W. Romani, G. B. Taylor, E. M. Sadler, R. Ricci, T. Murphy, J. S. Ulvestad, and J. N. Winn. CRATES: An All-Sky Survey of Flat-Spectrum Radio Sources. *ApJS*, 171:61–71, July 2007. doi: 10.1086/513742.
- Carl Heiles, Phil Perillat, Michael Nolan, Duncan Lorimer, Ramesh Bhat, Tapasi Ghosh, Ellen Howell, Murray Lewis, Karen O’Neil, Chris Salter, and Snezana Stanimirovic. All-Stokes Parameterization of the Main Beam and First Sidelobe for the Arecibo Radio Telescope. *PASP*, 113(788):1247–1273, October 2001. doi: 10.1086/323290.
- J. N. Hendriks, C. Jidling, A. Wills, and T. B. Schön. Evaluating the squared-exponential covariance function in Gaussian processes with integral observations. *arXiv e-prints*, art. arXiv:1812.07319, Dec 2018.
- James Hensman, Nicolo Fusi, and Neil D. Lawrence. Gaussian Processes for Big Data. *arXiv e-prints*, art. arXiv:1309.6835, Sep 2013.
- S. L. G. Hess and P. Zarka. Modeling the radio signature of the orbital parameters, rotation, and magnetic field of exoplanets. *A&A*, 531:A29, July 2011. doi: 10.1051/0004-6361/201116510.
- A. Hewish. The Diffraction of Radio Waves in Passing through a Phase-Changing Ionosphere. *Proceedings of the Royal Society of London Series A*, 209(1096):81–96, Oct 1951. doi: 10.1098/rspa.1951.0189.
- A. Hewish. The Diffraction of Galactic Radio Waves as a Method of Investigating the Irregular Structure of the Ionosphere. *Proceedings of the Royal Society of London Series A*, 214(1119):494–514, Oct 1952. doi: 10.1098/rspa.1952.0185.
- A. Hewish, S. J. Bell, J. D. H. Pilkington, P. F. Scott, and R. A. Collins. Observation of a Rapidly Pulsating Radio Source. *Nature*, 217(5130):709–713, February 1968. doi: 10.1038/217709a0.
- J. A. Högbom. Aperture Synthesis with a Non-Regular Distribution of Interferometer Baselines. *A&AS*, 15:417, June 1974.
- C. Hollitt, M. Johnston-Hollitt, S. Dehghan, M. Frean, and T. Butler-Yeoman. *An Overview of the SKA Science Analysis Pipeline*, volume 512 of *Astronomical Society of the Pacific Conference Series*, page 367. 2017.
- N. Hurley-Walker, J. R. Callingham, P. J. Hancock, T. M. O. Franzen, L. Hindson, A. D. Kapińska, J. Morgan, A. R. Offringa, R. B. Wayth, C. Wu, Q. Zheng, T. Murphy, M. E. Bell, K. S. Dwarakanath, B. For, B. M. Gaensler, M. Johnston-Hollitt, E. Lenc, P. Procopio, L. Staveley-Smith, R. Ekers, J. D. Bowman, F. Briggs, R. J. Cappallo, A. A. Deshpande, L. Greenhill, B. J. Hazelton, D. L. Kaplan, C. J. Lonsdale, S. R. McWhirter, D. A. Mitchell, M. F. Morales, E. Morgan, D. Oberoi, S. M. Ord, T. Prabu, N. U. Shankar, K. S. Srivani, R. Subrahmanyam, S. J. Tingay, R. L. Webster, A. Williams, and C. L. Williams. GaLactic and Extragalactic All-sky Murchison Widefield Array (GLEAM) survey - I. A low-frequency extragalactic catalogue. *MNRAS*, 464:1146–1167, January 2017. doi: 10.1093/mnras/stw2337.

-
- I. H. Hutchinson. *Principles of Plasma Diagnostics*. 2005.
- H. T. Intema. SPAM: Source Peeling and Atmospheric Modeling. Astrophysics Source Code Library, August 2014.
- H. T. Intema, S. van der Tol, W. D. Cotton, A. S. Cohen, I. M. van Bemmelen, and H. J. A. Röttgering. Ionospheric calibration of low frequency radio interferometric observations using the peeling scheme. I. Method description and first results. *A&A*, 501:1185–1205, July 2009. doi: 10.1051/0004-6361/200811094.
- H. T. Intema, P. Jagannathan, K. P. Mooley, and D. A. Frail. The GMRT 150 MHz All-sky Radio Survey: First Alternative Data Release TGSS ADR1. *ArXiv e-prints*, March 2016.
- E. A. Isaev and P. A. Tarasov. Transmission of large amounts of scientific data using laser technology. In *Journal of Physics Conference Series*, volume 740 of *Journal of Physics Conference Series*, page 012015, August 2016. doi: 10.1088/1742-6596/740/1/012015.
- W. J. Jaffe and G. C. Perola. Dynamical Models of Tailed Radio Sources in Clusters of Galaxies. *A&A*, 26:423, August 1973.
- Karl G. Jansky. Radio Waves from Outside the Solar System. *Nature*, 132(3323):66, July 1933. doi: 10.1038/132066a0.
- Wojciech Jarosz. *Efficient Monte Carlo Methods for Light Transport in Scattering Media*. PhD thesis, UC San Diego, September 2008.
- H. Jeffreys. On certain approximate solutions of lineae differential equations of the second order*. *Proceedings of the London Mathematical Society*, s2-23(1):428–436, 1925. ISSN 1460-244X. doi: 10.1112/plms/s2-23.1.428.
- Carl Jidling, Johannes Hendriks, Niklas Wahlström, Alexander Gregg, Thomas B. SchÄun, Christopher Wensrich, and Adrian Wills. Probabilistic modelling and reconstruction of strain. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 436:141 – 155, 2018. ISSN 0168-583X. doi: <https://doi.org/10.1016/j.nimb.2018.08.051>. URL <http://www.sciencedirect.com/science/article/pii/S0168583X1830524X>.
- R. C. Jones. 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, July 1941.
- C. H. Jordan, S. Murray, C. M. Trott, R. B. Wayth, D. A. Mitchell, M. Rahimi, B. Pindor, P. Procopio, and J. Morgan. Characterization of the ionosphere above the Murchison Radio Observatory using the Murchison Widefield Array. *MNRAS*, 471(4):3974–3987, Nov 2017. doi: 10.1093/mnras/stx1797.
- H. Kang and D. Ryu. Diffusive Shock Acceleration at Cosmological Shock Waves. *ApJ*, 764:95, February 2013. doi: 10.1088/0004-637X/764/1/95.
- N. S. Kardashev. On the Possibility of Detection of Allowed Lines of Atomic Hydrogen in the Radio-Frequency Spectrum. *Soviet Ast.*, 3:813, October 1959.

-
- N. S. Kardashev. Nonstationarity of Spectra of Young Sources of Nonthermal Radio Emission. *Soviet Ast.*, 6:317, December 1962.
- S. Kazemi, S. Yatawatta, S. Zaroubi, P. Lampropoulos, A. G. de Bruyn, L. V. E. Koopmans, and J. Noordam. Radio interferometric calibration using the SAGE algorithm. *MNRAS*, 414: 1656–1666, June 2011. doi: 10.1111/j.1365-2966.2011.18506.x.
- J.B. Keller, R. Bellman, and American Mathematical Society. *Stochastic Equations and Wave Propagation in Random Media*. Proceedings of symposia in applied mathematics. American Mathematical Society, 1964. URL <https://books.google.nl/books?id=evQW0gAACAAJ>.
- K. O. Kiepenheuer. On the relations between ionosphere, sunspots and solar corona. *MNRAS*, 106:515, Jan 1946. doi: 10.1093/mnras/106.6.515.
- Diederik P. Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization. *arXiv e-prints*, art. arXiv:1412.6980, December 2014.
- M. G. Kivelson and C. T. Russell. *Introduction to Space Physics*. Cambridge University Press, April 1995. ISBN 0521451043.
- A. N. Kolmogorov. 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, July 1991. doi: 10.1098/rspa.1991.0075.
- Andrey N. Kolmogorov. *Foundations of the theory of probability*. Chelsea Pub Co, 2 english edition, 1956.
- Andrey N. Kolmogorov. *Foundations of the Theory of Probability*. Chelsea Pub Co, 2 edition, June 1960. URL [http://www.clrc.rhul.ac.uk/resources/fop/Theory%20of%20Probability%20\(small\).pdf](http://www.clrc.rhul.ac.uk/resources/fop/Theory%20of%20Probability%20(small).pdf).
- S. S. Komissarov and A. G. Gubanov. Relic radio galaxies: evolution of synchrotron spectrum. *A&A*, 285:27–43, May 1994.
- L. V. E. Koopmans. Ionospheric Power-spectrum Tomography in Radio Interferometry. *ApJ*, 718:963–971, August 2010. doi: 10.1088/0004-637X/718/2/963.
- Benjamin Krüger, Thomas Brenner, and Alwin Kienle. Solution of the inhomogeneous Maxwell’s equations using a Born series. *Optics Express*, 25(21):25165, October 2017. doi: 10.1364/OE.25.025165.
- Mark Kuiack, Folkert Huizinga, Gijs Molenaar, Peeyush Prasad, Antonia Rowlinson, and Ralph A. M. J. Wijers. AARTFAAC flux density calibration and Northern hemisphere catalogue at 60 MHz. *Monthly Notices of the Royal Astronomical Society*, 482(2):2502–2514, Jan 2019. doi: 10.1093/mnras/sty2810.
- M. J. Kurtz and D. J. Mink. RVSAO 2.0: Digital Redshifts and Radial Velocities. *PASP*, 110: 934–977, August 1998. doi: 10.1086/316207.

-
- R. R. Lindner, A. J. Baker, J. P. Hughes, N. Battaglia, N. Gupta, K. Knowles, T. A. Marriage, F. Menanteau, K. Moodley, E. D. Reese, and R. Srianand. The Radio Relics and Halo of El Gordo, a Massive $z = 0.870$ Cluster Merger. *ApJ*, 786:49, May 2014. doi: 10.1088/0004-637X/786/1/49.
- Robert R. Lindner, Carlos Vera-Ciro, Claire E. Murray, Snežana Stanimirović, Brian Babler, Carl Heiles, Patrick Hennebelle, W. M. Goss, and John Dickey. Autonomous Gaussian Decomposition. *AJ*, 149(4):138, April 2015. doi: 10.1088/0004-6256/149/4/138.
- S. T. Loi, T. Murphy, I. H. Cairns, F. W. Menk, C. L. Waters, P. J. Erickson, C. M. Trott, N. Hurley-Walker, J. Morgan, E. Lenc, A. R. Offringa, M. E. Bell, R. D. Ekers, B. M. Gaensler, C. J. Lonsdale, L. Feng, P. J. Hancock, D. L. Kaplan, G. Bernardi, J. D. Bowman, F. Briggs, R. J. Cappallo, A. A. Deshpande, L. J. Greenhill, B. J. Hazelton, M. Johnston-Hollitt, S. R. McWhirter, D. A. Mitchell, M. F. Morales, E. Morgan, D. Oberoi, S. M. Ord, T. Prabu, N. U. Shankar, K. S. Srivani, R. Subrahmanyam, S. J. Tingay, R. B. Wayth, R. L. Webster, A. Williams, and C. L. Williams. Real-time imaging of density ducts between the plasmasphere and ionosphere. *Geophys. Res. Lett.*, 42:3707–3714, May 2015. doi: 10.1002/2015GL063699.
- R. J. Lyon, B. W. Stappers, L. Levin, M. B. Mickaliger, and A. Scaife. A processing pipeline for high volume pulsar candidate data streams. *Astronomy and Computing*, 28:100291, July 2019. doi: 10.1016/j.ascom.2019.100291.
- M. A. Malkov and L. O’C Drury. Nonlinear theory of diffusive acceleration of particles by shock waves. *Reports on Progress in Physics*, 64:429–481, April 2001. doi: 10.1088/0034-4885/64/4/201.
- S. Mandal, H. T. Intema, T. W. Shimwell, R. J. van Weeren, A. Botteon, H. J. A. Röttgering, D. N. Hoang, G. Brunetti, F. de Gasperin, S. Giacintucci, H. Hoekstra, A. Stroe, M. Brüggen, R. Cassano, A. Shulevski, A. Drabent, and D. Rafferty. Ultra-steep spectrum emission in the merging galaxy cluster Abell 1914. *A&A*, 622:A22, February 2019. doi: 10.1051/0004-6361/201833992.
- S. Mandal, H. T. Intema, R. J. van Weeren, T. W. Shimwell, A. Botteon, G. Brunetti, F. de Gasperin, M. Brüggen, G. Di Gennaro, R. Kraft, H. J. A. Röttgering, M. Hardcastle, and C. Tasse. Revived fossil plasma sources in galaxy clusters. *A&A*, 634:A4, February 2020. doi: 10.1051/0004-6361/201936560.
- L. Mandel and E. Wolf. Coherence Properties of Optical Fields. *Reviews of Modern Physics*, 37(2):231–287, April 1965. doi: 10.1103/RevModPhys.37.231.
- Alexander G. de G. Matthews, Mark van der Wilk, Tom Nickson, Keisuke. Fujii, Alexis Boukouvalas, Pablo León-Villagrà, Zoubin Ghahramani, and James Hensman. GPflow: A Gaussian process library using TensorFlow. *Journal of Machine Learning Research*, 18(40):1–6, apr 2017. URL <http://jmlr.org/papers/v18/16-537.html>.
- B. J. Maughan, C. Jones, W. Forman, and L. Van Speybroeck. Images, Structural Properties, and Metal Abundances of Galaxy Clusters Observed with Chandra ACIS-I at $0.1 \ll z \ll 1.3$. *ApJS*, 174:117–135, January 2008. doi: 10.1086/521225.

-
- F. Mernier, J. de Plaa, L. Lovisari, C. Pinto, Y.-Y. Zhang, J. S. Kaastra, N. Werner, and A. Simionescu. Abundance and temperature distributions in the hot intra-cluster gas of Abell 4059. *A&A*, 575:A37, March 2015. doi: 10.1051/0004-6361/201425282.
- F. G. Mertens, M. Mevius, L. V. E. Koopmans, A. R. Offringa, G. Mellema, S. Zaroubi, M. A. Brentjens, H. Gan, B. K. Gehlot, V. N. Pandey, A. M. Sardarabadi, H. K. Vedantham, S. Yatawatta, K. M. B. Asad, B. Ciardi, E. Chapman, S. Gazagnes, R. Ghara, A. Ghosh, S. K. Giri, I. T. Iliev, V. Jelić, R. Kooistra, R. Mondal, J. Schaye, and M. B. Silva. Improved upper limits on the 21 cm signal power spectrum of neutral hydrogen at $z \approx 9.1$ from LOFAR. *MNRAS*, 493(2): 1662–1685, April 2020. doi: 10.1093/mnras/staa327.
- Christopher A. Metzler, Martin White, and Chris Loken. The Effect of the Cosmic Web on Cluster Weak Lensing Mass Estimates. *ApJ*, 547(2):560–573, February 2001. doi: 10.1086/318406.
- M. Mevius, S. van der Tol, V. N. Pandey, H. K. Vedantham, M. A. Brentjens, A. G. de Bruyn, F. B. Abdalla, K. M. B. Asad, J. D. Bregman, W. N. Brouw, S. Bus, E. Chapman, B. Ciardi, E. R. Fernandez, A. Ghosh, G. Harker, I. T. Iliev, V. Jelić, S. Kazemi, L. V. E. Koopmans, J. E. Noordam, A. R. Offringa, A. H. Patil, R. J. van Weeren, S. Wijnholds, S. Yatawatta, and S. Zaroubi. Probing ionospheric structures using the LOFAR radio telescope. *Radio Science*, 51:927–941, Jul 2016. doi: 10.1002/2016RS006028.
- G. K. Miley, H. van der Laan, and K. J. Wellington. Recent Westerbork Observations of Head-Tail Galaxies. In J. R. Shakeshaft, editor, *The Formation and Dynamics of Galaxies*, volume 58 of *IAU Symposium*, page 109, 1974.
- C. A. Muller and J. H. Oort. Observation of a Line in the Galactic Radio Spectrum: The Interstellar Hydrogen Line at 1,420 Mc./sec., and an Estimate of Galactic Rotation. *Nature*, 168(4270):357–358, September 1951. doi: 10.1038/168357a0.
- Jan E. Noordam. *LOFAR calibration challenges*, volume 5489 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 817–825. 2004. doi: 10.1117/12.544262.
- S. E. Nuza, M. Hoeft, R. J. van Weeren, S. Gottlöber, and G. Yepes. How many radio relics await discovery? *MNRAS*, 420:2006–2019, March 2012. doi: 10.1111/j.1365-2966.2011.20118.x.
- A. R. Offringa. Compression of interferometric radio-astronomical data. *A&A*, 595:A99, November 2016. doi: 10.1051/0004-6361/201629565.
- Gerwin Osnabrugge, Saroch Leedumrongwatthanakun, and Ivo M. Vellekoop. A convergent Born series for solving the inhomogeneous Helmholtz equation in arbitrarily large media. *Journal of Computational Physics*, 322:113–124, October 2016. doi: 10.1016/j.jcp.2016.06.034.
- Donald E. Osterbrock. *Astrophysics of gaseous nebulae and active galactic nuclei*. 1989.
- A. G. Pacholczyk. *Radio astrophysics. Nonthermal processes in galactic and extragalactic sources*. 1970.

-
- Gregory Paciga, Joshua G. Albert, Kevin Bandura, Tzu-Ching Chang, Yashwant Gupta, Christopher Hirata, Julia Odegova, Ue-Li Pen, Jeffrey B. Peterson, Jayanta Roy, J. Richard Shaw, Kris Sigurdson, and Tabitha Voytek. A simulation-calibrated limit on the H I power spectrum from the GMRT Epoch of Reionization experiment. *MNRAS*, 433(1):639–647, July 2013. doi: 10.1093/mnras/stt753.
- N. G. Parke. Optical algebra. *Journal of Mathematics and Physics*, 28(1-4):131–139, 1949. doi: 10.1002/sapm1949281131. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/sapm1949281131>.
- A. H. Patil, S. Yatawatta, L. V. E. Koopmans, A. G. de Bruyn, M. A. Brentjens, S. Zaroubi, K. M. B. Asad, M. Hatef, V. Jelić, M. Mevius, A. R. Offringa, V. N. Pandey, H. Vedantham, F. B. Abdalla, W. N. Brouw, E. Chapman, B. Ciardi, B. K. Gehlot, A. Ghosh, G. Harker, I. T. Iliev, K. Kakiichi, S. Majumdar, G. Mellema, M. B. Silva, J. Schaye, D. Vrbanec, and S. J. Wijnholds. Upper Limits on the 21 cm Epoch of Reionization Power Spectrum from One Night with LOFAR. *ApJ*, 838:65, March 2017. doi: 10.3847/1538-4357/aa63e7.
- T. J. Pearson and A. C. S. Readhead. Image Formation by Self-Calibration in Radio Astronomy. *ARA&A*, 22:97–130, January 1984. doi: 10.1146/annurev.aa.22.090184.000525.
- Richard A. Perley, Frederic R. Schwab, and Alan H. Bridle. *Synthesis imaging in radio astronomy: a collection of lectures from the third NRAO synthesis imaging summer school*, volume 6. 1989.
- H. J. Pesch and M. Plail. The maximum principle of optimal control: A history of ingenious ideas and missed opportunities. *Control and Cybernetics*, Vol. 38, no 4A:973–995, 2009.
- G. J. Phillips. Measurement of winds in the ionosphere. *Journal of Atmospheric and Terrestrial Physics*, 2(3):141–154, Jan 1952. doi: 10.1016/0021-9169(52)90059-7.
- Planck Collaboration, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. Balbi, A. J. Banday, R. B. Barreiro, M. Bartelmann, and et al. Planck early results. IX. XMM-Newton follow-up for validation of Planck cluster candidates. *A&A*, 536:A9, December 2011. doi: 10.1051/0004-6361/201116460.
- Planck Collaboration, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, H. Aussel, C. Baccigalupi, and et al. Planck 2013 results. XXIX. The Planck catalogue of Sunyaev-Zeldovich sources. *A&A*, 571:A29, November 2014. doi: 10.1051/0004-6361/201321523.
- G. B. Poole, M. A. Fardal, A. Babul, I. G. McCarthy, T. Quinn, and J. Wadsley. The impact of mergers on relaxed X-ray clusters - I. Dynamical evolution and emergent transient structures. *MNRAS*, 373:881–905, December 2006. doi: 10.1111/j.1365-2966.2006.10916.x.
- P. Prasad, F. Huizinga, E. Kooistra, D. van der Schuur, A. Gunst, J. Romein, M. Kuiack, G. Molenaar, A. Rowlinson, J. D. Swinbank, and R. A. M. J. Wijers. The AARTFAAC All-Sky Monitor: System Design and Implementation. *Journal of Astronomical Instrumentation*, 5:1641008, December 2016. doi: 10.1142/S2251171716410087.

-
- L. H. Quiroga-Nuñez, H. T. Intema, J. R. Callingham, J. Villadsen, H. J. van Langevelde, P. Jagannathan, T. W. Shimwell, and E. P. Boven. Differences in radio emission from similar M dwarfs in the binary system Ross 867-8. *A&A*, 633:A130, January 2020. doi: 10.1051/0004-6361/201936491.
- L. Rabiner and B. Juang. An introduction to hidden markov models. *IEEE ASSP Magazine*, 3 (1):4–16, Jan 1986. ISSN 1558-1284. doi: 10.1109/MASSP.1986.1165342.
- J. Radon. Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten. *Berichte über die Verhandlungen der Königlich-Sächsischen Gesellschaft der Wissenschaften zu Leipzig*, 69:262–277, 1917.
- Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press, 2006a. ISBN 026218253X.
- Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press, 2006b. ISBN 026218253X.
- J. A. Ratcliffe. Some Aspects of Diffraction Theory and their Application to the Ionosphere. *Reports on Progress in Physics*, 19(1):188–267, Jan 1956. doi: 10.1088/0034-4885/19/1/306.
- H. E. Rauch. Solution to the linear smoothing problem. *Automatic Control, IEEE Transactions on*, AC-8:371 – 372, 11 1963. doi: 10.1109/TAC.1963.1105600.
- T. H. Reiprich and H. Böhringer. The Mass Function of an X-Ray Flux-limited Sample of Galaxy Clusters. *ApJ*, 567:716–740, March 2002. doi: 10.1086/338753.
- C. J. Riseley, A. M. M. Scaife, M. W. Wise, and A. O. Clarke. Diffuse radio emission in MACS J0025.4-1222: the effect of a major merger on bulk separation of ICM components. *A&A*, 597:A96, January 2017. doi: 10.1051/0004-6361/201629530.
- K. Roettiger, J. M. Stone, and J. O. Burns. Magnetic Field Evolution in Merging Clusters of Galaxies. *ApJ*, 518:594–602, June 1999. doi: 10.1086/307298.
- George B. Rybicki and Alan P. Lightman. *Radiative Processes in Astrophysics*. 1986.
- M. Ryle and D. D. Vonberg. Solar Radiation on 175 Mc./s. *Nature*, 158(4010):339–340, September 1946. doi: 10.1038/158339b0.
- J. S. Santos, P. Rosati, P. Tozzi, H. Böhringer, S. Ettori, and A. Bignamini. Searching for cool core clusters at high redshift. *A&A*, 483:35–47, May 2008. doi: 10.1051/0004-6361/20078815.
- A. Saxena, M. Marinello, R. A. Overzier, P. N. Best, H. J. A. Röttgering, K. J. Duncan, I. Prandoni, L. Pentericci, M. Magliocchetti, D. Paris, F. Cusano, F. Marchi, H. T. Intema, and GK Miley. Discovery of a radio galaxy at $z = 5.72$. *MNRAS*, 480(2):2733–2742, October 2018. doi: 10.1093/mnras/sty1996.
- Jelena A. Schmalz, Gerd Schmalz, Timur E. Gureyev, and Konstantin M. Pavlov. On the derivation of the Green's function for the Helmholtz equation using generalized functions. *American Journal of Physics*, 78(2):181–186, February 2010. doi: 10.1119/1.3253655.

-
- T. W. Shimwell, H. J. A. Röttgering, P. N. Best, W. L. Williams, T. J. Dijkema, F. de Gasperin, M. J. Hardcastle, G. H. Heald, D. N. Hoang, A. Horneffer, H. Intema, E. K. Mahony, S. Mandal, A. P. Mechev, L. Morabito, J. B. R. Oonk, D. Rafferty, E. Retana-Montenegro, J. Sabater, C. Tasse, R. J. van Weeren, M. Brüggen, G. Brunetti, K. T. Chyży, J. E. Conway, M. Haverkorn, N. Jackson, M. J. Jarvis, J. P. McKean, G. K. Miley, R. Morganti, G. J. White, M. W. Wise, I. M. van Bemmell, R. Beck, M. Brienza, A. Bonafede, G. Calistro Rivera, R. Cassano, A. O. Clarke, D. Cseh, A. Deller, A. Drabent, W. van Driel, D. Engels, H. Falcke, C. Ferrari, S. Fröhlich, M. A. Garrett, J. J. Harwood, V. Heesen, M. Hoeft, C. Horellou, F. P. Israel, A. D. Kapińska, M. Kunert-Bajraszewska, D. J. McKay, N. R. Mohan, E. Orrú, R. F. Pizzo, I. Prandoni, D. J. Schwarz, A. Shulevski, M. Sipior, D. J. B. Smith, S. S. Sridhar, M. Steinmetz, A. Stroe, E. Varenus, P. P. van der Werf, J. A. Zensus, and J. T. L. Zwart. The LOFAR Two-metre Sky Survey. I. Survey description and preliminary data release. *A&A*, 598:A104, Feb 2017. doi: 10.1051/0004-6361/201629313.
- T. W. Shimwell, C. Tasse, M. J. Hardcastle, A. P. Mechev, W. L. Williams, P. N. Best, H. J. A. Röttgering, J. R. Callingham, T. J. Dijkema, F. de Gasperin, D. N. Hoang, B. Hugo, M. Mirmont, J. B. R. Oonk, I. Prandoni, D. Rafferty, J. Sabater, O. Smirnov, R. J. van Weeren, G. J. White, M. Atemkeng, L. Bester, E. Bonnassieux, M. Brüggen, G. Brunetti, K. T. Chyży, R. Cochrane, J. E. Conway, J. H. Croston, A. Danezi, K. Duncan, M. Haverkorn, G. H. Heald, M. Iacobelli, H. T. Intema, N. Jackson, M. Jamrozy, M. J. Jarvis, R. Lakhoo, M. Mevius, G. K. Miley, L. Morabito, R. Morganti, D. Nisbet, E. Orrú, S. Perkins, R. F. Pizzo, C. Schrijvers, D. J. B. Smith, R. Vermeulen, M. W. Wise, L. Alegre, D. J. Bacon, I. M. van Bemmell, R. J. Beswick, A. Bonafede, A. Botteon, S. Bourke, M. Brienza, G. Calistro Rivera, R. Cassano, A. O. Clarke, C. J. Conselice, R. J. Dettmar, A. Drabent, C. Dumba, K. L. Emig, T. A. Enßlin, C. Ferrari, M. A. Garrett, R. T. Génova-Santos, A. Goyal, G. Gürkan, C. Hale, J. J. Harwood, V. Heesen, M. Hoeft, C. Horellou, C. Jackson, G. Kokotanekov, R. Kondapally, M. Kunert-Bajraszewska, V. Mahatma, E. K. Mahony, S. Mandal, J. P. McKean, A. Merloni, B. Mingo, A. Miskolczi, S. Mooney, B. Nikiel-Wroczyński, S. P. O’Sullivan, J. Quinn, W. Reich, C. Roskowiński, A. Rowlinson, F. Savini, A. Saxena, D. J. Schwarz, A. Shulevski, S. S. Sridhar, H. R. Stacey, S. Urquhart, M. H. D. van der Wiel, E. Varenus, B. Webster, and A. Wilber. The LOFAR Two-metre Sky Survey. II. First data release. *A&A*, 622:A1, February 2019. doi: 10.1051/0004-6361/201833559.
- I. S. Shklovsky. The Problem of Cosmic Radioemission. ‘Problema Kosmičeskogo Radioizlučeniâ’. *AZh*, 30:15–36, January 1953.
- R. H. Shumway and D. S. Stoffer. An approach to time series smoothing and forecasting using the em algorithm. *Journal of Time Series Analysis*, 3(4):253–264, 1982. doi: 10.1111/j.1467-9892.1982.tb00349.x. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9892.1982.tb00349.x>.
- C. Sifón, F. Menanteau, J. P. Hughes, M. Carrasco, and L. F. Barrientos. Strong lensing analysis of PLCK G004.5-19.5, a Planck-discovered cluster hosting a radio relic at $z = 0.52$. *A&A*, 562:A43, February 2014. doi: 10.1051/0004-6361/201321638.
- O. M. Smirnov and C. Tasse. Radio interferometric gain calibration as a complex optimization problem. *MNRAS*, 449(3):2668–2684, May 2015. doi: 10.1093/mnras/stv418.

-
- George C. Southworth. Early History of Radio Astronomy. *The Scientific Monthly*, 82(2):55–66, February 1956.
- T. A. T. Spoelstra. The influence of ionospheric refraction on radio astronomy interferometry. *A&A*, 120:313–321, April 1983.
- Woodruff T. Sullivan. The history of radio telescopes, 1945–1990. *Experimental Astronomy*, 25(1-3):107–124, August 2009. doi: 10.1007/s10686-009-9140-2.
- Terence Tao. Recent progress on the restriction conjecture. *arXiv Mathematics e-prints*, art. math/0311181, November 2003.
- C. Tasse. Nonlinear Kalman filters for calibration in radio interferometry. *Astronomy and Astrophysics*, 566:A127, Jun 2014a. doi: 10.1051/0004-6361/201423503.
- C. Tasse. Applying Wirtinger derivatives to the radio interferometry calibration problem. *arXiv e-prints*, October 2014b.
- C. Tasse, S. van der Tol, J. van Zwieten, G. van Diepen, and S. Bhatnagar. Applying full polarization A-Projection to very wide field of view instruments: An imager for LOFAR. *A&A*, 553:A105, May 2013. doi: 10.1051/0004-6361/201220882.
- C. Tasse, B. Hugo, M. Mirmont, O. Smirnov, M. Atemkeng, L. Bester, M. J. Hardcastle, R. Lakhoo, S. Perkins, and T. Shimwell. Faceting for direction-dependent spectral deconvolution. *A&A*, 611:A87, April 2018. doi: 10.1051/0004-6361/201731474.
- S. J. Tingay, R. Goeke, J. D. Bowman, D. Emrich, S. M. Ord, D. A. Mitchell, M. F. Morales, T. Boller, B. Crosse, R. B. Wayth, C. J. Lonsdale, S. Tremblay, D. Pallot, T. Colegate, A. Wicenc, N. Kudryavtseva, W. Arcus, D. Barnes, G. Bernardi, F. Briggs, S. Burns, J. D. Bunton, R. J. Cappallo, B. E. Corey, A. Deshpande, L. Desouza, B. M. Gaensler, L. J. Greenhill, P. J. Hall, B. J. Hazelton, D. Herne, J. N. Hewitt, M. Johnston-Hollitt, D. L. Kaplan, J. C. Kasper, B. B. Kincaid, R. Koenig, E. Kratzenberg, M. J. Lynch, B. McKinley, S. R. McWhirter, E. Morgan, D. Oberoi, J. Pathikulangara, T. Prabu, R. A. Remillard, A. E. E. Rogers, A. Rosh, J. E. Salah, R. J. Sault, N. Udaya-Shankar, F. Schlagenhauser, K. S. Srivani, J. Stevens, R. Subrahmanyam, M. Waterson, R. L. Webster, A. R. Whitney, A. Williams, C. L. Williams, and J. S. B. Wyithe. The Murchison Widefield Array: The Square Kilometre Array Precursor at Low Radio Frequencies. *PASA*, 30:e007, January 2013. doi: 10.1017/pasa.2012.007.
- R. van de Weygaert and W. Schaap. *The Cosmic Web: Geometric Analysis*, volume 665, pages 291–413. 2009. doi: 10.1007/978-3-540-44767-2_11.
- Rien van de Weygaert and Tjeerd S. van Albada. The Square Kilometer Array: new challenges for cosmology. *arXiv e-prints*, art. astro-ph/9602096, February 1996.
- S. van der Tol. *Bayesian estimation for ionospheric calibration in radio astronomy*. PhD thesis, TU Delft, 2009.
- M. P. van Haarlem, M. W. Wise, A. W. Gunst, G. Heald, J. P. McKean, J. W. T. Hessels, A. G. de Bruyn, R. Nijboer, J. Swinbank, R. Fallows, M. Brentjens, A. Nelles, R. Beck, H. Falcke, R. Fender, J. Hörandel, L. V. E. Koopmans, G. Mann, G. Miley, H. Röttgering, B. W. Stappers,

-
- R. A. M. J. Wijers, S. Zaroubi, M. van den Akker, A. Alexov, J. Anderson, K. Anderson, A. van Ardenne, M. Arts, A. Asgekar, I. M. Avruch, F. Batejat, L. Bähren, M. E. Bell, M. R. Bell, I. van Bemmell, P. Bennema, M. J. Bentum, G. Bernardi, P. Best, L. Bîrzan, A. Bonafede, A.-J. Boonstra, R. Braun, J. Bregman, F. Breitling, R. H. van de Brink, J. Broderick, P. C. Broekema, W. N. Brouw, M. Brüggen, H. R. Butcher, W. van Cappellen, B. Ciardi, T. Coenen, J. Conway, A. Coolen, A. Corstanje, S. Damstra, O. Davies, A. T. Deller, R.-J. Dettmar, G. van Diepen, K. Dijkstra, P. Donker, A. Doorduïn, J. Dromer, M. Drost, A. van Duin, J. Eislöffel, J. van Enst, C. Ferrari, W. Frieswijk, H. Gankema, M. A. Garrett, F. de Gasperin, M. Gerbers, E. de Geus, J.-M. Grießmeier, T. Grit, P. Gruppen, J. P. Hamaker, T. Hassall, M. Hoeft, H. A. Holties, A. Horneffer, A. van der Horst, A. van Houwelingen, A. Huijgen, M. Iacobelli, H. Intema, N. Jackson, V. Jelic, A. de Jong, E. Juette, D. Kant, A. Karastergiou, A. Koers, H. Kollen, V. I. Kondratiev, E. Kooistra, Y. Koopman, A. Koster, M. Kuniyoshi, M. Kramer, G. Kuper, P. Lambropoulos, C. Law, J. van Leeuwen, J. Lemaître, M. Loose, P. Maat, G. Macario, S. Markoff, J. Masters, R. A. McFadden, D. McKay-Bukowski, H. Meijering, H. Meulman, M. Mevius, E. Middelberg, R. Millenaar, J. C. A. Miller-Jones, R. N. Mohan, J. D. Mol, J. Morawietz, R. Morganti, D. D. Mulcahy, E. Mulder, H. Munk, L. Nieuwenhuis, R. van Nieuwpoort, J. E. Noordam, M. Norden, A. Noutsos, A. R. Offringa, H. Olofsson, A. Omar, E. Orrú, R. Overeem, H. Paas, M. Pandey-Pommier, V. N. Pandey, R. Pizzo, A. Polatidis, D. Rafferty, S. Rawlings, W. Reich, J.-P. de Reijer, J. Reitsma, G. A. Renting, P. Riemers, E. Rol, J. W. Romein, J. Roosjen, M. Ruiter, A. Scaife, K. van der Schaaf, B. Scheers, P. Schellart, A. Schoenmakers, G. Schoonderbeek, M. Serylak, A. Shulevski, J. Sluman, O. Smirnov, C. Sobey, H. Spreeuw, M. Steinmetz, C. G. M. Sterks, H.-J. Stiepel, K. Stuurwold, M. Tagger, Y. Tang, C. Tasse, I. Thomas, S. Thoudam, M. C. Toribio, B. van der Tol, O. Usov, M. van Veelen, A.-J. van der Veen, S. ter Veen, J. P. W. Verbiest, R. Vermeulen, N. Vermaas, C. Vocks, C. Vogt, M. de Vos, E. van der Wal, R. van Weeren, H. Weggemans, P. Weltevrede, S. White, S. J. Wijnholds, T. Wilhelmsson, O. Wucknitz, S. Yatawatta, P. Zarka, A. Zensus, and J. van Zwieten. LOFAR: The LOw-Frequency ARray. *A&A*, 556:A2, August 2013. doi: 10.1051/0004-6361/201220873.
- R. J. van Weeren, H. J. A. Röttgering, M. Brüggen, and A. Cohen. Diffuse radio emission in the merging cluster MACS J0717.5+3745: the discovery of the most powerful radio halo. *A&A*, 505:991–997, October 2009. doi: 10.1051/0004-6361/200912528.
- R. J. van Weeren, H. J. A. Röttgering, and M. Brüggen. Diffuse steep-spectrum sources from the 74 MHz VLSS survey. *A&A*, 527:A114, March 2011. doi: 10.1051/0004-6361/201015991.
- R. J. van Weeren, W. L. Williams, M. J. Hardcastle, T. W. Shimwell, D. A. Rafferty, J. Sabater, G. Heald, S. S. Sridhar, T. J. Dijkema, G. Brunetti, M. Brüggen, F. Andrade-Santos, G. A. Ogrean, H. J. A. Röttgering, W. A. Dawson, W. R. Forman, F. de Gasperin, C. Jones, G. K. Miley, L. Rudnick, C. L. Sarazin, A. Bonafede, P. N. Best, L. Bîrzan, R. Cassano, K. T. Chyży, J. H. Croston, T. Ensslin, C. Ferrari, M. Hoeft, C. Horellou, M. J. Jarvis, R. P. Kraft, M. Mevius, H. T. Intema, S. S. Murray, E. Orrú, R. Pizzo, A. Simionescu, A. Stroe, S. van der Tol, and G. J. White. LOFAR Facet Calibration. *ApJS*, 223:2, March 2016. doi: 10.3847/0067-0049/223/1/2.
- R. J. van Weeren, F. de Gasperin, H. Akamatsu, M. Brüggen, L. Feretti, H. Kang, A. Stroe, and F. Zandanel. Diffuse Radio Emission from Galaxy Clusters. *Space Sci. Rev.*, 215:16, February 2019. doi: 10.1007/s11214-019-0584-z.
- Reinout J. van Weeren, Huub J. A. Röttgering, Marcus Brüggen, and Matthias Hoeft. Particle

-
- Acceleration on Megaparsec Scales in a Merging Galaxy Cluster. *Science*, 330(6002):347, October 2010. doi: 10.1126/science.1194293.
- F. Vazza. How complex is the cosmic web? *MNRAS*, 491(4):5447–5463, February 2020. doi: 10.1093/mnras/stz3317.
- F. Vazza, G. Brunetti, C. Gheller, R. Brunino, and M. Brüggen. Massive and refined. II. The statistical properties of turbulent motions in massive galaxy clusters with high spatial resolution. *A&A*, 529:A17, May 2011. doi: 10.1051/0004-6361/201016015.
- F. Vazza, M. Brüggen, C. Gheller, and G. Brunetti. Modelling injection and feedback of cosmic rays in grid-based cosmological simulations: effects on cluster outskirts. *MNRAS*, 421: 3375–3398, April 2012. doi: 10.1111/j.1365-2966.2012.20562.x.
- H. K. Vedantham and L. V. E. Koopmans. Scintillation noise in widefield radio interferometry. *MNRAS*, 453(1):925–938, October 2015. doi: 10.1093/mnras/stv1594.
- H. K. Vedantham and L. V. E. Koopmans. Scintillation noise power spectrum and its impact on high-redshift 21-cm observations. *MNRAS*, 458(3):3099–3117, May 2016. doi: 10.1093/mnras/stw443.
- T. Venturi, S. Giacintucci, D. Dallacasa, R. Cassano, G. Brunetti, S. Bardelli, and G. Setti. GMRT radio halo survey in galaxy clusters at $z = 0.2\text{--}0.4$. II. The eBCS clusters and analysis of the complete sample. *A&A*, 484(2):327–340, June 2008. doi: 10.1051/0004-6361:200809622.
- T. Vernstrom, B. M. Gaensler, S. Brown, E. Lenc, and R. P. Norris. Low-frequency radio constraints on the synchrotron cosmic web. *MNRAS*, 467:4914–4936, Jun 2017. doi: 10.1093/mnras/stx424.
- Y. Weiss and W. T. Freeman. Correctness of belief propagation in gaussian graphical models of arbitrary topology. *Neural Computation*, 13(10):2173–2200, Oct 2001. ISSN 0899-7667. doi: 10.1162/089976601750541769.
- S. J. Wijnholds, A. G. Willis, and S. Salvini. Baseline-dependent averaging in radio interferometry. *MNRAS*, 476(2):2029–2039, May 2018. doi: 10.1093/mnras/sty360.
- C. H. Wilcox. Wave propagation in a random medium (lev a. chernov). *SIAM Review*, 4(1): 55–55, 1962. doi: 10.1137/1004017. URL <https://doi.org/10.1137/1004017>.
- W. L. Williams, G. Calistro Rivera, P. N. Best, M. J. Hardcastle, H. J. A. Röttgering, K. J. Duncan, F. de Gasperin, M. J. Jarvis, G. K. Miley, E. K. Mahony, L. K. Morabito, D. M. Nisbet, I. Prandoni, D. J. B. Smith, C. Tasse, and G. J. White. LOFAR-Boötes: properties of high- and low-excitation radio galaxies at $0.5 < z < 2.0$. *MNRAS*, 475(3):3429–3452, April 2018. doi: 10.1093/mnras/sty026.
- D. Wittor, M. Hoeft, F. Vazza, M. Brüggen, and P. Domínguez-Fernández. Polarization of radio relics in galaxy clusters. *MNRAS*, 490(3):3987–4006, December 2019. doi: 10.1093/mnras/stz2715.
- E. Wolf. Three-dimensional structure determination of semi-transparent objects from holographic data. *Optics Communications*, 1:153–156, September 1969. doi: 10.1016/0030-4018(69)90052-2.

- Lingfei Wu, Dashun Wang, and James A. Evans. Large teams develop and small teams disrupt science and technology. *Nature*, 566:1, Feb 2019. doi: 10.1038/s41586-019-0941-9.
- K. C. Yeh. Propagation of spherical waves through an ionosphere containing anisotropic irregularities. *Journal of Research of National Bureau of Standards*, 5:621–636, sep 1962. doi: 10.6028/jres.066D.062.
- K. C. Yeh and G. W. Swenson, Jr. The scintillation of radio signals from satellites. *J. Geophys. Res.*, 64:2281–2286, December 1959. doi: 10.1029/JZ064i012p02281.
- Fisher Yu and Vladlen Koltun. Multi-scale context aggregation by dilated convolutions. In *International Conference on Learning Representations (ICLR)*, May 2016.

Appendix A

Derivation of tomographic equivalence

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

$$\int P(\boldsymbol{\tau} | \mathbf{n}_e) P(\mathbf{n}_e | \boldsymbol{\tau}_{\text{obs}}) d\mathbf{n}_e = \int P(\mathbf{n}_e, \boldsymbol{\tau} | \boldsymbol{\tau}_{\text{obs}}) d\mathbf{n}_e. \quad (\text{A.1})$$

We note that we sometimes use the notation $\mathcal{N}[a | m_a, C_a]$ 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, $\Delta_* \mathbf{n}_e = \{\Delta_{\mathbf{x}}^{\hat{\mathbf{k}}} n_e | (\mathbf{x}, \hat{\mathbf{k}}) \in \mathcal{S}_*\}$, 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 $\mathbf{K} = \{K(\mathbf{x}, \mathbf{x}') | \mathbf{x}, \mathbf{x}' \in \mathcal{X}\}$. Using these matrix representation we have the following joint distribution,

$$P(\mathbf{n}_e, \boldsymbol{\tau}, \boldsymbol{\tau}_{\text{obs}}) = \mathcal{N} \begin{bmatrix} \bar{n}_e & \mathbf{K} & \mathbf{K} \Delta_*^T & \mathbf{K} \Delta^T \\ 0 & \Delta_* \mathbf{K} & \Delta_* \mathbf{K} \Delta_*^T & \Delta_* \mathbf{K} \Delta^T \\ 0 & \Delta \mathbf{K} & \Delta \mathbf{K} \Delta_*^T & \Delta \mathbf{K} \Delta^T + \sigma^2 \mathbf{I} \end{bmatrix}. \quad (\text{A.2})$$

Let us first work out the left-hand side (LHS) of Eq. A.1. Because $\boldsymbol{\tau} = \Delta_* \mathbf{n}_e$, and using standard Gaussian identities we have,

$$P(\boldsymbol{\tau} | \mathbf{n}_e) = \mathcal{N} \left[\underbrace{\Delta_* \mathbf{K} \mathbf{K}^{-1} (\mathbf{n}_e - \bar{n}_e)}_{\Delta_*(\mathbf{n}_e - \bar{n}_e)}, \underbrace{\Delta_* \mathbf{K} \Delta_* - \Delta_* \mathbf{K} \mathbf{K}^{-1} \mathbf{K} \Delta_*}_{\mathbf{0}} \right]. \quad (\text{A.3})$$

Similarly, the second distribution on the LHS is,

$$P(\mathbf{n}_e | \boldsymbol{\tau}_{\text{obs}}) = \mathcal{N} [\bar{n}_e + \mathbf{K} \Delta^T (\Delta \mathbf{K} \Delta^T + \sigma^2 \mathbf{I})^{-1} \boldsymbol{\tau}_{\text{obs}}, \mathbf{K} - \mathbf{K} \Delta^T (\Delta \mathbf{K} \Delta^T + \sigma^2 \mathbf{I})^{-1} \Delta \mathbf{K}]. \quad (\text{A.4})$$

We now can use standard Gaussian identities [e.g. Weiss and Freeman, 2001] to evaluate the

integral on the LHS,

$$\int P(\boldsymbol{\tau} | \mathbf{n}_e) P(\mathbf{n}_e | \boldsymbol{\tau}_{\text{obs}}) d\mathbf{n}_e = \mathcal{N}[\boldsymbol{\Delta}_* \mathbf{K} \boldsymbol{\Delta}^T (\boldsymbol{\Delta} \mathbf{K} \boldsymbol{\Delta}^T + \sigma^2 \mathbf{I})^{-1} \boldsymbol{\tau}_{\text{obs}}, \boldsymbol{\Delta}_* \mathbf{K} \boldsymbol{\Delta}_*^T - \boldsymbol{\Delta}_* \mathbf{K} \boldsymbol{\Delta}^T (\boldsymbol{\Delta} \mathbf{K} \boldsymbol{\Delta}^T + \sigma^2 \mathbf{I})^{-1} \boldsymbol{\Delta} \mathbf{K} \boldsymbol{\Delta}_*^T] \quad (\text{A.5})$$

In order to work out the right-hand side (RHS), we simply condition Eq. A.2 on $\boldsymbol{\tau}_{\text{obs}}$ and then marginalise $\mathbf{n}_e$ by selecting the corresponding sub-block of the Gaussian,

$$P(\mathbf{n}_e, \boldsymbol{\tau} | \boldsymbol{\tau}_{\text{obs}}) = \mathcal{N} \left[\begin{pmatrix} \bar{\mathbf{n}}_e \\ 0 \end{pmatrix} + \begin{pmatrix} \mathbf{K} \boldsymbol{\Delta}^T \\ \boldsymbol{\Delta}_*^T \mathbf{K} \boldsymbol{\Delta}^T \end{pmatrix} (\boldsymbol{\Delta} \mathbf{K} \boldsymbol{\Delta}^T + \sigma^2 \mathbf{I})^{-1} \boldsymbol{\tau}_{\text{obs}}, \begin{pmatrix} \bar{\mathbf{K}} & \boldsymbol{\Delta} \mathbf{K} \boldsymbol{\Delta}_*^T \\ \boldsymbol{\Delta}_* \mathbf{K} \boldsymbol{\Delta}^T & \boldsymbol{\Delta}_* \mathbf{K} \boldsymbol{\Delta}_*^T \end{pmatrix} - \begin{pmatrix} \mathbf{K} \boldsymbol{\Delta}^T \\ \boldsymbol{\Delta}_* \mathbf{K} \boldsymbol{\Delta}^T \end{pmatrix} (\boldsymbol{\Delta} \mathbf{K} \boldsymbol{\Delta}^T + \sigma^2 \mathbf{I})^{-1} \begin{pmatrix} \boldsymbol{\Delta} \mathbf{K} & \boldsymbol{\Delta} \mathbf{K} \boldsymbol{\Delta}_*^T \end{pmatrix} \right] \quad (\text{A.6})$$

Marginalising over $\mathbf{n}_e$ is equivalent to neglecting the sub-block corresponding to $\mathbf{n}_e$. Therefore, the RHS is,

$$\int P(\mathbf{n}_e, \boldsymbol{\tau} | \boldsymbol{\tau}_{\text{obs}}) d\mathbf{n}_e = \mathcal{N}[\boldsymbol{\Delta}_* \mathbf{K} \boldsymbol{\Delta}^T (\boldsymbol{\Delta} \mathbf{K} \boldsymbol{\Delta}^T + \sigma^2 \mathbf{I})^{-1} \boldsymbol{\tau}_{\text{obs}}, \boldsymbol{\Delta}_* \mathbf{K} \boldsymbol{\Delta}_*^T - \boldsymbol{\Delta}_* \mathbf{K} \boldsymbol{\Delta}^T (\boldsymbol{\Delta} \mathbf{K} \boldsymbol{\Delta}^T + \sigma^2 \mathbf{I})^{-1} \boldsymbol{\Delta} \mathbf{K} \boldsymbol{\Delta}_*^T]. \quad (\text{A.7})$$

■

Appendix B

Derivation of the ΔTEC variance function and its limits

We derive the ΔTEC variance function $\sigma_{\Delta\text{TEC}}^2(d)$ for zenith observations ($\mathbf{k} = \mathbf{k}' = \hat{\mathbf{z}}$) by considering a baseline between an antenna-of-interest at $\mathbf{x}_i = \mathbf{x}_j$ and a reference antenna at $\mathbf{x}_0 = \mathbf{0}$. 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 $K_{\Delta\text{TEC}}([\mathbf{x}_i, \mathbf{x}_0, \hat{\mathbf{k}}], [\mathbf{x}_j, \mathbf{x}_0, \hat{\mathbf{k}}'])$, and find that in this particular case

$$\sigma_{\Delta\text{TEC}}^2(d) \triangleq K_{\Delta\text{TEC}}([\mathbf{x}_i, \mathbf{x}_0, \hat{\mathbf{z}}], [\mathbf{x}_i, \mathbf{x}_0, \hat{\mathbf{z}}]) \quad (\text{B.1})$$

$$= \sum_{p_1=0}^1 \sum_{p_2=0}^1 (-1)^{p_1+p_2} \int_0^b \int_0^b K_{n_e}(\|\mathbf{x}_{(1-p_1)i} - \mathbf{x}_{(1-p_2)i} + \hat{\mathbf{z}}(s_1 - s_2)\|) \, ds_1 \, ds_2, \quad (\text{B.2})$$

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

$$\sigma_{\Delta\text{TEC}}^2(d) = 2 \int_0^b \int_0^b K_{n_e}(\|s_1 - s_2\|) - K_{n_e}(\sqrt{d^2 + (s_1 - s_2)^2}) \, ds_1 \, ds_2. \quad (\text{B.3})$$

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 $\sigma_{\Delta\text{TEC}}^2(d)$ 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 $K_{n_e} = \sigma_{n_e}^2 C_{n_e}$, 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.

$$\sigma_{\Delta\text{TEC}}^2(d) = 2\sigma_{n_e}^2 h^2 \int_0^{\frac{b}{h}} \int_0^{\frac{b}{h}} C_{n_e}(h|u_1 - u_2|) - C_{n_e}\left(h\sqrt{\left(\frac{d}{h}\right)^2 + (u_1 - u_2)^2}\right) du_1 du_2. \quad (\text{B.4})$$

We first note that the variance of ΔTEC is simply proportional to the variance of n_e . Secondly, we note that $h|u_1 - u_2| < h\sqrt{\left(\frac{d}{h}\right)^2 + (u_1 - u_2)^2}$ for any non-zero d , so that $C_{n_e}(h|u_1 - u_2|) > C_{n_e}\left(h\sqrt{\left(\frac{d}{h}\right)^2 + (u_1 - u_2)^2}\right)$ 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. $\frac{d}{h} \rightarrow 0$, we have $C_{n_e}\left(h\sqrt{\left(\frac{d}{h}\right)^2 + (u_1 - u_2)^2}\right) \rightarrow C_{n_e}(h|u_1 - u_2|)$. We therefore find that $\sigma_{\Delta\text{TEC}}^2 \rightarrow 0$ 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 $\sqrt{\left(\frac{d}{h}\right)^2 + (u_1 - u_2)^2} \approx \frac{d}{h}$, since $(u_1 - u_2)^2 < \left(\frac{b}{h}\right)^2 \ll \left(\frac{d}{h}\right)^2$. Assuming $C_{n_e}(d) \approx 0$ when $\frac{d}{h} \gg 1$, the integrand reduces to $C_{n_e}(h|u_1 - u_2|) - C_{n_e}\left(h \cdot \frac{d}{h}\right) \approx C_{n_e}(h|u_1 - u_2|)$. We find that in this case,

$$\sigma_{\Delta\text{TEC}}^2 \approx 2\sigma_{n_e}^2 h^2 \int_0^{\frac{b}{h}} \int_0^{\frac{b}{h}} C_{n_e}(h|u_1 - u_2|) du_1 du_2. \quad (\text{B.5})$$

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 B.5 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

$$\mathbb{V}(\tau_i^{\hat{z}}) = \sigma_{n_e}^2 h^2 \int_0^{\frac{b}{h}} \int_0^{\frac{b}{h}} C_{n_e}(h|u_1 - u_2|) du_1 du_2. \quad (\text{B.6})$$

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

$$\sigma_{\Delta\text{TEC}}^2 = \mathbb{V}(\tau_i^{\hat{z}} - \tau_0^{\hat{z}}) = \mathbb{V}(\tau_i^{\hat{z}}) + \mathbb{V}(\tau_0^{\hat{z}}), \quad (\text{B.7})$$

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 B.6 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 B.4 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 $\inf_{\mathbb{R}} \{C_{n_e}(r) : r \in \mathbb{R}_{>0}\} \triangleq I$, we find the inequality,

$$\sigma_{\Delta\text{TEC}}^2 \leq 2\sigma_{n_e}^2 h^2 \int_0^{\frac{b}{h}} \int_0^{\frac{b}{h}} 1 - I \, du_1 du_2 = 2(1 - I) \sigma_{n_e}^2 b^2. \quad (\text{B.8})$$

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 $\sigma_{\Delta\text{TEC}}^2 \leq 2\sigma_{n_e}^2 b^2$. Kernels resulting in anti-correlated FEDs produce the constraint $\sigma_{\Delta\text{TEC}}^2 \leq 4\sigma_{n_e}^2 b^2$ or tighter. By measuring $\sigma_{\Delta\text{TEC}}(d)$, one can bound the product $\sigma_{n_e} b$ from below. The strongest bound is obtained for large d .

Appendix C

Factoring commutative DI dependence from the RIME

We consider the effect of directionally referencing spatially referenced commutative Jones scalars. In particular, we study phases of the Jones scalars, and assume amplitudes of one, but the same idea extends to amplitude by considering log-amplitudes that can be treated like a pure-imaginary phase.

Let $g(\mathbf{x}, \mathbf{k}) = e^{i\phi(\mathbf{x}, \mathbf{k})}$ be a Jones scalar, and consider the necessarily non-unique decomposition of phase into DD and DI components,

$$\phi(\mathbf{x}, \mathbf{k}) = \phi^{\text{DD}}(\mathbf{x}, \mathbf{k}) + \phi^{\text{DI}}(\mathbf{x}). \quad (\text{C.1})$$

This functional form only specifies that the DI term is not dependent on $\mathbf{k}$. The differential phase to which the RIME is sensitive is found by spatially referencing the phase,

$$\Delta_0 \phi(\mathbf{x}, \mathbf{k}) = \phi(\mathbf{x}, \mathbf{k}) - \phi(\mathbf{x}_0, \mathbf{k}) \quad (\text{C.2})$$

$$= \Delta_0 \phi^{\text{DD}}(\mathbf{x}, \mathbf{k}) + \Delta_0 \phi^{\text{DI}}(\mathbf{x}). \quad (\text{C.3})$$

Directionally referencing the differential phase to direction $\mathbf{k}_0$, we have

$$\Delta_0^2 \phi(\mathbf{x}, \mathbf{k}) = \Delta_0 \phi(\mathbf{x}, \mathbf{k}) - \Delta_0 \phi(\mathbf{x}, \mathbf{k}_0) \quad (\text{C.4})$$

$$= \Delta_0^2 \phi^{\text{DD}}(\mathbf{x}, \mathbf{k}). \quad (\text{C.5})$$

We see that the DI phase has disappeared, and we are left with the doubly differential phase for the DD term. We then assume that there was a remnant DI component in $\phi^{\text{DD}}(\mathbf{x}, \mathbf{k})$. Then by induction we have that $\Delta_0^2 \phi^{\text{DD}}(\mathbf{x}, \mathbf{k})$ must be free of DI terms, and $\Delta_0^2 \phi(\mathbf{x}, \mathbf{k})$ must be free of all DI components. It follows that directionally referencing phase guarantees that all DI components are removed.

Furthermore, assuming we were to directionally reference a doubly differential phase. We assume that the doubly differential phase is referenced to direction $\mathbf{k}'_0$, with $\Delta_0'^2 \phi(\mathbf{x}, \mathbf{k}) \triangleq$

$\Delta_0\phi(\mathbf{x}, \mathbf{k}) - \Delta_0\phi(\mathbf{x}, \mathbf{k}'_0)$. Then we see

$$\Delta_0'^2\phi(\mathbf{x}, \mathbf{k}) - \Delta_0'^2\phi(\mathbf{x}, \mathbf{k}_0) = \Delta_0\phi(\mathbf{x}, \mathbf{k}) - \Delta_0\phi(\mathbf{x}, \mathbf{k}'_0) - (\Delta_0\phi(\mathbf{x}, \mathbf{k}_0) - \Delta_0\phi(\mathbf{x}, \mathbf{k}'_0)) \quad (\text{C.6})$$

$$= \Delta_0\phi(\mathbf{x}, \mathbf{k}) - \Delta_0\phi(\mathbf{x}, \mathbf{k}_0) \quad (\text{C.7})$$

$$= \Delta_0^2\phi(\mathbf{x}, \mathbf{k}). \quad (\text{C.8})$$

That is, directionally referencing a doubly differential phase produces a new doubly differential phase, referenced to the new direction, $\mathbf{k}_0$. This trick thus allows us to set a well-defined reference direction to phases that have undergone DI calibration before DD calibration.

Appendix D

Recursive Bayesian estimation

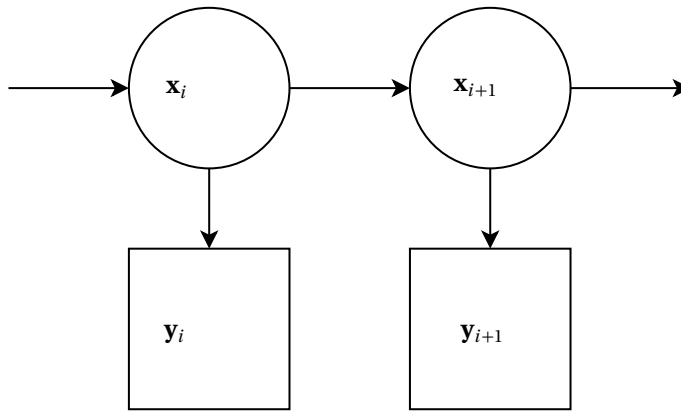


Figure D.1: Casual graph depicting a hidden Markov model.

Recursive Bayesian estimation is a method of performing inference on a hidden Markov model [HMM; Rabiner and Juang, 1986]. Let $\mathbf{y}$ be an observable and $\mathbf{x}$ be a hidden variable. The HMM assumptions on $\mathbf{x}$ and $\mathbf{y}$ are, firstly, that the hidden random variable is only conditionally dependent on its previous state, and secondly, the observable is conditionally independent of all other random variables except for the current hidden state. These assumptions are depicted in **Figure D.1**, where i is the sequence index. This paradigm is often given a notion of causality or time, but this is not necessary in any way. The sequence index is an abstract notion that simply explains how the set of hidden variables are traversed. For example, we assume that the observations are frames of a movie, and the hidden variable is the plot contained in each movie frame. The HMM assumption is that the movie plot is linear, and the picture encodes what is going on in the movie at a given point in time. With recursive Bayesian estimation the movie can be watched with frames randomly ordered and the complete plot is still rendered completely.

There are two distinct types of information propagation in a hidden Markov model. Information can flow in the direction of the arrows, or against them. This leads to the notion

of the forward and backwards equations that describe how belief in hidden variables is propagated forward, and for revising, we believe in previously visited hidden variables.

The joint distribution of the hidden random variables and observables in a chain of length T can be written out as a product of conditional distributions and a marginal using the product rule of probability distributions [Kolmogorov, 1960]. Because of the HMM assumptions the joint distribution is

$$p(\mathbf{x}_{0:T}, \mathbf{y}_{0:T}) = p(\mathbf{x}_0) \prod_{i=1}^T p(\mathbf{x}_i | \mathbf{x}_{i-1}) p(\mathbf{y}_i | \mathbf{x}_i). \quad (\text{D.1})$$

We first consider propagating information forward. This is done in two steps, typically called the predict and update steps. For the predict step, we consider how belief in the absence of new observables is propagated. For this, we apply the Chapman-Kolmogorov identity¹ for Markovian processes,

$$p(\mathbf{x}_{i+1} | \mathbf{y}_{0:i}) = \mathbb{E}_{\mathbf{x}_i | \mathbf{y}_{0:i}} [p(\mathbf{x}_{i+1} | \mathbf{x}_i)], \quad (\text{D.2})$$

which gives us the probability distribution of the hidden variables at time $i + 1$ in terms of the so-called state transition distribution $p(\mathbf{x}_{i+1} | \mathbf{x}_i)$ and posterior distribution $p(\mathbf{x}_i | \mathbf{y}_{0:i})$ at index i . The current prior belief can be understood as the expectation of the state transition distribution over the measure of the current posterior belief.

The update step is simply an application of the Bayes theorem with the prior defined by the predict step,

$$p(\mathbf{x}_i | \mathbf{y}_{0:i}) = \frac{p(\mathbf{y}_i | \mathbf{x}_i) \mathbb{E}_{\mathbf{x}_{i-1} | \mathbf{y}_{0:i-1}} [p(\mathbf{x}_i | \mathbf{x}_{i-1})]}{p(\mathbf{y}_i | \mathbf{y}_{0:i-1})}, \quad (\text{D.3})$$

where the denominator is the Bayesian evidence of the newly arrived data given all previous data, and it is independent of the hidden variables. Equation D.3 gives a recurrence relation for propagating our belief forward, therefore this is called the forward equation.

We now assume that we are at index T , and wish to use all acquired information to revise our belief in the previously visited hidden variables at indices $i < T$. The trick is to realise that $p(\mathbf{x}_i | \mathbf{x}_{i+1} | \mathbf{y}_{0:T}) = p(\mathbf{x}_i | \mathbf{x}_{i+1} | \mathbf{y}_{0:i})$ as a result of the Markov properties. In this case, again using the product rule, we find that the joint conditional distribution of a pair of sequential hidden states given the whole sequence of data is

$$p(\mathbf{x}_i, \mathbf{x}_{i+1} | \mathbf{y}_{0:T}) = \frac{p(\mathbf{x}_{i+1} | \mathbf{x}_i) p(\mathbf{x}_i | \mathbf{y}_{0:i}) p(\mathbf{x}_{i+1} | \mathbf{y}_{0:T})}{p(\mathbf{x}_{i+1} | \mathbf{y}_{0:i})}. \quad (\text{D.4})$$

Marginalising the second hidden parameter, we arrive at the recurrence relation

$$p(\mathbf{x}_i | \mathbf{y}_{0:T}) = p(\mathbf{x}_i | \mathbf{y}_{0:i}) \int \frac{p(\mathbf{x}_{i+1} | \mathbf{x}_i) p(\mathbf{x}_{i+1} | \mathbf{y}_{0:T})}{p(\mathbf{x}_{i+1} | \mathbf{y}_{0:i})} d\mathbf{x}_{i+1}. \quad (\text{D.5})$$

This can be solved by starting at T and solving this equation iteratively backwards, therefore Equation D.5 is called the backward equation. Most importantly, we note that the backward equation does not require conditioning on data, as was done in the update step.

¹If b is conditionally independent of a , then $p(a | c) = \int p(a | b) p(b | c) db = \mathbb{E}_{b|c} [p(a | b)]$ is the Chapman-Kolmogorov identity.

When the transition and likelihood are assumed to be Gaussian, for example, in a linear dynamical system, the forward and backward equations are equivalent to the well-known Kalman filter equations and Rauch smoother equations [Rauch, 1963], respectively.

Appendix E

Jones scalar variational expectation

Let $\mathbf{g} \in \mathbb{C}^{N_{\text{freq}}}$ be an observed complex Jones scalar vector, with amplitudes $g \in \mathbb{R}^{N_{\text{freq}}}$ and phases $\phi \in \mathbb{R}^{N_{\text{freq}}}$. We assume that the Jones scalars have complex Gaussian noise, described by the observational covariance matrix Σ . Thus, we have that the observational likelihood of the Jones scalars (*cf* Eq. 3.9) is

$$p(\mathbf{g} | \phi, g, \Sigma) = \mathcal{N}_{\mathbb{C}}[\mathbf{g} | g e^{i\phi}, \Sigma], \quad (\text{E.1})$$

where $\mathcal{N}_{\mathbb{C}}$ is the complex Gaussian distribution, which is defined as the Gaussian distribution of the stacked real and imaginary components. We define the residuals $\delta R = \text{Re}[\mathbf{g}] - g \cos \phi$ and $\delta I = \text{Im}[\mathbf{g}] - g \sin \phi$, and the stacked residuals $\delta \mathbf{g} = (\delta R, \delta I)^T$. Then the log-likelihood becomes

$$\log p(\mathbf{g} | \phi, g, \Sigma) = -N_{\text{freq}} \log 2\pi - \frac{1}{2} \log |\Sigma| - \frac{1}{2} \delta \mathbf{g}^T \Sigma^{-1} \delta \mathbf{g}, \quad (\text{E.2})$$

$$= -N_{\text{freq}} \log 2\pi - \frac{1}{2} \log |\Sigma| - \frac{1}{2} \text{Tr}[\Sigma^{-1} \delta \mathbf{g} \delta \mathbf{g}^T] \quad (\text{E.3})$$

$$= -N_{\text{freq}} \log 2\pi - \frac{1}{2} \log |\Sigma| - \frac{1}{2} \text{vec}[\Sigma^{-1}]^T \text{vec}[\delta \mathbf{g} \delta \mathbf{g}^T] \quad (\text{E.4})$$

where we used the fact that the trace of a scalar is a scalar, and that $\text{Tr}[A^T B] = \text{Tr}[B A^T] = \text{vec}[A]^T \text{vec}[B]$.

Now we assume that the phases are linearly modelled according to

$$\phi(\nu) = \sum_i^M f_i(\nu) a_i, \quad (\text{E.5})$$

$$\triangleq f^T(\nu) a \quad (\text{E.6})$$

where $f_i(\nu)$ is the i -th basis function of a set of M linearly independent functions depending on frequency ν , and a_i is the i -th parameter in the linear model. In the second line we have

written the sum as a dot product. We furthermore assume that the parameter vector, a , is a (potentially correlated) Gaussian random variable,

$$a \sim \mathcal{N}[\mu, \Gamma]. \quad (\text{E.7})$$

This induces a distribution on phase via Eq. E.5.

The variational expectation of the Jones scalars with respect to a is the expectation of the log-likelihood,

$$\begin{aligned} \mathbb{E}_a [\log p(\mathbf{g} | \phi, g, \Sigma)] = & -N_{\text{freq}} \log 2\pi - \frac{1}{2} \log |\Sigma| \\ & - \frac{1}{2} \text{vec}[\Sigma^{-1}]^T \text{vec}[\mathbb{E}_a[\delta \mathbf{g} \delta \mathbf{g}^T]], \end{aligned} \quad (\text{E.8})$$

where we used the fact that expectation and vectorisation commute. Now we need to evaluate the expectation, $\mathbb{E}_a[\delta \mathbf{g} \delta \mathbf{g}^T]$. In order to do this, we realise that there are three explicit expectations that must be done corresponding to sub-blocks of $\delta \mathbf{g} \delta \mathbf{g}^T$ consisting of the real-real, real-imaginary, and imaginary-imaginary outer products of residual vectors,

$$\mathbb{E}_a[\delta R \delta R^T] = \int \delta R \delta R^T dp(a) \quad (\text{E.9})$$

$$\triangleq I_{rr} \quad (\text{E.10})$$

$$\mathbb{E}_a[\delta R \delta I^T] = \int \delta R \delta I^T dp(a) \quad (\text{E.11})$$

$$\triangleq I_{ri} \quad (\text{E.12})$$

$$\mathbb{E}_a[\delta I \delta I^T] = \int \delta I \delta I^T dp(a) \quad (\text{E.13})$$

$$\triangleq I_{ii}, \quad (\text{E.14})$$

where $p(a) = \mathcal{N}[\mu, \Gamma]$. All three of the above integrals involve Gaussian expectations of trigonometric functions. For example, the (i, j) -th element of the integral I_{rr} is,

$$(I_{rr})_{(i,j)} = \int (\text{Re}[\mathbf{g}]_i - g_i \cos \phi(\nu_i))(\text{Re}[\mathbf{g}]_j - g_j \cos \phi(\nu_j)) dp(a). \quad (\text{E.15})$$

To evaluate this integral, we use angle addition formulae to reduce all powers of trigonometric functions. We then use the characteristic function of the Gaussian distribution, which gives the relation, $\mathbb{E}_{X \sim \mathcal{N}[\mu, \Gamma]}[e^{if^T X}] = e^{if^T \mu - \frac{1}{2} f^T \Gamma f}$, on the exponential form of cosine and sine. Applying the above, we evaluate all integral matrix elements,

$$\begin{aligned} (I_{rr})_{(i,j)} = & \text{Re}[\mathbf{g}]_i \text{Re}[\mathbf{g}]_j - g_i \text{Re}[\mathbf{g}]_j h_1(f(\nu_i)) - \text{Re}[\mathbf{g}]_i g_j h_1(f(\nu_j)) \\ & + g_i g_j h_3(f(\nu_i), f(\nu_j)) \end{aligned} \quad (\text{E.16})$$

$$\begin{aligned} (I_{ri})_{(i,j)} = & \text{Re}[\mathbf{g}]_i \text{Im}[\mathbf{g}]_j - g_i \text{Im}[\mathbf{g}]_j h_1(f(\nu_i)) - \text{Re}[\mathbf{g}]_i g_j h_2(f(\nu_j)) \\ & + g_i g_j h_5(f(\nu_i), f(\nu_j)) \end{aligned} \quad (\text{E.17})$$

$$\begin{aligned} (I_{ii})_{(i,j)} = & \text{Im}[\mathbf{g}]_i \text{Im}[\mathbf{g}]_j - g_i \text{Im}[\mathbf{g}]_j h_2(f(\nu_i)) - \text{Im}[\mathbf{g}]_i g_j h_2(f(\nu_j)) \\ & + g_i g_j h_4(f(\nu_i), f(\nu_j)) \end{aligned} \quad (\text{E.18})$$

where,

$$h_1(x) = \mathbb{E}_a[\cos x^T a] = e^{-\frac{1}{2}x^T \Gamma x} \cos x^T \mu \quad (\text{E.19})$$

$$h_2(x) = \mathbb{E}_a[\sin x^T a] = e^{-\frac{1}{2}x^T \Gamma x} \sin x^T \mu \quad (\text{E.20})$$

$$h_3(x, y) = \mathbb{E}_a[\cos x^T a \cos y^T a] = \frac{1}{2}(h_1(x + y) + h_1(x - y)) \quad (\text{E.21})$$

$$h_4(x, y) = \mathbb{E}_a[\sin x^T a \sin y^T a] = \frac{1}{2}(h_1(x - y) - h_1(x + y)) \quad (\text{E.22})$$

$$h_5(x, y) = \mathbb{E}_a[\cos x^T a \sin y^T a] = \frac{1}{2}(h_2(x + y) + h_2(x - y)). \quad (\text{E.23})$$

This completes the evaluation of Eq. E.8 for an arbitrary linear phase model. For example, in a phase model consisting only of DDTEC, we have $M = 1$, $f_1(\nu) = \frac{\kappa}{\nu}$ and $a_1 = \Delta_0^2 \tau$. It is simple to calculate the variational expectation of other linear phase models including terms such as clock-like terms, constant-in-frequency terms, and higher order ionospheric terms.

A simplification follows if the observational errors are assumed uncorrelated, which leads to a reduction of the number of operations required to compute the variational expectation. In this case, only the diagonals of I_{rr} and I_{ii} are needed, and I_{ri} is not needed at all.

Appendix F

Derivation of the M-step

Here we derive the maximisation step of Algorithm 1. This is an extension of Shumway and Stoffer [1982] to time varying observational covariance and state transition covariance. In order to do this we must linearise the forward equation. Let the set of HMM states be $X = \{x_i \mid i = 1..n\}$ and the set of observations be $Y = \{y_i \mid i = 1..n\}$. Following similar notation as in Shumway and Stoffer [1982] we'll term x_i^s be the state at step i conditioned on data up to an including step s . Therefore if $i > s$ then x_i^s is the forecast of the state. Consequently the conditional mean and covariance is μ_i^s and Γ_i^s . We'll also need the joint-covariance $\Gamma_{i,i-1}^s$ which is the covariance of $\{x_i, x_{i-1}\}$ conditioned on data up to step s . The recursive equations for all these components are given in [e.g. Shumway and Stoffer, 1982] and follow from the forward and backward equations in Appendix D.

Since at each time step the state is a Gaussian variable, using standard identities it follows that the posterior distribution of $x_i - x_{i-1}$ conditioned up to step n is given by,

$$x_i - x_{i-1} \sim \mathcal{N}(x_i - x_{i-1} \mid \mu_i^n - \mu_{i-1}^n, \Gamma_i^n + \Gamma_{i-1}^n - \Gamma_{i,i-1}^n - \Gamma_{i-1,i}^n). \quad (\text{E.1})$$

The joint log-likelihood of the data and state is the sum of the prior component and the likelihood component,

$$\begin{aligned} \log L(X, Y) &\triangleq -\frac{1}{2} \log |\Gamma_0| - \frac{1}{2} (x_0 - \mu_0)^T \Gamma_0^{-1} (x_0 - \mu_0) \\ &\quad - \frac{1}{2} \sum_{i=1}^n \log |\Omega_i| - \frac{1}{2} \sum_{i=1}^n (x_i - x_{i-1})^T \Omega_i^{-1} (x_i - x_{i-1}) \\ &\quad - \frac{1}{2} \sum_{i=1}^n \log |\Sigma_i| - \frac{1}{2} \sum_{i=1}^n (y_i - f_i(x_i))^T \Sigma_i^{-1} (y_i - f_i(x_i)) + \text{const.} \end{aligned} \quad (\text{E.2})$$

where f_i is the forward function from the HMM state to observable and time i . In general, the function f_i can change over time, e.g. deterministic input control parameters to f_i can be time dependent. Let us define the linear map $D_i(y) = \nabla_x f_i|_{\tilde{x}_i} y$ for some $\tilde{x}_i$. The forward

function then satisfies $f_i(x_i) = f_i(\bar{x}_i) + D_i(x_i - \bar{x}_i) + o(|x_i - \bar{x}_i|^2)$. From this Eq. F2 becomes,

$$\begin{aligned} \log L(X, Y) = & -\frac{1}{2} \log |\Gamma_0| - \frac{1}{2} (x_0 - \mu_0)^T \Gamma_0^{-1} (x_0 - \mu_0) \\ & - \frac{1}{2} \sum_{i=1}^n \log |\Omega_i| - \frac{1}{2} \sum_{i=1}^n (x_i - x_{i-1})^T \Omega_i^{-1} (x_i - x_{i-1}) \\ & - \frac{1}{2} \sum_{i=1}^n \log |\Sigma_i| \\ & - \frac{1}{2} \sum_{i=1}^n (y_i - f_i(\bar{x}_i) - D_i(x_i - \bar{x}_i))^T \Sigma_i^{-1} (y_i - f_i(\bar{x}_i) - D_i(x_i - \bar{x}_i)) \\ & + \text{const.} \end{aligned} \quad (\text{F3})$$

which is exactly equal to Eq. F2 when $x_i = \bar{x}_i$ and f_i quadratic in the neighbourhood of $\bar{x}_i$. Defining $\delta y_i = y_i - f_i(\bar{x}_i) + D_i \bar{x}_i$ then Eq. F3 becomes,

$$\begin{aligned} \log L(X, Y) = & -\frac{1}{2} \log |\Gamma_0| - \frac{1}{2} (x_0 - \mu_0)^T \Gamma_0^{-1} (x_0 - \mu_0) \\ & - \frac{1}{2} \sum_{i=1}^n \log |\Omega_i| - \frac{1}{2} \sum_{i=1}^n (x_i - x_{i-1})^T \Omega_i^{-1} (x_i - x_{i-1}) \\ & - \frac{1}{2} \sum_{i=1}^n \log |\Sigma_i| \\ & - \frac{1}{2} \sum_{i=1}^n (\delta y_i - D_i x_i)^T \Sigma_i^{-1} (\delta y_i - D_i x_i) \\ & + \text{const.} \end{aligned} \quad (\text{F4})$$

Taking the expectation of Eq. F4 with respect to the state conditioned on all data we find [cf. Shumway and Stoffer, 1982],

$$\begin{aligned} \mathbb{E}_X[\log L(X, Y)] = & -\frac{1}{2} \log |\Gamma_0| - \frac{1}{2} \sum_{i=1}^n \log |\Omega_i| - \frac{1}{2} \sum_{i=1}^n \log |\Sigma_i| \\ & - \frac{1}{2} (\mu_0^n - \mu_0)^T \Gamma_0^{-1} (\mu_0^n - \mu_0) - \frac{1}{2} \text{Tr}[\Gamma_0^{-1} \Gamma_0^n] \\ & - \frac{1}{2} \sum_{i=1}^n (\mu_i^n - \mu_{i-1}^n)^T \Omega_i^{-1} (\mu_i^n - \mu_{i-1}^n) \\ & - \frac{1}{2} \sum_{i=1}^n \text{Tr}[\Omega_i^{-1} (\Gamma_i^n + \Gamma_{i-1}^n - \Gamma_{i,i-1}^n - \Gamma_{i-1,i}^n)] \\ & - \frac{1}{2} \sum_{i=1}^n (\delta y_i - D_i \mu_i^n)^T \Sigma_i^{-1} (\delta y_i - D_i \mu_i^n) \\ & - \frac{1}{2} \sum_{i=1}^n \text{Tr}[\Sigma_i^{-1} D_i \Gamma_i^n D_i^T] \\ & + \text{const.} \end{aligned} \quad (\text{F5})$$

We next take the derivative with respect to μ_0 , Γ_0 , Ω_i , and Σ_i and set to zero and solve. Some derivative identities are required. Assuming A is symmetric and positive definite, then using Einstein summation notation we have,

$$\nabla_{A_{ij}} A_{lq}^{-1} = -A_{li}^{-1} A_{jq}^{-1} \quad (\text{E.6})$$

$$\nabla_{A_{ij}} \log|A| = A_{ij}^{-1} \quad (\text{E.7})$$

$$\nabla_{A_{ij}} x_l A_{lq}^{-1} y_q = -(A_{il}^{-1} x_l)(A_{jq}^{-1} y_q) = -(A^{-1} x)(A^{-1} y)^T \quad (\text{E.8})$$

$$\nabla_{A_{ij}} \text{Tr}[A^{-1} B] = -A_{il}^{-1} B_{lq}^T A_{qj}^{-1} = -A^{-1} B^T A^{-1}. \quad (\text{E.9})$$

Applying these identities as needed we have,

$$\nabla_{\mu_0} \mathbb{E}_X[\log L(X, Y)] = \Gamma_0^{-1}(\mu_0^n - \mu_0) = 0 \quad (\text{E.10})$$

$$\implies \mu_0 = \mu_0^n \quad (\text{E.11})$$

$$\begin{aligned} \nabla_{\Gamma_0} \mathbb{E}_X[\log L(X, Y)] &= -\frac{1}{2} \nabla_{\Gamma_0} \log|\Gamma_0| - \frac{1}{2} \nabla_{\Gamma_0} (\mu_0^n - \mu_0)^T \Gamma_0^{-1} (\mu_0^n - \mu_0) \\ &\quad - \frac{1}{2} \nabla_{\Gamma_0} \text{Tr}[\Gamma_0^{-1} \Gamma_0^n] \end{aligned} \quad (\text{E.12})$$

$$= -\frac{1}{2} \Gamma_0^{-1} + \frac{1}{2} \Gamma_0^{-1} \Gamma_0^n \Gamma_0^{-1} = 0 \quad (\text{E.13})$$

$$\implies \Gamma_0 = \Gamma_0^n \quad (\text{E.14})$$

$$\begin{aligned} \nabla_{\Omega_i} \mathbb{E}_X[\log L(X, Y)] &= -\frac{1}{2} \sum_{i=1}^n \nabla_{\Omega_i} \log|\Omega_i| \\ &\quad - \frac{1}{2} \sum_{i=1}^n \nabla_{\Omega_i} (\mu_i^n - \mu_{i-1}^n)^T \Omega_i^{-1} (\mu_i^n - \mu_{i-1}^n) \\ &\quad - \frac{1}{2} \sum_{i=1}^n \nabla_{\Omega_i} \text{Tr}[\Omega_i^{-1} (\Gamma_i^n + \Gamma_{i-1}^n - \Gamma_{i,i-1}^n - \Gamma_{i-1,i}^n)] \end{aligned} \quad (\text{E.15})$$

$$\begin{aligned} &= -\frac{1}{2} \Omega_i^{-1} + \frac{1}{2} \Omega_i^{-1} (\mu_i^n - \mu_{i-1}^n) (\mu_i^n - \mu_{i-1}^n)^T \Omega_i^{-1} \\ &\quad + \frac{1}{2} \Omega_i^{-1} (\Gamma_i^n + \Gamma_{i-1}^n - \Gamma_{i,i-1}^n - \Gamma_{i-1,i}^n) \Omega_i^{-1} = 0 \end{aligned} \quad (\text{E.16})$$

$$\begin{aligned} \implies \Omega_i &= (\mu_i^n - \mu_{i-1}^n) (\mu_i^n - \mu_{i-1}^n)^T \\ &\quad + (\Gamma_i^n + \Gamma_{i-1}^n - \Gamma_{i,i-1}^n - \Gamma_{i-1,i}^n) \end{aligned} \quad (\text{E.17})$$

$$\begin{aligned} \nabla_{\Sigma_i} \mathbb{E}_X[\log L(X, Y)] &= -\frac{1}{2} \sum_{i=1}^n \nabla_{\Sigma_i} \log|\Sigma_i| \\ &\quad - \frac{1}{2} \sum_{i=1}^n \nabla_{\Sigma_i} (\delta y_i - D_i \mu_i^n)^T \Sigma_i^{-1} (\delta y_i - D_i \mu_i^n) \end{aligned}$$

$$-\frac{1}{2} \sum_{i=1}^n \nabla_{\Sigma_i} \text{Tr}[\Sigma_i^{-1} D_i \Gamma_i^n D_i^T] \quad (\text{E18})$$

$$= -\frac{1}{2} \Sigma_i^{-1} + \frac{1}{2} \Sigma_i^{-1} (\delta y_i - D_i \mu_i^n) (\delta y_i - D_i \mu_i^n)^T \Sigma_i^{-1} \\ + \frac{1}{2} \Sigma_i^{-1} D_i^T \Gamma_i^n D_i \Sigma_i^{-1} \quad (\text{E19})$$

$$= 0 \quad (\text{E20})$$

$$\Rightarrow \Sigma_i = (\delta y_i - D_i \mu_i^n) (\delta y_i - D_i \mu_i^n)^T + D_i^T \Gamma_i^n D_i \quad (\text{E21})$$

This completes the derivation of the update equations of the M-step.

English Summary

What we know is a drop, and what we don't know, an ocean. Some claim that the only things we can know are those which are seen with the naked eye. To take in light, and produce knowledge. That is the way information, raw and without interpretation, makes its way to places like Wikipedia, books, schools, and into the minds of the next generation. The next time you experience a blackout at night in your city go outside and look up. If luck favours you, you won't find clouds but rather the cosmos staring down at you. You'll note that most of those lights twinkle, like a distant candle in the wind. An odd few you'll note do not twinkle, and, as you may already know, those are the planets of our solar system. Mars, Venus, Jupiter... these characters of myth conjure up scenes of the gods' wars, loves, and vengeance, and it doesn't take long before your mind starts to wander. When was it, the last time you looked up in the night sky?

If you could see in the radio spectrum, that is if your eyes could sense wavelengths greater than 1 cm, you would see a night sky that looks vastly different from the star-strewn scene you're used to in the visible spectrum. The radio sky contains large diffuse structures, compact pulsating sources, filamentary structures, ring like structures, and more. These shapes correspond to a diversity of physical processes at play throughout the Universe.

One such physical process that we discuss in this thesis is synchrotron emission resulting from the merging of galaxy clusters. Synchrotron emission is radio frequency light that results from fast moving electric charges in the presence of an ambient magnetic field. Galaxy clusters, on a scale some 100 millions times larger than our solar system, are the largest gravitational wells of the Universe with masses on the order of $10^{14} M_{\odot}$. Galaxy clusters form the vertices of the cosmic web, with elongated filaments and sheets connecting them. Since the scales are so large, they present an interesting laboratory for testing physics.

Occasionally, galaxy clusters merge with matter in-falling from the cosmic web with relative speeds exceeding 1000 km s^{-1} . Such merger events can be thought of as the analogue of particle accelerators here on Earth, where we test theories of physics on the smallest scales. A merger event produces large outward travelling shocks in the intra-cluster medium (ICM), releasing gigantic amounts of kinetic energy (10^{56} J) in the form of turbulence and acceleration of cosmic ray electrons (CRE) to ultra-relativistic speeds. These CRE emit synchrotron emission that traces the dynamics of the mergers, and with advanced computational methods allow us to test theories of physics on the largest scales.

In Chapter 1 we investigate one particular galaxy cluster (called PLCK G004.5-19.5) that has the interesting property of being at large cosmological time (with a redshift of $z = 0.52$). We analyse the cluster at 150 MHz, 325 MHz, and 610 MHz with the Giant Metrewave Radio Telescope [GMRT; Gupta et al., 2017] and find emission that suggests the cluster has under-

gone a merger in the past, however the arrangement of emission is too complicated to say much more. An interesting finding is that the central region of the cluster appears to have a diffuse component that resembles what is called a radio halo. Radio halos are thought to be driven by turbulence, however their existence at high redshift is rare. Their rarity comes from the fact that radiant energy is lost through a process called inverse Compton scattering, and this energy loss goes as $\propto (1+z)^4$. Therefore, it's not that they don't exist at high redshift, they are merely too dim. In fact, this cut-off in cosmological redshift acts as a valuable means of testing cosmology. The prevalence of faint complicated emission in high-redshift galaxy clusters, and the hidden treasure trove of information, prompts the need for high-resolution, low-frequency, wide-bandwidth, high-sensitivity radio telescopes.

Radio astronomy has changed a lot since it began. Early pioneers of radio interferometry [Ryle and Vonberg, 1946] acquired so-called 'visibilities' with analog-driven pens on paper. They could touch their data with their hands. Today, a radio astronomer ironically rarely sees the visibilities, and the trend is increasing in that direction. Radio telescopes are some of the biggest data experiments of humankind. One such telescope that occupies a large portion of this thesis is the Low-Frequency Array [LOFAR; van Haarlem et al., 2013].

Since there is a clear need for deeper and better radio images, a large part of radio astronomy is technical in nature and known as calibration and imaging. In short, the visibilities characterise how light from two very distant, but neighbouring, sources in the sky interferes with each other. Much like how the interference pattern in Thomas Young's double-slit experiment told him how far apart the slits were, the visibilities tell us how far apart radio sources are on the sky. Calibration and imaging is the process of using computers to invert large sets of visibilities.

Chapters 2 to 4 deal with the aspect of calibration. Calibration has to do with the fact that the telescope and medium between the telescope and sources are constantly changing. Physical properties like temperature and humidity can change how a radio antenna receives light, and something known as the ionosphere alters how radio waves propagate to the telescope. In this thesis we are concerned with calibration of the ionosphere.

One of the challenges of ionospheric calibration is that you can't directly see the ionosphere. There is a long history of trying to solving this challenging problem. There are two main types of methods for ionospheric calibration, field-based calibration [Cotton et al., 2004] which models image domain apparent source shifts, and facet-based calibration [Intema et al., 2009, van Weeren et al., 2016, Tasse et al., 2018] which performs calibration on disjoint patches of the sky called facets. In this thesis we focus entirely on the facet-based approach.

Facet-based calibration has several orders of complexity, however we will just sketch the basic idea here. The idea is that when radio frequency light traverses a medium with a low density of electric charges – as in the case of the ionosphere – the light wave takes on a small time delay. This time delay is proportional to how many electrons the wave passes, and so one can imagine counting all the electrons between the radio antenna and the source and using this quantity to calibrate the effect of the ionosphere. In a way that is what facet-based calibration does, except once for each bright source in the field of view.

How does this work though, since one cannot see the ionosphere? Indeed, the challenge all lies in inferring a physical quantity based only on the light that the antennas receive. In this thesis we borrow a concept from an entirely different field called seismic imaging. In seismic imaging the goal is to see inside the Earth by bouncing acoustic waves off of structures inside and recording them on the surface again. This method is built on the principle of

tomography.

Tomography and probability theory is used in Chapter 2 to derive a method that can infer the ionosphere's effect, and we show using simulations that it is superior to an array of other methods that aim to do the same thing. The superiority follows because we have used physics to place strong constraints on the result. In general, using physics to constrain the possible outcomes of a process is very powerful.

With this powerful new method proven to work on simulated data, in Chapter 3 we apply it to real data for the first time. In this case we apply it to an 8-hour LOFAR observation that previously showed large calibration errors in the image. We pleasantly find that our method is able to remove these image errors and we have a much clearer view of that small piece of the sky.

We naturally want to know if that success was just luck, and if the method is robust. Therefore, in Chapter 4 we apply the same method to twelve 8 hour LOFAR observations all pointed at the Ursa Major (The Big Dipper) constellation. This area is known as the Lockman Hole, for it shows very little HI emission in the galactic foreground. In this case we observe that there is a general improvement in ten out of twelve cases, however it also brings to our attention a low-level systematic in the data that we were not aware of. This systematic proves to be very difficult to quantify.

Our method has shown great initial success and there are clear directions in which to improve it in order to account for this systematic. This process is typical in science. As methods get better, uncertainty decreases and, in the process, reveals lower-level phenomenon. Often this is the threshold where discoveries and advances are made.

Nederlandse Samenvatting

Vertaald door Martijn Oei

Wat we weten is een druppel, en wat we niet weten, een oceaan. Sommige mensen zeggen dat de enige dingen die we kunnen weten, dingen zijn die je met het blote oog kunt zien. Licht opnemen, en kennis produceren. Dat is de manier waarop informatie, in eerste instantie primitief en zonder interpretatie, uiteindelijk terecht komt op plekken als Wikipedia, in boeken, op scholen, en in de hoofden van de volgende generatie. Ga de volgende keer dat er 's nachts in jouw stad een stroomstoring is eens naar buiten, en kijk dan omhoog. Met een beetje geluk is het niet bewolkt, en zie je de kosmos naar beneden staren. Het zal je dan opvallen dat de meeste lichtjes die je ziet, twinkelen, als verafgelegen kaarsen in de wind. Van een paar zal het je opvallen dat ze niet twinkelen, en – zoals je misschien al zult weten – dat zijn de planeten van ons zonnestelsel. Mars, Venus, Jupiter... Deze mythische figuren doen denken aan verhalen over de oorlogen, liefde, en wraak van goden, en het duurt niet lang voordat je gedachten met je aan de wandel gaan. Wanneer was het voor het laatst dat je 's nachts naar de hemel hebt gekeken?

Als je licht van het radiospectrum zou kunnen zien, dus als je ogen gevoelig waren voor golflengten groter dan een centimeter, dan zou je een hemel zien die er volkomen anders uitziet dan het met-sterren-bezaaide tafereel waaraan je gewend bent vanuit het zichtbare spectrum. De radiohemel bevat grote diffuse structuren, compacte pulserende bronnen, filamentaire structuren, ringachtige structuren, en meer. Deze vormen corresponderen met een scala aan fysische processen die zich in het Universum voltrekken.

Eén zo'n fysisch proces dat we in dit proefschrift bespreken, is synchrotronemissie afkomstig van het samensmelten van clusters van sterrenstelsels. Synchrotronemissie is licht met radiofrequentie dat ontstaat door snelbewegende elektrische ladingen in een omringend magnetisch veld. Clusters van sterrenstelsels, die honderd miljoen keer groter kunnen worden dan ons zonnestelsel, zijn de voornaamste zwaartekrachtspuiten in het Universum, met massa's van zo'n honderdduizend miljard zonsmassa's. Clusters van sterrenstelsels vormen de knopen van het kosmische web, die verbonden worden met uitgerekte filamenten en dunne vliezen. Omdat ze zo ontzettend groot zijn, vormen ze interessante laboratoria om natuurkundige wetten te toetsen.

Soms nemen clusters van sterrenstelsels materiaal op dat met relatieve snelheden van meer dan duizend kilometer per seconde vanuit de rest van het kosmische web komt aanstormen. Zulke botsingen en samensmeltingen kunnen worden gezien als het analogon van deeltjesversnellers hier op aarde, waarmee we fysische theorieën testen die over het allerkleinste gaan. Samensmeltingen van clusters onderling zorgen voor grote naar buiten bewegende

schokgolven in het intraclustermedium (ICM), waarbij gigantische hoeveelheden kinetische energie (10^{56} J) vrijkomen in de vorm van turbulentie en de versnelling van kosmische straling (in het bijzonder elektronen - CRe) naar ultrarelativistische snelheden. Deze CRe stralen synchrotronemissie uit waarmee de dynamica van de samensmeltingen af te leiden is, en waarmee, aan de hand van geavanceerde computationele technieken, natuurkundige theorieën over de grootste schalen getoetst kunnen worden.

In Hoofdstuk 1 onderzoeken we één bepaald cluster van sterrenstelsels (met de naam PLCK G004.5-19.5), dat de interessante eigenschap heeft dat we hem zien zoals hij kosmologisch ver terug in de tijd eruit zag (op een roodverschuiving van $z = 0.52$). We onderzoeken het cluster op 150, 325 en 610 MHz met de Giant Metrewave Radio Telescope [GMRT; Gupta et al., 2017], en vinden emissie die erop wijst dat het cluster in het verleden een samensmelting met een ander ondergaan heeft, maar het patroon van de emissie is te ingewikkeld om er veel meer over te kunnen zeggen. Een interessante bevinding is dat het centrale gebied van het cluster een diffuse component lijkt te hebben die veel weg heeft van wat een radiohalo wordt genoemd. Men denkt dat radiohalos ontstaan door turbulentie, en dat hun bestaan op hoge roodverschuiving zeldzaam is. Die zeldzaamheid is het gevolg van het feit dat stralingsenergie verloren gaat door een proces dat inverse comptonverstrooiing wordt genoemd, en dat energieverlies schaalt als $\propto (1+z)^4$. Daarom is het eigenlijk niet zo dat radiohalos niet bestaan op hoge roodverschuiving, maar meer dat ze te lichtzwak zijn om waargenomen te worden. Het is zelfs zo dat deze observationele begrenzing in kosmologische roodverschuiving een waardevol instrument is om kosmologische modellen te toetsen. De aanwezigheid van lichtzwakke, complexe emissie in clusters van sterrenstelsels op hoge roodverschuiving, en de bijbehorende verborgen schat aan informatie, vraagt om de ontwikkeling van radiotelescopie met hoge resolutie, ontvangst bij lage frequenties, brede waarneembanden en hoge gevoeligheden.

De radiosterrenkunde is flink veranderd sinds de beginnagen. Vroege pioniers van radiointerferometrie [Ryle and Vonberg, 1946] verzamelden data, ‘visibilities’ genoemd, met analoog aangedreven pennen op papier. Ze konden de data aanraken met hun handen. Vandaag de dag ziet een radioastronoom echter zelden de visibilities. Radiotelescopie zijn enkele van de grootste data-experimenten van de mensheid. Eén zo’n telescoop, welke in een groot gedeelte van dit proefschrift centraal zal staan, is de Low-Frequency Array [LOFAR; van Haarlem et al., 2013].

Aangezien er een duidelijke behoefte is aan diepere en betere radiobeelden, is een groot deel van de radiosterrenkunde technisch van aard, en heeft te maken met calibratie en beeldvorming. Kort gezegd karakteriseren de visibilities hoe licht van twee verafgelegen, maar qua hoek nabijgelegen, bronnen aan de hemel met elkaar interfereren. Net zoals het interferentiepatroon in het tweespletenexperiment van Thomas Young hem vertelde hoe ver de spleten uit elkaar stonden, zo vertellen de visibilities ons hoe ver radiobronnen uit elkaar staan aan de hemel. Calibratie en beeldvorming is het proces waarbij computers gebruikt worden om grote hoeveelheden visibilities te inverteren.

Hoofdstukken 2 tot en met 4 behandelen het calibratieaspect. Calibratie is noodzakelijk omdat de telescoop en het medium tussen het instrument en de lichtbronnen voortdurend aan verandering onderhevig zijn. Zaken als temperatuur en luchtvochtigheid kunnen de manier waarop een radioantenne licht ontvangt, veranderen, en een atmosferische laag die bekend staat als de ionosfeer beïnvloedt hoe radiogolven naar de telescoop propageren. In dit proefschrift houden we ons bezig met calibratie van ionosferische effecten.

Eén van de uitdagingen van ionosferische calibratie is dat je de ionosfeer niet direct kunt zien. Er is een lange geschiedenis van mensen die dit uitdagende probleem hebben geprobeerd op te lossen. Er zijn, hoofdzakelijk, twee klassen van methoden voor ionosferische calibratie: veldgebaseerde calibratie [Cotton et al., 2004], waarin ogenschijnlijk bronverschuivingen in het beelddomein worden gemodelleerd, en facetgebaseerde calibratie [Intema et al., 2009, van Weeren et al., 2016, Tasse et al., 2018], waarin men ontkoppelde stukjes hemel calibreert die facetten genoemd worden. In dit proefschrift focussen we ons volledig op de facetgebaseerde aanpak.

Facetgebaseerde calibratie kent verschillende lagen aan complexiteit, maar we zullen hier vooral het basisidee schetsen. Het idee is dat wanneer licht met radiofrequentie door een medium reist met een lage dichtheid aan elektrische ladingen – zoals in de ionosfeer – de lichtgolf een kleine tijdsvertraging oploopt. Deze vertraging is evenredig met het aantal elektronen dat de golf tegenkomt, en je zou je daarom kunnen voorstellen dat je alle elektronen zou kunnen tellen tussen de radioantenne en de lichtbron, en de uitkomst zou kunnen gebruiken om het effect van de ionosfeer weg te calibreren. In zekere zin is dat wat facetgebaseerde calibratie klaarspeelt, maar dan voor iedere heldere bron in het blikveld.

Maar hoe werkt dat dan, aangezien we de ionosfeer niet kunnen zien? Het is inderdaad zo dat de uitdaging er volledig in ligt dat je een fysische grootheid moet afleiden puur en alleen op basis van het licht dat de antennes ontvangen. In dit proefschrift lenen we een concept van een geheel ander wetenschappelijk veld: seismische beeldvorming. Binnen de seismische beeldvorming is het doel om het inwendige van de aarde zichtbaar te maken door akoestische golven te laten weerkaatsen door structuren in de aarde, en de echos weer te meten aan het oppervlak. Deze methode is gebouwd rond het principe van tomografie.

Tomografie en kansrekening worden in Hoofdstuk 2 gebruikt om een methode af te leiden waarmee ionosferische effecten kunnen worden bepaald, en we laten aan de hand van simulaties zien dat deze methode superieur is aan een klasse andere methoden waarmee hetzelfde beoogd wordt. Deze superioriteit is het gevolg van het feit dat fysische wetmatigheden gebruikt zijn om de resultaten van de methode te vormen. In zijn algemeenheid is het erg krachtig om fysica te gebruiken om de verscheidenheid aan mogelijke uitkomsten van een proces in te perken tot alleen de plausibele.

Nadat duidelijk is dat deze krachtige nieuwe methode werkt op gesimuleerde data, passen we haar in Hoofdstuk 3 voor het eerst toe op echte data. In dit geval gebruiken we een achturige LOFAR-observatie, waarvan het beeld in eerste instantie grote calibratiefouten bevatte. Prettig genoeg constateren we dat onze methode in staat is om deze beeldfouten weg te werken, zodat we een stuk duidelijker beeld krijgen van dit kleine stukje hemel.

Logischerwijs willen we weten of dit succes een gelukje betrof, en of de methode robuust is. Daarom passen we in Hoofdstuk 4 dezelfde methode toe op twaalf achturige LOFAR-observaties, die allemaal gericht zijn op een deel van het sterrenbeeld Ursa Major (Grote Beer). Dit gebied staat bekend als het Lockmangat, omdat het erg weinig HI-emissie vertoont in de Galactische voorgrond. In dit geval blijkt er een algemene verbetering te zijn in tien van de twaalf gevallen, maar ook wordt duidelijk dat er een kleine systematische fout in de data zit waar we in eerste instantie niet van op de hoogte waren. Deze systematische fout blijkt erg moeilijk te kwantificeren.

Onze methode heeft groot initieel succes getoond en er zijn duidelijke vervolgstappen om haar te verbeteren en voornoemde systematische fout aan te pakken. Dit proces is typisch voor de wetenschap. Als methodes beter worden, neemt de onzekerheid in de resultaten af

en als gevolg wordt het bestaan van subtielere fenomenen blootgelegd. Vaak is dit de grens waar ontdekkingen worden gedaan en vooruitgang geboekt wordt.

Curriculum Vitae

Joshua George Albert was born in London, Ontario, Canada on 27 April, 1990 and grew up in the small town of Ingersoll, Ontario. In June 2008 he graduated from Ingersoll District Collegiate Institute. He had applied for an engineering degree at University of Toronto and was denied, however he was offered to enter the mathematics and physics programs. In September 2008 Josh began an Honours Bachelor of Science in Physics at the University of Toronto Scarborough Campus, and graduated in June, 2013. In the summer of 2010 – 2012 he undertook extracurricular projects which expanded his interests. The most formative research project was in the summer of 2011 when he spent two months at the Giant Metrewave Radio Telescope, 80km north of Pune, India, where he was first exposed to radio astronomy research. During this time his work on radio frequency interference mitigation earned him his first co-authorship. In the fifth year of his Bachelor he took a number of auxiliary graduate level physics courses,

which lead him to choose astronomy as an academic path, and ultimately to apply to Leiden University for a Master of Science in astronomy. In September 2013 Josh began his Master of astronomy at Leiden University. His first Master thesis was on the topic of radio astronomy and galaxy clusters, and his second Master thesis was on the topic of predicting the future detection of hyper-velocity stars with Gaia. In August 2015 he graduated from the M.Sc. and began a Ph.D. at Leiden Observatory. During this time he focused on the problem of ionosphere calibration in radio interferometric data. An updated list of publications can be found at <https://orcid.org/0000-0001-5327-3705>.



List of Publications

- J. G. Albert, R. J. van Weeren, H. T. Intema, and H. J. A. Röttgering. On the feasibility of probabilistic direction-dependent ionospheric screens for LoTSS. (*in preparation*).
- J. G. Albert, C. Sifón, A. Stroe, F. Mernier, H. T. Intema, H. J. A. Röttgering, and G. Brunetti. Complex diffuse emission in the $z = 0.52$ cluster PLCK G004.5-19.5. *A&A*, 607:A4, October 2017. doi: 10.1051/0004-6361/201730496.
- J. G. Albert, M. S. S. L. Oei, R. J. van Weeren, H. T. Intema, and H. J. A. Röttgering. A probabilistic approach to direction-dependent ionospheric calibration. *A&A*, 633:A77, Jan 2020a. doi: 10.1051/0004-6361/201935668.
- J. G. Albert, R. J. van Weeren, H. T. Intema, and H. J. A. Röttgering. Probabilistic direction-dependent ionospheric calibration for LOFAR-HBA. *A&A*, 635:A147, March 2020b. doi: 10.1051/0004-6361/201937424.
- T. Marchetti, O. Contigiani, E. M. Rossi, J. G. Albert, A. G. A. Brown, and A. Sesana. Predicting the hypervelocity star population in Gaia. *MNRAS*, 476:4697–4712, June 2018. doi: 10.1093/mnras/sty579.
- Gregory Paciga, Joshua G. Albert, Kevin Bandura, Tzu-Ching Chang, Yashwant Gupta, Christopher Hirata, Julia Odegova, Ue-Li Pen, Jeffrey B. Peterson, Jayanta Roy, J. Richard Shaw, Kris Sigurdson, and Tabitha Voytek. A simulation-calibrated limit on the H I power spectrum from the GMRT Epoch of Reionization experiment. *MNRAS*, 433(1):639–647, July 2013. doi: 10.1093/mnras/stt753.

Acknowledgements

To all the people who were there along the way. I had hoped that when I started my Ph.D. it would be with a great perspective, and that it would allow me to tackle whatever problem lay ahead with independence. Instead, what I discovered, and what I should have guessed, is that with each stone turned over, a new mystery awoke. I grew to enjoy collaboration as much as independence. Not only that but I was inspired by both friends and colleagues.

Kim, Jit, Alex, my academic siblings, thank you for your companionship and shared struggle. In fact, to all of the Sterrewacht, thank you for providing a great place to work. There is a strong sense of community at the STRW, a place of friendly faces and good conversation. Where one can find intellect, anecdote, and jokes. It's a place one wants to work (except in the summer months in the Oort building because there's no A/C) and that is something that I'll take with me when I create work spaces of my own.

Next to my Ph.D. was, most importantly, the support, love, and patience of my friends and family. My tribe in all honesty. You are all so smart, beautiful, happy, and funny, and I'm so lucky to have you.

Firstly, and singularly, Mark, thank you so much for being my brother. I wouldn't have done it without you and I'm looking forward to the rest. Tyler Clark, thank you for your patience and friendship, and our deep conversations and journey together. Tyler Sage, thank you for the camaraderie, late night walks, words, and hugs. Ron and Nisa thank you for your precious spirit and laughter, and the special times that will last forever. Roy Kamienchik, thank you for your understanding and friendship, and your confidence during the tough times. Mikko, thank you for our times front and centre, and also the silent times just taking in the beauty. Delivis, thank you for your pure individuality and strength, and the care you entreat your family. Roy Klein, thank you for your kind soul, which anyone can see. Martijn, thank you for the sincerity and intensity you give to all your thoughts and relationships. You are also among the most positive of people, and I'm so happy to have you as my friend. Lolita, thank you for your creative and gentle spirit, and the love you bring Mark. Marije, thank you for your eccentricity and the love you bring Tyler Sage. Daria, thank you for your beautiful soul and perspective on life. Peter, thank you for your giving in the truest sense. You are one of the kindest souls I know. Koen, thank you for your sensitivity and creativity, and for all the times spent either playing music together, or just talking. Rasmus and Katya, thank you for the care and dedication you gave us in teaching us to dance. Tyron, and in fact all my gym and yoga friends (you know who you are), thank you for sharing my mental sanctuary and giving me motivation. Rik, thank you for all the attention you give to the small things, and Nina for all your overflowing creativity and strength. Ramsay, thank you for the joy you bring to the life of everyone around you. It's a serious special talent of yours, and I hope you've

rubbed off on me. Cas, thank you for your sweet positivity in life, from the guitar playing well past midnight to just cooking together. Jacob, dear buddy, thank you for your determination and kindness and for all the rumpus times together.

Jane, mom, thank you for teaching me kindness and creativity (yes I learned that), and about the aspect of eternal love. Denis, dad, thank you for teaching me the rules of life, and for my inspiration for how things work (it all came from you). Ainsley and Bailey, my beautiful twin sisters, I'm so thankful to have you and your smiles. To have your trust, confidence, and courage. Ava, my niece, for letting me be the first one to make you laugh. Joanne and Ronald, grandma and grandpa, thank you for the sense of wonder you bestowed in me while growing up. From gifts and stories of your travels to magical Christmases and to our discussions on life and existence. Thank you for your everlasting support. Natalia and Pawel, mom and dad of my love, thank you for raising a gem and for giving me a warm place from where I finished this thesis.

Now, as with author lists in academia, the order of all names except the first and last are quite meaningless. Lastly, I'd like to thank Aleksandrina, Alina, my fate and love, for everything. You've made the last *schrikkeljaar*¹ unforgettable and I look forward to the rest.

¹Leap year in Dutch.

