

Universiteit
Leiden
The Netherlands

Big simulations for big problems

Kugel, R.H.

Citation
Kugel, R. H. (2025, January 16). *Big simulations for big problems*. Retrieved from https://hdl.handle.net/1887/4176251

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/1887/4176251

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

Big simulations for big problems

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr.ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op donderdag 16 januari 2025
klokke 10:00 uur

door

Roi Kugel

geboren te Assen, Nederland
in 1998

Promotor:
Prof.dr. J. Schaye

Co-promotor:
Dr. M. Schaller

Promotiecommissie:

Dr. S. Bocquet	(Ludwig-Maximilians-Universität München)
Dr. E. Bulbul	(Max Planck Institute for Extraterrestrial Physics)
Prof.dr. H. Hoekstra	
Prof.dr. K. Kuijken	
Prof.dr. I. Snellen	

Copyright © *2024, Roi Kugel*

ISBN: 978-94-6496-313-7

Cover design by Annely Arends

Printed by: Gildeprint



Universiteit Leiden

Contents

1 Introduction **1**
1.1 The cosmic web and ΛCDM . . . 1
1.1.1 The cosmic web . . . 3
1.1.2 Towards a standard model . . . 4
1.1.3 ΛCDM . . . 4
1.1.4 The Cosmic Microwave Background . . . 6
1.1.5 Galaxy surveys and the S_8 tension . . . 7
1.2 Models for cosmology . . . 10
1.2.1 Initial conditions . . . 10
1.2.2 Halo models . . . 11
1.2.3 Dark matter only simulations . . . 12
1.2.4 Populating dark halos . . . 14
1.2.5 Emulators for cosmology . . . 14
1.2.6 Cosmological hydrodynamical simulations . . . 15
1.2.7 Radiative cooling and star formation in hydro simulations . 17
1.2.8 Feedback in hydro simulations . . . 19
1.2.9 Calibration of hydro simulations . . . 22
1.2.10 Baryons and cosmology . . . 25
1.2.11 The FLAMINGO simulations . . . 25
1.3 Galaxy Cluster Cosmology . . . 27
1.3.1 Current constraints . . . 29
1.3.2 Selection effects . . . 31
1.3.3 Models for cluster cosmology . . . 32
1.3.4 Future surveys . . . 35
1.4 This thesis . . . 35
1.5 Outlook . . . 36

References **38**

2 Calibration using emulators **43**
2.1 Introduction . . . 44
2.2 Simulations . . . 47
2.2.1 Stellar feedback . . . 51
2.2.2 Black hole growth . . . 52

2.2.3 AGN feedback . . . 53
2.3 Observational data and biases . . . 59
2.3.1 The galaxy stellar mass function . . . 59
2.3.2 The cluster gas mass fractions . . . 64
2.4 Emulator construction . . . 67
2.4.1 Training sets . . . 67
2.4.2 Obtaining the required simulation output . . . 68
2.4.3 Training using Gaussian processes . . . 70
2.4.4 Error estimation . . . 71
2.5 Using the emulator for parameter estimation . . . 74
2.6 Results . . . 77
2.6.1 Parameter sweeps . . . 77
2.6.2 The best-fitting intermediate-resolution model . . . 79
2.6.3 The best-fitting subgrid high- and low-resolution models . 81
2.6.4 Feedback variations . . . 84
2.7 Conclusions . . . 89

References **93**

Appendix **97**
2.A Different apertures . . . 97
2.B Posteriors for high- and low-resolution . . . 97
2.C Parameter ranges for the AGN Jet model . . . 99
2.D swift-emulator: A Python package for emulation of simulated scaling relations - Roi kugel & Josh Borrow . . . 99
2.D.1 Summary . . . 99
2.D.2 Statement of need . . . 99
2.D.3 Emulator Requirements . . . 102
2.D.4 Why swift-emulator? . . . 103
2.D.5 Acknowledgements . . . 104

3 FLAMINGO: Cluster selection effects **107**
3.1 Introduction . . . 107
3.2 Methods . . . 111
3.2.1 FLAMINGO . . . 111
3.2.2 Observables used for selection . . . 113
3.2.3 Sample Selection . . . 115
3.3 Results . . . 115
3.3.1 Scatter at fixed mass . . . 118
3.3.2 Correlations between cluster properties . . . 119
3.3.3 Characteristic mass as a function of the cut in observable space . . . 120
3.3.4 Selection at fixed comoving number density . . . 121
3.3.5 Sample mass bias as a function of the selection limit . . . 124

3.3.6 Sample mass bias as a function of the target mass limit . . 126
3.3.7 The effect of modelling uncertainty 131
3.3.8 Biases in properties other than mass 134
3.4 Conclusions 137

References **140**

Appendix **143**
3.A Fits at different redshifts 143
3.B Using Compton-Y within R_{500c} 145
3.C Convergence with numerical resolution and simulation box size . 145

4 Theoretical uncertainties in cluster counts **151**
4.1 Introduction 151
4.2 Methods 156
4.2.1 FLAMINGO 156
4.2.2 The cluster count model 157
4.2.3 Definition of samples 162
4.2.4 The likelihood 163
4.3 Results 164
4.3.1 DMO Halo mass function models 165
4.3.2 Baryonic effects on the halo mass function 167
4.3.3 Fit to the observable-mass scaling relation 169
4.3.4 Systematic uncertainty in the observable-mass scaling relation 174
4.3.5 Quantifying the biases in the number counts 176
4.3.6 Biases in cosmological parameters 179
4.4 Conclusions 181

References **187**

5 Cosmic Web Dynamics: Forces and Strains **191**
5.1 Introduction 191
5.1.1 The Cosmic Web: structure and detection 192
5.1.2 The Cosmic Web: dynamics, formation and evolution . . . 196
5.1.3 Morphology of the Cosmic Web 198
5.1.4 Cosmic Web Dynamical Inventory 200
5.1.5 Outline 201
5.2 Data and Analysis: N-body simulations 201
5.2.1 N-body Simulation 202
5.2.2 DTFE Density and Velocity maps 202
5.2.3 MMF/NEXUS+ 202
5.3 Force and Strain Field: Procedure 206
5.3.1 Gravitational Force field 207
5.3.2 Tidal field 209

5.4 Gravitational Force Field 211
5.4.1 Filaments 214
5.4.2 Voids 215
5.4.3 Walls 215
5.4.4 Cluster Nodes 216
5.4.5 Gravity Field strength: statistical analysis 216
5.5 Gravitational Force Field: individual structures 221
5.5.1 Case study 1: void force field 221
5.5.2 Case study 2: cluster node force field & flowlines 225
5.6 Tidal field 228
5.6.1 Tidal Field illustrated: tidal impact voids 228
5.6.2 Tidal Field Structure 232
5.6.3 Tidal field strength: statistical analysis 233
5.6.4 Tidal field alignment 237
5.7 Cosmic Web Force & Tidal Evolution 240
5.8 Summary and Discussion 242

References **249**

Appendix **253**
5.A MMF/Nexus 253
5.A.1 Hessian Geometry and Morphological Identity 253
5.A.2 Scale Space and Multiscale Structure 254
5.A.3 Signature & Versions 254
5.A.4 NEXUS+ 255

Nederlandse samenvatting **257**
Introductie 257
Dit proefschrift 259

Publications **263**

Curriculum Vitae **267**

Acknowledgements **269**

1 Introduction

At first there was something we will probably never know. Then, something happened which we might never understand. This led to the creation of the Universe as we know it. The first thing our Universe did was to expand, a process that appears to still be ongoing. The study of how the cosmos as a whole is acting is the topic of the field of cosmology.

Over the history of mankind many attempts have been made to describe our Universe. However, when we talk about cosmology as a science, it starts with Einstein's formulation of general relativity (Einstein, 1916; Einstein, 1917). Specifically, Einstein's equations can be solved in the special case of the cosmological principle. The cosmological principle states that, at the largest scales, the Universe is homogeneous and isotropic. If you assume a homogeneous and isotropic density distribution, Einstein's equations simplify to give expanding or contracting solutions. In these universes the metric that describes space-time takes the form of the Friedmann-Lemaître-Robertson-Walker metric (Friedmann, 1922, 1924; Lemaitre, 1931; Robertson, 1935, 1936; Walker, 1937)

$$-\mathrm{ds}^2 = -\mathrm{c}^2\mathrm{dt}^2 + \mathrm{a(t)}^2(\mathrm{dr}^2 + \mathrm{r}^2\mathrm{d}\Omega^2) \tag{1.1}$$

where $a(t)$ is the expansion factor of the universe. The validity of this equation, and its prediction of a universe that is expanding was validated by the observations of Hubble (1929). As noted by Lemaître, the expansion of the universe can be traced back to a singular point in time, where the Universe emerged from a singularity. This "Big Bang" lies at the start of our Universe. With current theories we can explain most of the processes the Big Bang (and inflation). However, the Big Bang and the processes that precede it are likely to remain a mystery for a very long time, and possibly forever.

1.1 The cosmic web and ΛCDM

As demanded by the cosmological principles, according to the Friedmann equations, the Universe is perfectly isotropic and homogeneous. This is true on the very largest scales. However, the small inhomogeneities originating from quantum fluctuations that are grown to cosmic proportions by inflation form the seed of all the structures that we see in the Universe today. With the passing of time

1

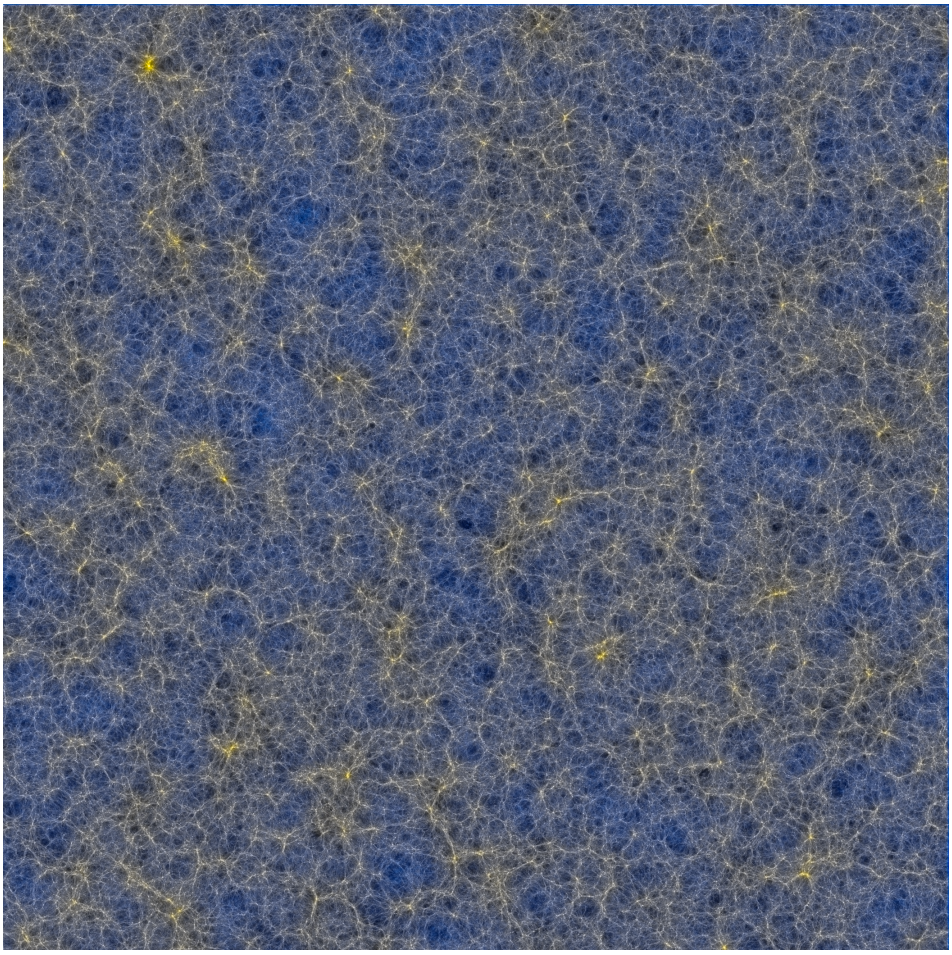


Figure 1.1: The dark matter distribution of the universe as predicted from the FLAMINGO simulations. The intensity shows the dark matter density at each point, while the colour shows the neutrino density at each point. The prominent connected structure is what is called the cosmic web.

these small overdensities grow to become more and more dense. At first this happens linearly, the density fluctuations growing with the same linear factor on all scales, until eventually they undergo non-linear collapse and become virialized objects, which we refer to as haloes. It is only on the scale of haloes that non-gravitational effects start playing a big role. Large scale structure formation can be described almost completely using gravity only.

1.1.1 The cosmic web

At intermediate scales, in between uniformity and virialisation, a particularly striking pattern emerges: the cosmic web (Bond et al., 1996). The cosmic web is an interconnected multi-scale structure that permeates the galaxy distribution. The cosmic web structure is shown for the FLAMINGO simulation in Fig. 1.1. The structure of a large scale cosmic web can be predicted from first order perturbation theory (Zel'Dovich, 1970) on the Lagrangian hydrodynamic equations in an expanding universe. The prominent structures of the cosmic web originate from the different axes along which gravity can lead to collapse. Collapse along all three axes leads to dense nodes, collapse along two axes leads to extended filaments, collapse along a single axis leads to flattened walls and expanding regions become empty voids. Much like haloes, these individual structures form hierarchically, and nodes, filaments, walls and voids can all merge and form bigger structures. As the cosmic web is both the first structure to emerge and the structure within which galaxies are distributed, this makes it an interesting field of study.

The cosmic web emerges as the result of gravitational collapse in a near uniform medium. Because this process happens on many different scales simultaneously, it gives the cosmic web somewhat of a fractal morphology, only broken at the point where haloes start virialising. This multi-scale nature of the cosmic web makes it naturally difficult to quantify and identify, as there are no a priori set boundaries between each different part of it, and each individual element of the it can appear at a multitude of different scales.

To identify the cosmic web, there is a range of methods available that vary greatly in how they distinguish the different components. Because the cosmic web is by nature multi-scale, it comes naturally to define the components of the cosmic web morphologically or according to the local dynamics. The morphology of the density field is what is used for methods like the multi-scale morphology finder (MMF) and NEXUS(+) (Aragón-Calvo et al., 2010; Cautun et al., 2013). In these methods the morphology of the density field is determined by using the Hessian or tidal tensor. The sign of the eigenvalues of this tensor holds information about the local morphology, showing whether there has been collapse along one of the axes. This processes leads to a morphological identification for the Hessian of the density field, and a dynamical identification for the tidal tensor. NEXUS+ further pre-processes the density field by making use of a multi-scale, log-space filtering method to bring out the structures at all scales, leading to a detailed rendering of the cosmic web. One alternative to these methods is to make use of the topology of the field (see e.g. Sousbie, 2013; Bermejo et al., 2024) but these are not used in this thesis.

By making use of the method NEXUS+, Cautun et al. (2014) make an inventory of the properties of the cosmic web and of the haloes that reside in it. In

1

the Universe, about half the mass is in cosmic filaments, while voids make up most of the cosmic volume at about 77%. The nodes take up less than 0.1% of the volume but still make up about 11% of the mass. Nearly all halos with a mass higher than M_{200c}[1] $> 10^{13.5}$ $M_\odot$ are found only in the nodes of the cosmic web. The different elements will also create different morphologically distinct tidal and force fields, that can, for example, impact the orientation of halos (Jones & van de Weygaert, 2009; Ganeshaiah Veena et al., 2019, 2021). A more complete understanding of the cosmic web will further inform us about both cosmology and galaxy formation.

1.1.2 Towards a standard model

The existence of a cosmic web has been verified observationally. Some of the earliest confirmations came in the form of the "Stickman" found in the Centre of Astrophysics (CfA) redshift survey (de Lapparent et al., 1986), shown in Fig 1.2. This slice of the Universe shows a collection of filaments connected to a cluster in the center in the slice. The image also shows the "fingers of god" redshift distortions, where objects with a high velocity dispersion, like clusters, are stretched out along the viewing direction.

At around the same time, the first large-scale structure formation simulations where being done (Davis et al., 1985; Frenk et al., 1985; White et al., 1987). These simulations where predicting structures similar to what was found in observations. At the time, cosmologies with $\Omega_m \neq 1$, let alone with a cosmological constant, where still quite non-standard, and only a limited exploration was done of different models. However, in comparisons with the redshift surveys (Efstathiou et al., 1990) and galaxy clusters (Eke et al., 1996) it already seemed that a universe with $\Omega \approx 0.3$ matched the data better. Together with observations from galaxy rotation curves (Rubin et al., 1980) indicating that there was potentially a large dark component of the matter density. The field was moving towards a shift in its standard model. The observation that many see as the establishment of the new paradigm are the the results from the supernova cosmology project (Perlmutter et al., 1999). By making use of the supernova type 1a standard candle, they made a definite detection of accelerated expansion driven by a cosmological constant Λ that accounted for $\Omega_\Lambda \approx 0.7$ of the energy density of the universe. This paved the way for our current standard model of cosmology: ΛCDM.

1.1.3 ΛCDM

Over two decades later, ΛCDM is still the gold standard of cosmology. As a model, ΛCDM indicates a universe with a flat geometry, that contains both a

[1] M_{200c} is defined as the mass inside a spherical aperture where the average matter density is 200 times the critical density of the universe.

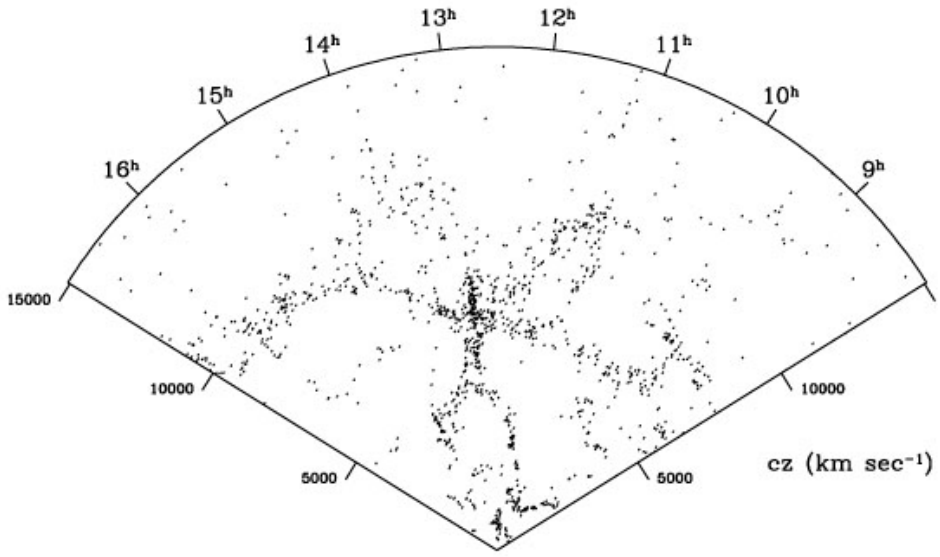


Figure 1.2: Early results from the CfA redshift survey, showing the first clear rendering of the cosmic web. The image is often referred to as "the stickman", due to the appearance of the redshift-distorted cluster in the center of the frame. Image taken from de Lapparent et al. (1986).

cosmological constant and a cold dark component of matter that only interact via gravity. In ΛCDM the early universe is characterised by a period of rapid inflation. During this period of inflation, the seeds of structure are planted via quantum fluctuation and the universe becomes geometrically flat. In its simplest form ΛCDM can predict structure formation along the entirety of cosmic time, according to six input cosmological parameters

- $\Omega_c h^2$ - The cold dark matter content in the universe.
- $\Omega_b h^2$ - The baryonic matter content in the universe.
- t_0 - The age of the universe.
- n_s - The power-law scalar index of the power spectrum of the primordial fluctuation field.
- Δ^2_R - The amplitude of curvature fluctuations.
- τ - The optical depth to reionisation.

From these six the other more commonly used parameters can be directly derived. For example the Hubble constant H_0, which describes the expansion rate

1

of the universe, and σ_8 which describes the level of matter fluctuations on a scale of $8h^{-1}$ Mpc. With just these parameters, most of the last two decades of observations can be accurately modelled and predicted. Up to this point, there has seemed to be only one very clear downside to ΛCDM. None of the three pillars of the model, dark energy, dark matter and inflation, currently have a clear physical mechanism that drives their existence. In essence, this makes ΛCDM mostly an empirical model.

In base ΛCDM there are a number of parameters that are kept constant at their most natural values. However, there are certain extensions that can be made to the model that are self-contained with ΛCDM. Some examples are the effective number of neutrinos species N_{eff} and the assumption of a flat universe. Both of these have been independently verified to be consistent with their natural values. There are some extensions of ΛCDM that add new processes to it not included in the base model. One example of this is the addition of massive neutrinos which we know exist, but are not part of standard ΛCDM. Another example is to add a parametric equation of state for dark energy

$$w_{\text{DE}} = w_0 + (1-a)w_a, \tag{1.2}$$

where $w_{\text{DE}} = \frac{P_{\text{DE}}}{\rho_{DE}}$ relates the pressure and density of dark energy and a is the expension factor. For this addition, the extension is more like a perturbation. This extension is sometimes refered to as $w_a w_0$CDM. Finding a significant deviation from $(w_0, w_a) = (-1, 0)$, the values which correspond to a cosmological constant, would indicate additional limitations within base ΛCDM.

1.1.4 The Cosmic Microwave Background

One of the biggest successes of the hot Big Bang model is the prediction of a cosmic background radiation. Just after the Big Bang the universe is extremely hot and dense, and light is unable to free stream over large distances. Once the universe cools down enough for the hydrogen to become neutral, the period referred to as the era of recombination, the photons that were trapped in the Big-Bang plasma are able to free stream for the first time. The radiation from these photons is still detectable today in the Cosmic Microwave Background (CMB). The CMB gives us a window to how the universe looked like about 300.000 years after the Big Bang. The accurate measurements of this radiation are one of the main pillars of modern cosmology.

The original discovery of the CMB by Penzias & Wilson (1965) went paired with its description by Dicke et al. (1965). Since then, multiple satellites like WMAP (Komatsu et al., 2003) and Planck (Planck Collaboration et al., 2020a), and also ground based observatories like SPT (Story et al., 2013) and ACT (Sievers et al., 2013) have revolutionised the level of accuracy with which we can constrain cosmological parameters. The Planck cosmological parameters are measured at

better than per cent level accuracy. Because of the robustness of the modelling of the CMB fluctuation signal, the consistency in the fitting of cosmological parameters, and the high sensitivity of the signal to the cosmological parameters, the results from the Planck satellite form the reference model of cosmology.

Besides the primary CMB, in other words, the angular power spectrum of primordial fluctuations, the radiation of the CMB holds a wealth of other probes of cosmology and structure formation. Of particular interest are the Sunyaev-Zel'Dovich (SZ) effect and weak lensing of the CMB. The SZ effect is caused by CMB photons interacting with high energy electrons between the time of emission and the time of measuring. This leads to a distinct spectral feature in the CMB. As the effects are caused by high energy electrons, the SZ effect is stronger for path-lines that cross galaxy groups and clusters, where there is a high temperature electron plasma. The use of the SZ effect for galaxy cluster cosmology will be covered later in this chapter. Besides cluster-finding, the SZ effect can also be mapped over the entire sky (Planck Collaboration et al., 2014; Bleem et al., 2022). When looking at the power spectrum of these maps, there is a (slight) internal tension within the Planck results (McCarthy et al., 2018, 2023). At large scales, simulations predict SZ to have higher power at matched cosmology, and a lower σ_8 would partially resolve the tension. At small scales, there is a general mismatch between the power spectrum and any predictions. For both scales there is a likelihood that this is due to systematics in dealing with foreground removal, which is highly non-trivial for SZ. For example a large shift was seen with the Planck results between the 2013 and 2015 releases of the SZ signal maps. Such an inconsistency is not found for the Planck lensing results (Planck Collaboration et al., 2020b). Even though CMB lensing also probes the structures along the line of sight, the cosmological inference is perfectly consistent with the Planck cosmology, though it is sensitive to higher redshift than SZ.

1.1.5 Galaxy surveys and the S_8 tension

When investigating the large-scale structure of the Universe, we are limited to matter that emits light (or potentially gravitational waves or non-photon cosmic rays, but those are outside the scope of this thesis.) Because of this limitation, our main probes of structure are galaxies. Galaxies are the lighthouses that populate the large-scale structure. For cosmology, the distribution of galaxies holds information on the large-scale structure and hence can be used to constrain cosmological parameters. Much like the CMB, a lot of the information is contained in the first order clustering, measured using the correlation function.

When it comes to measuring galaxies for cosmology, there are a few approaches that can be taken. In the simplest form, galaxies can be counted as a function of their position on the sky to measure their angular correlations. If the redshift is also measured, 3D correlation function analyses are possible. Because

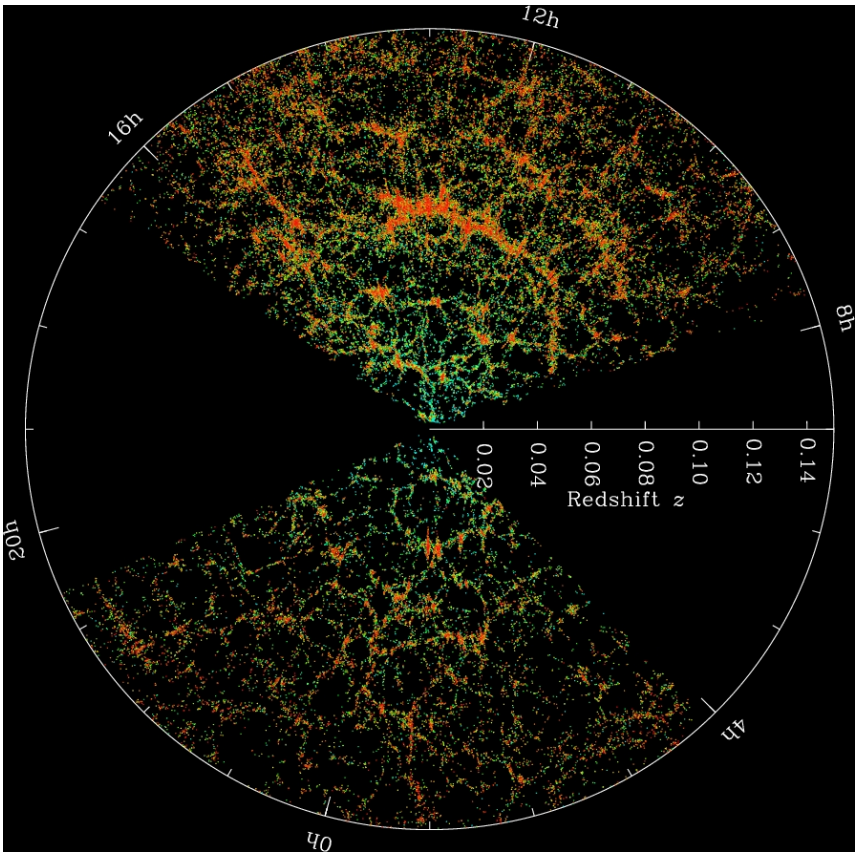


Figure 1.3: SDSS Galaxy distribution. Each dot represents a galaxy, color indicates high number density. Image credit: M. Blanton and SDSS

of the peculiar velocities of the galaxies, the redshifts will give a distorted view of the large-scale structure. However, the redshift-space distortions themselves also hold cosmological information and can be used for cosmological inference. Obtaining spectra for every galaxy is not a trivial matter, and any galaxy survey that wants to use redshifts needs to be specifically designed to also measure them. Therefore many surveys now also rely on photometric redshifts, where a number of observing bands are used to give a rough estimate of the redshifts. This is particularly useful for weak lensing surveys. To measure week lensing, the most important observable is the shapes of galaxies, as their deformation gives statistical information about the matter distribution along the line of sight. By making use of photometric redshifts, the matter field can be analysed in a few tomographic bins. By studying lensing as a function of redshift, we can obtain additional cosmological constraints.

Major constraints on the baryonic acoustic oscillations, peaks in clustering at certain scales caused by acoustic waves in the primordial matter field, using galaxies are obtained with the Sloan Digital Sky Survey (SDSS), Baryon Oscillation Spectroscopic Survey (BOSS) (Dawson et al., 2013; Ross et al., 2017; Icaza-Lizaola et al., 2020) shown in Figure 1.3. BOSS provides some of the tightest constraints at the redshifts they are sensitive to. The BOSS results are also in agreement with the results from the Planck satellite. When it comes to spectroscopic surveys, the Dark Energy Spectroscopic Instrument (DESI) has also started reporting their first cosmological results (DESI Collaboration et al., 2024). In their initial data release they report finding a 3.9σ detection of a $waw0\Lambda$CDM signal when they combine their results with the results from the Planck satellite. It is however likely that this is driven by some of their $z_{\rm eff} = 0.5$ data bin, which might prove to be a statistical fluke, as the results from BOSS don't agree in this redshift range.

When it comes to galaxy surveys, many current and upcoming surveys have largely been designed to measure cosmology trough weak lensing. In general relativity, light-rays travel along geodesics, which are straight only when spacetime is flat. However, as spacetime can be curved, large gravitational bodies alter the the pathways of photons and distort the images of objects whose light passes trough their potential. For deep potentials, this usually leads to effects similar to light passing trough a lens, hence the name. The most massive galaxy clusters act as strong lenses, where the background light is heavily distorted. However, for large galaxy surveys, most objects are in the weak lensing regime, either around less massive objects, or at larger distances from the most massive clusters. In the weak lensing regime, the distortions of objects is too weak to pick up. Instead, what is measured is the statistical distortion due to the lensing of many galaxies. The signal is measured from sky-position correlated distortions of the shapes of a large sample of galaxies.

Many large galaxy surveys have measured the weak lensing signal, with interesting results. The results from the KiDS survey (Heymans et al., 2021) show a tension with the results from the Planck survey. The the value found for the clustering parameter σ_8 is $\approx 3\sigma$ lower than the Planck results. Results from the HSC surveys (Miyatake et al., 2023) find a 2σ tension and the results from the Dark Energy Survey (Abbott et al., 2022) are at about 1.5σ compared with Planck. This has led to what is now called the S_8 tension, named after the parameter S_8, which is the combination of σ_8 and $\Omega_{\rm m}$ that is maximally constrained by lensing surveys. The difference is relatively small, however, as weak lensing also probes the smaller scales, there is also astrophysical information to be learned. In order to further investigate the S_8 tension, we need further advances in the models we use to both compare the observations to and measure cosmology from.

1

1.2 Models for cosmology

In order to test our understanding of galaxy formation and cosmology, there is a plethora of models with which the data can be compared. These models are the backbone of any cosmological inference. With advances in both the depth and the observing area of cosmological surveys, it has become imperative that models keep up with the high accuracy demanded by the observations. Stage 4 surveys like LSST, Euclid, Nancy Roman, Simons observatory and CMB-S4 will demand accuracy in the power spectrum at the level of less than a percent. This level of precision not only makes accurate modelling a requirement, but processes like neutrinos and baryonic physics now need to be accounted for and modelled accurately. In this section different methods of modelling are discussed, starting from analytic models and leading to more and more complex models. The details of preparing these models for the high-accuracy constraints of upcoming surveys will be discussed where applicable.

1.2.1 Initial conditions

In order to initialise our cosmological models, we need to specify initial conditions. Within ΛCDM, the initial density field is described by a Gaussian random field, defined only by its power spectrum with random phases. The fluctuations in this random field originate from quantum fluctuations that were stretched to cosmic scales via inflation (Guth, 1981; Linde, 1982). While deviations from Gaussianity are one of the probes of deviations from standard ΛCDM, our current constraints are still consistent with a perfectly Gaussian random field (Planck Collaboration et al., 2020c).

Because the initial conditions are Gaussian, the main ingredient that needs to be specified is the matter power spectrum. In ΛCDM, the equations of the early universe can be fully specified, and then calculated numerically. This initial matter power spectrum is often referred to as the linear matter power spectrum. As structure formation happens power first grows linearly across all scales, hence the name, but as structures formation evolves it becomes non-linear and starts forming virialised structures. To calculate the linear power spectrum we make use of Boltzmann solvers that solve for the equations of the early universe. Two of the most widely used packages are CAMB (Lewis et al., 2000; Howlett et al., 2012) and CLASS (Lesgourgues, 2011a,b).

The linear matter field quickly evolves into the non-linear regime, for which the initial conditions become slightly more complex. For Lagrangian methods, the first proper initial conditions generation algorithm comes from the work by (Zel'Dovich, 1970). From the initial density field, the motions of a set of particles can be described via ballistic motion. This method is known as the Zel'Dovich approximation and can also be used to evolve the matter field to the current

time. However, the method becomes inaccurate once particles start crossing each other's paths. The Zel'Dovich approximation is the first order of Lagrangian perturbation theory applied to the cosmic density field. Since Zel'Dovich there have been developments that allow for the generation of nested random phases (Jenkins, 2013), that also make it easier to re-simulate smaller parts of the simulations at higher resolution. Initial conditions can now also be initiated to 2nd (Hahn & Abel, 2011) and 3rd order (Hahn et al., 2021).

Another major improvement for initial conditions is that there is no longer a need for "back-scaling" the power spectrum. When back-scaling, agreement between simulations and the linear power spectrum is guaranteed by first calculating the z=0 linear power spectrum with CLASS or CAMB, and scaling this power spectrum back to the starting redshift. With codes like monofonIC (Hahn et al., 2021), no longer use this approach. This makes the biggest difference at intermediate redshifts, as it leads to the proper evolution. Additionally, for simulations with gas and dark matter, monofonIC treats the different matter fields properly, and gives a different inital power spectrum to gas and dark matter. The use of higher order Lagrangian perturbation theory also allows the ICs to be evolved to lower redshifts, which reduces the noise incurred in numerical simulations at high redshifts. Extensions to the base ΛCDM model also have to be added to the ICs to achieve high precision, like for example massive neutrinos (Elbers et al., 2022).

1.2.2 Halo models

Now that we can generate initial conditions, let's start with an analytic formulation of structure formation: the halo model (For a recent review see Asgari et al., 2023). The halo model is a useful model to discuss as it nicely splits up the problems into smaller parts, which will make some of the choices made for the more complex models easier to understand. The halo model makes predictions for the matter power spectrum and is able to predict the clustering of matter in halos and subhalos with only a few ingredients: The linear matter power-spectrum, the halo mass function, the satellite number density profile and the halo occupation distribution.

The halo model splits the calculation into two terms: the two halo term and the one halo term. The two halo term describes the clustering between central halos. This thus describes clustering on large scales. The one halo term describes clustering within a single halo, between the central and its satellites. This prescription assumes that matter beyond the virial radius of halo clustering can be described via linear theory, while all matter inside haloes is described by and NFW profile. At both large and small scales the halo model can be very accurate. However, at intermediate scales, where neither the one or two halo terms dominates, it can become very inaccurate. Current halo models specifically add

1

correction terms that are fitted to simulations to decrease the uncertainty in the intermediate regime. This allows model like HMCode (Mead et al., 2021) to reach quite a high accuracy of a few per cent. The flexibility of halo models also make it possible to add beyond ΛCDM physics to the model, like the neutrinos, baryons and differing dark energy equations of state included in HMCode. Halo models are quick and easy to use. However, more complex and complete simulations are always needed to check that the results are accurate.

1.2.3 Dark matter only simulations

Dark matter only simulations make the simplifying assumption that all matter in the universe can be represented as dark matter that only interacts via gravity. This simplifies the calculations. The matter density field is usually represented as dark matter particles, that are evolved in a Lagrangian way. For systems with a small number of bodies gravity can be solved by calculating all the particle-to-particle forces. However, as these calculations scale as $\sim N^2$, this quickly becomes computationally infeasible. To solve this, the long range forces are usually implemented via a particle mesh (Hockney & Eastwood, 1981). Here the particles are binned onto a density grid, the mesh, allowing the forces to be calculated using an fast Fourier transform. This speeds up the calculation as this process scales as $N \log N$. With this method cosmological simulations become possible. This type of simulation was popularised by Davis et al. (1985); Frenk et al. (1985); White et al. (1987). Some of their results are shown in Figure 1.4

Particle mesh was the first step towards quick and efficient calculations of dark-matter simulations. Modern codes, like GreeM (Ishiyama et al., 2009), PKDgrav, (Stadel, 2001; Potter et al., 2017), GADGET-4 (Springel et al., 2021) and SWIFT (Schaller et al., 2023) make use of a mesh for large scale forces, and most opt for a tree-based fast-multiple-moment force calculation for the particle-particle forces at small scales. This approach allows for both the speed and easy periodic boundary conditions of a large scale mesh, and accurate forces on short scales using the fast-multiple-moment method. This way the evolution of the matter field can be calculated both efficiently and accurately.

It is exactly because of their relative ease that dark matter only simulations are widely used. Because all matter is assumed to be dark, these simulations make no direct predictions for galaxy properties besides the properties of the dark matter haloes, which are largely not observable. Dark matter only simulations are often populated with galaxies with different properties. To do this there are various methods that will be described in the next section. Their efficient nature also makes them the perfect candidate to create training sets for emulators that predict the effect of changing cosmologies.

Due to their flexible nature, there are a lot of big dark matter only projects that are still being used. Perhaps the most famous being the Millennium sim-

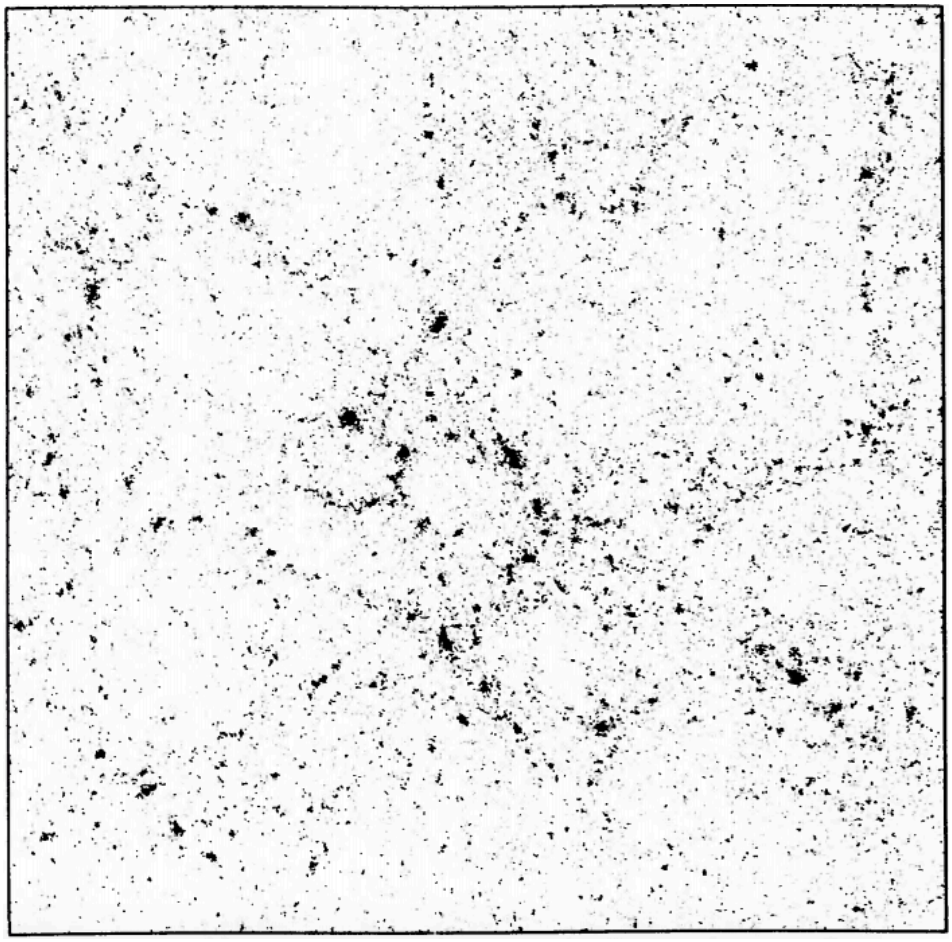


Figure 1.4: Results of a cosmological simulation by the 'Gang of four'. Image from Davis et al. (1985).

ulation (Springel et al., 2005a). The simulation encompassed a $(500h^{-1}\ \mathrm{Mpc})^3$ volume at a resolution of $8.6 \times 10^8\ \mathrm{M}_\odot$ and provides lots of insight into both cosmology and galaxy formation. In the era of precision cosmology, dark matter only simulations are the backbone for the creation of mock surveys for current and upcoming galaxy surveys. The simulations fall into two categories, depending on whether they are meant for large scale mock surveys, or for emulators. In the first category we have simulation like the Euclid flagship (Potter et al., 2017) and Uchuu (宇宙, Universe in Japanese) (Ishiyama et al., 2021). In this case the goal is to simulate an enormous volume. This usually leads to single simulations with a very large particle number. The large volumes are needed to be able to accommodate the depth and survey area of upcoming surveys. The other approach

1

is to run a large number of smaller simulations, like Mira Titan (Heitmann et al., 2016) and AbacusSummit (Maksimova et al., 2021). In this case the different simulations sample cosmological parameter-space probed by upcoming surveys. The final goal of such simulation suites is to provide the training data for cosmological emulators. As mentioned before, we will describe these in more detail later.

1.2.4 Populating dark halos

A lot of the use of dark matter only simulations comes from the models that are attached to the results of these simulations. The final distribution of halos can be populated with galaxies to create mock surveys. This way we can learn more about what is required for the build-up of galaxies. There are various methods for populating dark matter halos.

The simplest of these methods uses a Halo Occupation Distribution (HOD) (Berlind & Weinberg, 2002). The HOD describes the number of satellites halos per host halo, usually as a function of mass. However, these models can be made more complex to account for more complex, non-linear effects in the clustering, see for example the AbacusHOD (Hadzhiyska et al., 2023). While HODs are generally quite simple models, their flexibility makes them ideal for generating mocks. As HODs are applied after the simulation is done, one can quickly iterate over their parameters. For cosmology this means that the HOD parameters can be marginalised over, which may be preferred over assuming an explicit relation.

There are various, more complex, ways to make the connection between halos, subhaloes and galaxies. One of these methods is SubHalo Abundance Matching (SHAM), see for example Moster et al. (2010). For this method the primary assumption is to assume that bigger galaxies reside in bigger subhalos, which include both centrals and satellites. With this assumption, observed distributions of galaxies can be matched to a simulated subhalo distribution, thus populating the subhalos with galaxies. The most complex models fall into the category of semi-analytical models (SAM). These models are physics based, like for example Galform (Cole et al., 2000), L-GALAXIES (Yates et al., 2021) or Dark Sage (Stevens et al., 2023). For these models the full evolutionary histories of haloes is taken into account, and the galactic properties are calculated across cosmic time. SAMs come with free parameters that are tuned to replicate specific observations. Finally there are empirical models like UniverseMachine (Behroozi et al., 2019), where galaxy properties are fit to match observations.

1.2.5 Emulators for cosmology

While running ever larger simulations to account for the large observational volumes is very important for mocks, we need to be able to make very accurate

predictions for a large range of cosmologies, to constrain cosmological parameters with upcoming missions. While halo models can do this cheaply, the extreme accuracy requirements of upcoming surveys favour taking a more agnostic, approach that uses inputs from full simulations. As running full dark-matter-only simulations (let alone hydro simulations) for methods like Markov Chain Monte Carlo (MCMC) is not a viable strategy when it comes to the required computational time, this problem is overcome by making use of the wealth of machine learning literature that has recently become more and more popular. The main advantage for cosmology is that a limited amount of high accuracy simulations can be used to train a machine learning model. The trained model can then make prediction for how different observables respond to changes in cosmology.

To set up the emulators, a suite of simulations is run where the cosmological parameters of interest are efficiently varied to cover the parameter range. Instead of a regular lattice, the parameter space can be covered via more efficient methods like using a Latin hypercube (Sacks et al., 1989; Morris & Mitchell, 1995). In a Latin hypercube each dimension is sampled uniformly, i.e. the same value never appears more than once like in a regular grid. To optimise the volume coverage, the parameter vectors for each sampling is scrambled until an optimal volume filling is found. Using these designs, high accuracy can be obtain with a limited set of simulations. For example, using their own framework described by Heitmann et al. (2008), Heitmann et al. (2009) is able to achieve percent level accuracy when varying five cosmological parameters using only 37 training simulations.

Once the parameter space is designed, and the to-be emulated observable is obtained at each of the nodes, the emulator is trained using machine learning. For the machine learning method used, different projects use different methods, which are all able to reach sufficient accuracy. Some examples of used methods are Gaussian processes (Heitmann et al., 2009; Bocquet et al., 2020; Moran et al., 2022), polynomial chaos expansion (Euclid Collaboration et al., 2019) and neural networks (Angulo et al., 2021). At this moment emulation has been used mostly for the power spectrum and halo mass function, however, as we will explore in later sections and chapters, the ability to directly predict observables as a function of cosmology and simulation parameters is a powerful tool that should be used for more and more goals. An example of this is the emulator for weak lensing aperture masses by Debackere et al. (2022). Emulators using hydrodynamical simulations will be discussed later.

1.2.6 Cosmological hydrodynamical simulations

The most self-consistent way to model the universe is via the use of cosmological hydrodynamical simulations, which are the main type of simulations used in this thesis. As the name suggests, these simulations also model the evolution of the

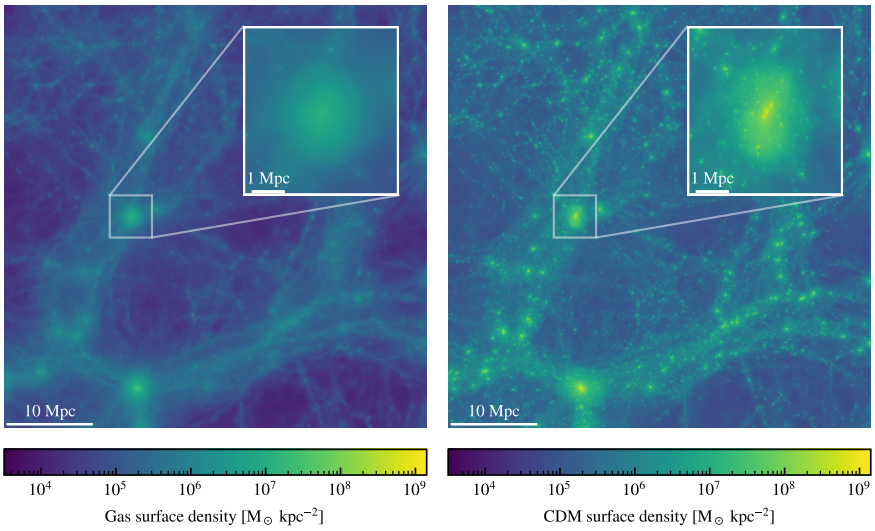


Figure 1.5: A comparison of the gas (left) and cold dark matter (right) distribution of a small cluster in the FLAMINGO suite of simulations. Figure taken form Schaye et al. (2023).

baryonic matter field using hydrodynamics. This complicates the modelling as this means that the hydro-forces now have to be solved on top of the gravity-only calculation described before for the dark matter. A comparison of gas and dark matter in such a simulation can be seen in Figure 1.5.

To solve the hydro forces, one can take one of two approaches: a Eulerian approach or a Langrangian approach. Both approaches, and methods that are in between, are widely used. One of the main numerical difficulties for cosmological simulations is that there is a very wide range of scales that need to be resolved. From the large-scale structure at scales of ~ 100 Mpc to galaxies at scales smaller than ~ 10 kpc there are already four orders of magnitude, without even resolving anything that happens inside galaxies. The main numerical challenge is focusing the computational effort in the dense regions where there is more to resolve, while degrading the resolution in regions where barely anything happens like in cosmic voids.

For Lagrangian particle methods this sort of optimisation comes quite naturally. As the resolution elements are represented as particles with similar mass, higher-density regions are resolved by more closely spaced particles. A popular lagrangian method is Smooth Particle Hydrodynamics (SPH) which is used by, for example, the codes GADGET-4 (Springel et al., 2021) and SWIFT (Schaller et al., 2023), with specific implementations like SPHENIX (Borrow et al., 2022) and Gizmo (Hopkins, 2015). In SPH, while particles have similar mass, the density

can vary. The smoothing length, which defines the area over which the particles mass is distributed, can become larger and smaller depending on the density. In high density regions the smoothing lengths are smaller and the resolution is higher. In low density regions the smoothing lengths are larger and the resolution is lower.

Because of the multi-scale nature of the large-scale structure, a grid code with a uniform grid is too expensive for most applications. One solution is to use adaptive mesh refinement, like the code RAMSES (Teyssier, 2002). In this case when a cell in the grid becomes too massive, the cell is split into smaller cells, increasing the local resolution. Again roughly conserving mass is the natural choice here, as this will lead to higher resolution in dense regions.

Additionally one can make use of a lagrangian grid code. One example of this is the code Arepo (Springel, 2011; Weinberger et al., 2020). Arepo makes use of a moving mesh. In this case advantages of both methods are combined. A moving adaptive mesh allows the equations of hydrodynamics to be solved as if it was a regular grid code, which is usually seen as more accurate than for example an SPH approximation. At the same time the mesh is allowed to moved, so the computational time can be focused in the higher density regions.

One of the main difficulties of doing cosmological hydrodynamics simulations comes from the fact that a lot of processes that are important for galaxy formation play out on much smaller scales than the simulation can resolve. From star formation, stellar feedback and black holes, that can play out on sub-parsec scales, to atomic cooling, it is unlikely that we will resolve the full range of scales in the foreseeable future. However, as these processes are able to affect the scales that are resolved by the simulation, they are added as subgrid/sub-resolution models, models for processes that are not resolved by the simulation. In the next few sections the most important subgrid models for cosmological hydro simulations will be detailed.

1.2.7 Radiative cooling and star formation in hydro simulations

The first model that will be discussed is what is referred to as radiative cooling, though models for cooling generally include the effects of both radiative heating and cooling. Cooling in cosmological simulations generally refers to processes that affect the internal energy of the gas elements via radiative processes. Because most simulations do not include ray-tracing of the light that is produced by star particles, stellar radiation fields are usually also added to the cooling.

The complexity of cooling models is set by the assumptions that are made when computing the cooling rates. Early models assumed collisional ionisation equilibrium and solar relative abundances (e.g. Sutherland & Dopita, 1993). However, there are many other processes that can influence what is called cooling in simulations. Examples of these are photo-ionisation, variations in the rel-

1

ative abundances of certain elements, a UV background, high density gas being affected by a stellar radiation field and self-shielding. Depending on the assumptions, the energy loss of particles can change by quite a bit, and it can have a big impact on, for example, the lowest temperatures gas can cool too. In order to calculate cooling rates, it is common to use packages that make use of theoretical calculations and lab results, like CLOUDY (Ferland et al., 2017). To incorporate a changing radiation field due to the UV background, stellar light and self-shielding, modern cooling tables make use of the local density (Ploeckinger & Schaye, 2020).

Once gas is allowed to cool, there could be instances of runaway collapse inside simulations. In this case gas can start cooling very rapidly and reach very high densities. Besides potentially leading to nonphysical situations, this will also cause the calculations of the hydro forces to become much more expensive. In the regime of large cosmological simulations, where it is unfeasible to add the feedback processes that can halt runaway collapse and create a realistic multi-phase ISM, this problem is solved by adding an artificial "equation of state" (EoS) via the use of an entropy floor. The form introduced by Schaye & Dalla Vecchia (2008) is given by

$$T_{\rm EoS} = 8000\ {\rm K}\left(\frac{n_{\rm H}}{10^{-1}\ {\rm cm}^{-3}}\right)^{1/3}, \tag{1.3}$$

where $n_{\rm H}$ is the hydrogen number density. Gas that cools onto the EoS effectively acts as a subgrid model for a properly pressurised multi-phase ISM.

The addition of the EoS also informs us about subgrid models for star formation. As described by Schaye & Dalla Vecchia (2008), the EoS in Eq. 1.3 corresponds to a constant Jeans mass of $10^7\ {\rm M}_\odot$. Part of the gas that is close to or on the equation of state should therefore be assumed to be able to form stars. When it comes to models of star formation, an equation of state is however not necessary. The simplest method, which can be sufficient for simulations that reach very high densities, is to have a density cut above which gas gets instantly converted into stars, see for example Schaye (2004). For models of star formation, a broad range of observations exist to create an empirical model for where stars can form. Specifically, the model described by Schaye & Dalla Vecchia (2008) uses the relation between star formation rate surface density and gas surface density observed by Kennicutt (1998) to prescribe the level of star formation as a function of the pressure of the gas. In this case, each gas particle that has a high enough density is assigned a star formation rate based on its pressure. Every time-step of the simulation the gas particles can then be stochastically converted into a star particle based on this star formation rate. The formation of star particles also naturally helps with the runaway collapse problem as very high density particles will be converted to stars. Star particles are usually treated as collisionless particles.

1.2.8 Feedback in hydro simulations

In cosmological simulations, the models that have the most impact, given their uncertainties, are models for feedback. The main feedback processes are stellar feedback, originating from young stars, and AGN feedback, originating from accreting supermassive black holes. They are called feedback models because these models are responsible for regulating the growth of galaxies. When a galaxy has a lot of gas, this will lead to an increase in star formation and/or black hole accretion. Without feedback, these processes wouldn't stop until the galaxy runs out of gas. Due to feedback, either in the form of supernovae or radiation and jets from AGN, these processes are regulated as gas is heated and transported out of the galaxy before it can lead to more star formation and black hole accretion. As the amount of feedback is directly proportional to the amount of star formation and black hole accretion, this causes the galaxy to go trough "feedback loops". In each of these loops star formation increases until enough supernovae go off to stop it again, after which either the gas that got heated by the supernovae slowly cools again, or when the ejected gas is replaced by newly accreted gas, repeating the cycle. A similar loop exists for supermassive black holes. In the following paragraphs both types of feedback will be described in detail.

To implement methods of feedback, the first central question is a numerical one. Both supernova and AGN feedback are characterised by violent behaviour on short timescales. As detailed earlier, feedback happens when a galaxy accretes a lot of gas. Feedback is thus more likely to happen when gas is very dense. Numerically, dense gas is most susceptible to undergo a "cooling catastrophe" (see Dalla Vecchia & Schaye, 2012). For both forms of feedback, if the energy that is added to the gas is coupled too gently, for example when the energy is added continuously at each time-step, it is likely that all the feedback energy is simply radiated away before doing any work. Note that this is mostly a resolution effect caused by relatively long timesteps. Many simulations therefore opt for injecting feedback energy in large stochastic bursts. Alternatively, the gas particles used for feedback can be decoupled from the hydrodynamics until they leave the galaxy. Here, the lower density CircumGalactic Medium (CGM) allows for much more efficient feedback. Another option is to temporarily suppres their cooling rates.

Let's first discuss supernova feedback and how it's implemented. The two main methods of transferring energy are via kinetic or thermal feedback. When the resolution gets very high, the choice doesn't matter, as the shock created when the energy is injected will equilibrate to the Sedov-Taylor solution. At this point it is important to take a moment to realise that for most cosmological simulations, supernova feedback does not work with single supernova. Instead, a stellar particle represents a population of stars and, based on the initial mass function, each particle injects the energy of a population of supernovae. This means that for the

1

lower-resolution simulations, supernova feedback has to describe how collections of supernovae couple to the rest of the galaxy.

When it comes to implementing this sort of feedback, both the total amount of energy that is injected, and the amount of energy that is put into a single (numerical) energy injection, are parameters. In the simplest model a stellar particle dumps the energy into a neighbouring gas particle or cell. For thermal models, the energy per injection is set by the change in internal energy, often characterised by the temperature jump ΔT, see for example the work by Dalla Vecchia & Schaye (2012). For kinetic models it is characterised by the change in velocity, often denoted as $\Delta v_{\rm kick}$ (Dalla Vecchia & Schaye, 2008). In both cases the chosen value in large part sets the efficiency of the energy injection. For values that are too low, the energy is dissipated before it can do any work. For large values, the feedback becomes more rare, giving more time for the gas to cool in between events or a single effect can completely destroy a galaxy. While arguably less physical, it is also possible to decouple the gas particles from the hydro until they reach the CGM, this is the method employed by Pillepich et al. (2018). At the lower CGM densities the feedback is much more efficient. For more complex models like the upcoming COLIBRE simulation, mixes of kinetic and thermal energy are use to drive both galactic outflows, and turbulence in the ISM (Chaikin et al., 2022).

When it comes to the total energy injected, there is a large amount of variety. At lower resolutions a single fixed energy usually suffices (Schaye et al., 2010; Kaviraj et al., 2017; Hirschmann et al., 2014; McCarthy et al., 2017; Schaye et al., 2023). However, for both numerical and physical reasons, the energy can be made a function of the properties of the halo/gas that is being shocked. For the IllustrisTNG model the energy depends on the velocity dispersion of the dark matter halo and redshift Pillepich et al. (2018). For the EAGLE model the energy is a function of both the metallicity and birth density of the stellar particle Schaye et al. (2015). These models also introduce additional parameters. We will discuss the calibration of free parameters in the next section.

The injection of AGN feedback energy into the gas has similar problems to supernova feedback. However, before a black hole can become an AGN and do feedback, it first needs to be seeded and grow via accretion and mergers with other black holes. The amount of feedback produced by an AGN is tied to its mass. Before an AGN can become dominant, it will first have to grow from a small seed to a mass that is large enough to start injecting feedback energy.

In order to seed black holes, most simulations frequently run a halo finder during the simulation (Springel et al., 2005b). When a halo is above a certain mass, and does not already have a black hole, a black hole is seeded in its center. The halo mass for seeding black holes and the black hole mass that is seeded are both free parameters. Both parameters have an influence on how quickly the black hole can grow. When black holes are small, one of the main growth mech-

anisms is via mergers with other black holes. In the perfect case, the dynamics of the black holes are resolved. However, in cosmological simulations the mass of the black hole particles is usually similar to or smaller than the other particles. This means that dynamical friction is not properly resolved by the simulation. Therefore black hole might wander too much within the galaxy and mergers may not happen on realistic timescales. To counter this, black holes are often explicitly repositioned to keep them at the centers of galaxy. This can have a big effect on the growth of black holes and the efficiency of their feedback (see e.g. Bahé et al., 2022, and references therein).

The other channel for black hole growth is via gas accretion. For hydro simulations for cosmology, the black hole is usually assumed to accrete gas mostly trough hot accretion. This is described by Bondi-Hoyle accretion

$$\dot{m}_{\rm accr} = \frac{4\pi G c^2 m_{\rm BH}^2 \rho}{\left(c_{\rm s}^2 + v_{\rm BH}^2\right)^{3/2}}, \tag{1.4}$$

where G is the gravitational constant, c is the speed of light, $m_{\rm BH}$ is the mass of the black hole, ρ is the gas density around the black hole, $c_{\rm s}$ is the local speed of sound and $v_{\rm BH}$ is the black hole velocity. One of the most important aspects of this accretion channel, is that it scales as the square of the black hole mass. This implies that when black holes get more massive, they are able to grow quadratically faster at a constant background density. Hence, when a black hole gets big enough, accretion-driven growth becomes more important than mergers.

In order to calculate the accretion rate of the black hole, there are a few additional models that are different from Bondi-Hoyle accretion. These can go from an angular momentum limiter, like the one used in the EAGLE simulation (Schaye et al., 2015), to direct accretion in very high resolution simulations (Anglés-Alcázar et al., 2021). For simulations at relatively low resolutions and simulations that employ an equation of state, the gas around the black hole may not be able to reach sufficiently high densities for the black hole to accrete efficiently. In that case Booth & Schaye (2009) describe that it might be preferred to add a boosting factor to Bondi-Hoyle accretion. This takes the form of

$$\alpha = \max\left[\left(\frac{n_{\rm H}}{n_{\rm H,*}}\right)^{\beta_{\rm BH}}, 1\right], \tag{1.5}$$

where $n_{\rm H}$ is the hydrogen number density of the gas, $n_{\rm H,*}$ is the hydrogen number density above which star formation is allowed and $\beta_{\rm BH}$ is a free parameter. The boost at high densities allows the black hole to accrete as if it is in a cloud of much higher density, allowing the black hole to grow on a more realistic timescale.

A fraction of the mass that is accreted onto the black hole is used to fuel AGN feedback. For AGN feedback the numerical problems are similar to those described for stellar feedback. Catastrophic overcooling has to be accounted for.

1

Compared with supernova feedback, AGN feedback is often implemented with more energetic effects, and it therefore also acts on larger scales. Similarly to supernova feedback, energy is distributed kinetically or thermally. A single mode of feedback can be sufficient, see for example both EAGLE (Schaye et al., 2015) and BAHAMAS (McCarthy et al., 2017). However, there are implementations that use multiple modes of AGN feedback, based on how close to the Eddington rate the black hole is accreting. In these models, used for example in SIMBA (Davé et al., 2019) and IllustriTNG (Pillepich et al., 2018; Pakmor et al., 2023; Nelson et al., 2023), the feedback is split up into two modes, usually referred to as jet mode at low accretion rates, and quasar mode at high accretion rates. In jet mode the AGN energy is often injected kinetically, along the axis of angular momentum of the black hole (Huško et al., 2022). In quasar mode, the energy is injected either thermally on kinetically, but isotropically to mimic a strong radiation field emanating from the black hole. Much like supernova feedback, these numerical recipes come with free parameters, the number depending on the complexity of the model, that need to be calibrated.

1.2.9 Calibration of hydro simulations

If there is one thing that should be clear from the previous sections, it is that there are not only many methods for each subgrid model, but additionally each subgrid model contains uncertain parameters that we cannot constrain from first principles. Therefore, when preparing to run large hydrodynamic simulations, care has to be taken to make sure the subgrid models are tuned in a way that the simulation is able to reproduce some relevant observational data. This sort of approach was applied to semi-analytic models (see e.g. Bower et al., 2006) before it was applied to hydro simulations (Schaye et al., 2015; Crain et al., 2015).

Depending on the goal of the hydro simulations, different observables are chosen as calibration targets. For galaxy formation simulations, properties like galaxy sizes and star formation rates are more important than for hydro simulations that investigate cosmological inference.For cosmology, the priorities are the properties of the gas in and around clusters, and ensuring that galaxies of a certain stellar mass occupy halos with the correct mass via the stellar mass function. In choosing the data to use for calibration, there are a few considerations to take into account. Some observables might look very promising for calibration, but they are observationally not very constrained. A good example is the stellar-mass-halo-mass relation. In theory constraining this relation would lead to the most constraining results. However, the stellar mass function is much better constrained observationally, so it is a better target. Similarly, for clusters, the baryon mass fraction would be much more constraining than the gas mass fraction, but much more data exists for the gas fraction, making it the better choice.

It is in the attempt of matching many observables that we learn the most about

our subgrid models. Inconsistencies, degeneracies between astrophysical processes and observables that we cannot match provide the biggest gateway into improving our models and hence our understanding of the physics of structure formation. It goes without saying that a model has to be robust to trust its predictions. This also makes it interesting for different simulation projects to take different approaches. Seeing how the differences between approaches manifest themselves when comparing predictions contains a wealth of information.

The differences in approaches come with changes in calibration targets, the resolution, box size and subgrid physics. Each can vary greatly between different simulations. One of the goals of some of the earlier simulations, like OWLS (Schaye et al., 2010), was to investigate the effects of the different subgrid models on, for example, the cosmic star formation history. Many higher resolution simulations focus on galaxy formation and evolution with different approaches and methods. EAGLE (Schaye et al., 2015), FABLE (Henden et al., 2018) and SIMBA (Davé et al., 2019) make use of SPH, HorizonAGN makes use of adaptive mesh refinement (Kaviraj et al., 2017) and IllustisTNG makes use of a moving mesh code (Pillepich et al., 2018). None of those simulations investigate the multiphase ISM, this is one of the main goals of the FIRE box simulations (Feldmann et al., 2023). To investigate clusters and groups of galaxies, it is necessary to simulate a large volume, at the cost of a lower resolution. This is the approach of simulations like Magneticum (Hirschmann et al., 2014), BAHAMAS (McCarthy et al., 2017) and the higher res MilleniumTNG (Pakmor et al., 2023).

One of the ways to investigate higher mass objects at a high resolution is to make use of cosmological zoom simulations. For these simulations a full cosmological volume is simulated, but only a very small fraction of the box is run at a high resolution. This approach can be used for Milky way galaxies, like in the APOSTLE project (Fattahi et al., 2016) and Auriga (Grand et al., 2017), but also for groups and clusters like in C-EAGLE (Barnes et al., 2017b) and Hydranga (Bahé et al., 2017). One disadvantage of the zoom approach is that you are limited to a small number of hand selected objects, making it impossible to look at things like selection effects. One recent approach to partially circumvent this is to do a zoom for many objects obtained from a single large volume, like for example the MACSIS simulations (Barnes et al., 2017a), the Three-hundred simulations (Cui et al., 2018) and ClusterTNG (Nelson et al., 2023).

The final simulation method, which has not yet been maximally explored but will likely become much more popular in the future, is to run suites of simulations for machine learning, much like how emulators are currently becoming more wide-spread for dark matter only simulations. The biggest example is the CAMELS suite of simulations (Villaescusa-Navarro et al., 2021), that has many thousands of variations of subgrid physics parameters, cosmology and numerical methods, but restricted to a small volume. The ANTILLES suite (Salcido et al., 2023a) also has many variations of subgrid physics parameters. As demonstrated

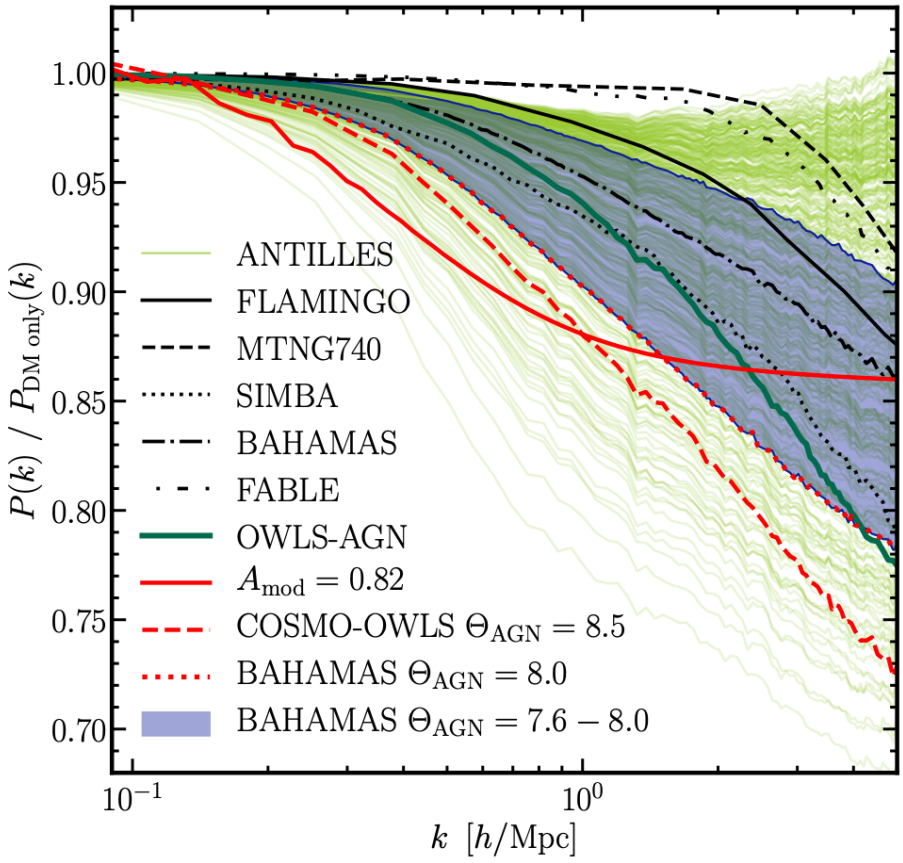


Figure 1.6: Ratio of the power-spectrum between full hydro simulations and the corresponding dark matter only simulations as a function of scale. The different lines show a large number of results of different hydro simulations. The blue region and the red lines are obtained from the HMCode (Mead et al., 2021) implementation of baryonic suppression. Plot taken from Bigwood et al. (2024).

by Brown et al. (2024), emulators can also be used to increase the interpretability of subgrid models. Emulators also form the basis of the calibration of the FLAMINGO simulations (Schaye et al., 2023), which is detailed in the next chapter of this thesis.

1.2.10 Baryons and cosmology

A lot of hydro work is done on galaxy formation. For cosmology, one of the main results of hydro simulations is that baryons often cannot be ignored for cosmological inference. One of the main effects is what is referred to as baryonic suppression. Because of AGN feedback in groups and clusters, gas is pushed out of groups, and the gas around clusters is heated to higher pressures, limiting the gas accretion rate. This leads to the matter distribution being "washed out" at scales of the order of a Megaparsec. Compared to dark matter only simulations, this leads to suppression of power, or in other words, a decrease in clustering on these scales. Because the baryons decrease the clustering in the late universe, this is seen as one of the candidate processes to explain the σ_8 tension. If baryonic suppression is powerful enough, it would bias the lensing surveys to find a lower σ_8 when compared to early universe measurements. This effect is shown as predicted by a large range of simulations in Fig. 1.6. As is clear from the figure, there is no clear estimate of how strong this effect should be.

It is however very clear that there is an effect, as the effect is not only predicted by hydrodynamical simulations (see e.g. Van Daalen et al., 2011; Semboloni et al., 2011, 2013), but also by by analytical models constrained by observations (Debackere et al., 2020). In order to model baryonic effects many approaches can be taken. From analytical fits (van Daalen et al., 2020; Mead et al., 2021), to baryonification algorithms applied to dark matter only simulations (Chisari et al., 2019; Aricò et al., 2021; Giri & Schneider, 2021), or emulators (Salcido et al., 2023b). In many of these models, the strength of the effect can be related to the baryon content in high mass haloes. As shown by van Daalen et al. (2020), the baryon fraction is a good probe for the level of baryonic suppression, allowing us to place external constraints on the baryon effects on the power spectrum.

Besides effects due to baryonic suppression there are other observables, less directly related to the power spectrum, that are sensitive to both baryons and cosmology. A prime example is the thermal SZ effect, which, as shown by McCarthy et al. (2018) is mostly sensitive to cosmology on large scales, but on the smaller scales baryons and neutrinos can have a big effect. The same is true for kSZ profiles (see e.g. Bigwood et al., 2024). However, in simulations, current X-ray measurements of the baryon properties of clusters and groups, and predictions for gas properties, clustering, lensing, thermal SZ and kinetic SZ are in tension. This tension might be partially due to σ_8 but there are hints towards a tension between baryonic observables as well.

1.2.11 The FLAMINGO simulations

Currently what would be extremely informative for the field is large-volume full-hydro simulations that explore potential systematic effects that our cosmological probes are sensitive to. This is the main goal of the simulations constructed and

1

used in this thesis: The FLAMINGO[2] simulations. The two biggest simulations are a $(2.8\ \mathrm{Gpc})^3$ simulation with 5040^3 gas resolution elements, leading to a resolution of $1.09\times10^9\ \mathrm{M}_\odot$, and a $(1\ \mathrm{Gpc})^3$ with 3600^3 resolution elements, leading to a resolution of $1.34\times10^8\ \mathrm{M}_\odot$. The $(2.8\ \mathrm{Gpc})^3$ simulation is currently the biggest cosmological hydrodynamics simulation that reached $z = 0$. One of FLAMINGO's defining features is the systematic variations that are done in the $(1\ \mathrm{Gpc})^3$ volume. In this volume the resolution is varied up and down by a factor eight, the cosmology is varied, including specific variations in the neutrino mass, the feedback is varied according to systematic shifts in the observables that where used for calibration, and there are runs with an alternative model for AGN feedback.

In addition to having many variations, FLAMINGO has a large set of outputs. Most notably these include full-sky lightcones that where produced while the simulation was ongoing. The lightcones come in two forms, maps of certain shells in redshift up to high z, and a full particle lightcone up to the redshift where the volume starts to repeat on itself, and up to even higher z for hot gas particles. This also includes a halo lightcone, with which the locations of halos in the volume can be aligned with the lightcone maps and particles. This gives many opportunities to forward model observations and make a direct comparison to simulated quantities. Additionally there are the standard simulation outputs, like finely spaced snapshots of all the particles and halo catalogs for each of them calculated using the code SOAP[3]. The use of SOAP, also allows for easy comparison between different halo finders, which include VELOCIraptor (Elahi et al., 2019) and HBT (Han et al., 2012). The simulation outputs also include X-ray luminosities for every particle, described by Braspenning et al. (2023), that are consistent with the cooling in the simulation.

For its subgrid models FLAMINGO uses an evolution of the subgrid models used for OWLS (Schaye et al., 2010) and BAHAMAS (McCarthy et al., 2017, 2018). It has models for radiative cooling (Ploeckinger & Schaye, 2020), star formation (Schaye & Dalla Vecchia, 2008), stellar mass loss (Wiersma et al., 2009), supernova feedback (Chaikin et al., 2022), black hole seeding and accretion (Bahé et al., 2022) and for thermal (Booth & Schaye, 2009) and kinetic jet-like AGN feedback (Huško et al., 2022). While the models are all similar in spirit to earlier simulations, nearly each one of them has received updates when compared with previous work. The calibration of the subgrid physics is done using machine learning, which is the topic of the next chapter of this thesis, so more details can be found there.

The FLAMINGO simulations have already lead to some remarkable results. The calibrated model of FLAMINGO has less baryonic suppression than BAHAMAS, especially at scales around and slightly below $k \approx 1\mathrm{Mpc}^{-1}$ (Schaye et al.,

[2]Full-hydro Large-scale structure simulations with All-sky Mapping for the Interpretation of Next Generation Observations.

[3]Spherical Overdensity and Aperture Processor, https://github.com/SWIFTSIM/SOAP

2023). FLAMINGO also predicts that for reasonable variations in the subgrid physics, the baryonic effects are not strong enough to resolve the σ_8 tension in lensing or SZ (McCarthy et al., 2023). The same can be said for the effect of massive neutrinos, while they have some impact, the total effect needed requires a neutrino mass that is larger than current constraints allow (Elbers et al., 2024).

1.3 Galaxy Cluster Cosmology

The final part of this introductory chapter concerns what many say are the biggest bound objects in the Universe, galaxy clusters. Unfortunately, that description is not quite correct, and a better description is the largest virialised objects in the Universe. As expected from hierarchical structure formation, they are also the last objects to form. Therefore, galaxy clusters are a powerful tool to explore both the cosmological parameters and galaxy formation. The study of galaxy clusters typically concerns objects of a mass M_{500c}[4]$> 10^{14}\ M_\odot$, though with the increasing sensitivity of current surveys, more and more objects at lower masses are also being considered. Objects with masses in the range $10^{13}\ M_\odot < M_{500c} < 10^{14}\ M_\odot$ are typically refereed to as galaxy groups.

As galaxy clusters are such extreme objects, observations of single clusters have also made a huge impact, especially when it comes to our thinking about dark matter. The first detection of a dark matter signature originates from observations by Zwicky (1933) of the Coma cluster. When comparing the observed mass of Coma with the velocity dispersion of the galaxies inside Coma, there was a large inconsistency between the two measurements when the dispersion would be due to virialisation, hinting at a dark matter accounting for additional gravity. Another big dark matter discovery was found via the bullet cluster (Clowe et al., 2006). The bullet is a merging cluster. By observing the cluster in both X-rays for the gas, optical for the stars, and by making use of weak lensing to reconstruct the mass distribution, the authors found a clear dark matter signal originating that was separated from the gas, where the bulk of the baryonic mass lies.

To understand how galaxy clusters can be used to infer cosmology, we must make the connection to the halo mass function (HMF). The HMF characterises the number of clusters per unit volume and mass. The expected shape of the HMF is a power-law, with an exponential turn-off at the highest masses. The HMF is very sensitive to the underlying cosmological parameters. Especially at the high mass end, at galaxy cluster masses, small changes in the clustering (via σ_8) or in the matter content (via Ω_m) can lead to large changes in the HMF. Additionally, as clusters form relatively late, they are also sensitive to changes in dark energy.

[4] M_{500c} is defined as the mass inside a spherical aperture where the average matter density is 500 times the critical density of the universe.

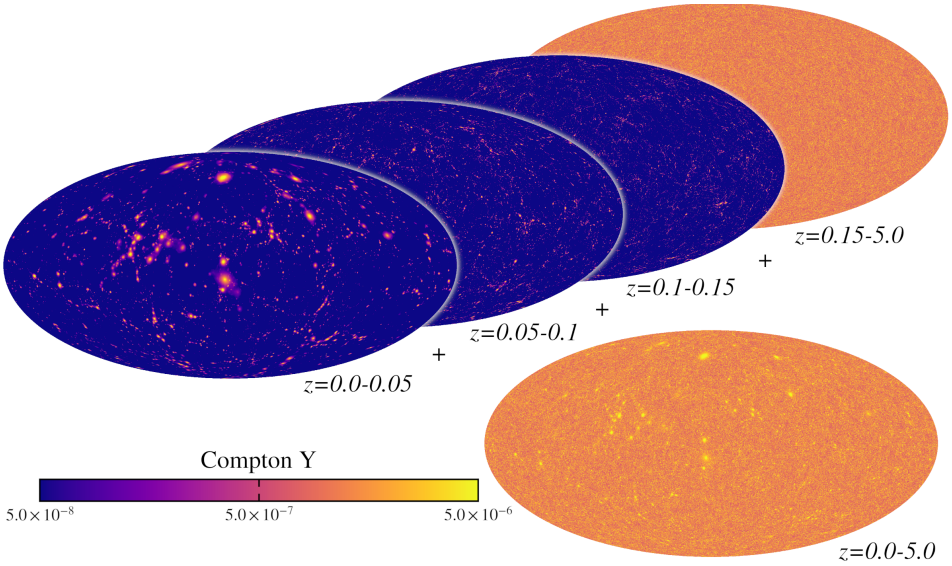


Figure 1.7: Full sky maps of Compton-Y in the FLAMINGO simulation. The figure shows a few individual shells and how they add up to the full signal up to $z = 5$. Figure taken from Schaye et al. (2023).

When observing clusters, with the exception of when we use lensing, we cannot directly observe their masses. Instead, we have to resort to using a mass-observable scaling relation. By making use of the scaling relation we can turn an observable mass proxy back into a mass, and then infer the underlying HMF. With clusters being such extreme objects, they can be observed across the entire electromagnetic spectrum. Some examples are: via radio (van Weeren et al., 2019), via the SZ effect (Sunyaev & Zeldovich, 1972) on the cosmic microwave background, via detecting cluster members in the optical (Rykoff et al., 2014, 2016; Black & Evrard, 2022), via X-rays (Pierre et al., 2016; Liu et al., 2022; Bulbul et al., 2024) and directly from weak lensing (Costanzi et al., 2019). A map of the SZ effect from the FLAMINGO simulation is shown in Figure 1.7, the clusters are clearly seen as large overdensities in the map. To get unbiased inference of cosmological parameters, each observable needs a well constrained mass-observable scaling relation, and we need a good grip on the selection effects inherent to each observable.

This section of the introduction will provide an overview of current constraints, selection effects found in cluster samples, how cluster counts are modelled and finally an outlook on the future.

1.3.1 Current constraints

In the last few years, many new surveys have released their cluster catalogues, and many have released cosmological constraints along with their catalogues. With the rapidly increasing sample sizes, and the ability to simulate ever increasing volumes, this is an exiting time to work on galaxy clusters, and cluster cosmology. While cluster cosmology can not yet reach the accuracy of the lensing and clustering surveys, it is becoming a competitive alternative. When it comes to large, recent cluster surveys this thesis focuses on surveys based on optical, X-ray and SZ selections.

Galaxy richness selection In galaxy clusters, the bulk of the mass is found in dark matter and gas. Therefore, in the optical, that is mostly sensitive to the stellar light, we cannot directly probe the majority of the mass. The observable that is most closely related to the mass is the number of satellite galaxies associated to the cluster. This is referred to as the galaxy richness, often denoted with λ. While the counting of galaxies might seem straightforward, there are many effects that make it difficult to accurately assign satellites to the correct center. The first complication is that spectroscopic samples are much more difficult to obtain, so members often have to be identified using photometric redshifts. Additionally, there are many fore and background galaxies that might contaminate the signal.

The most often used solution is to try and model all these effects simultaneously, and create a Bayesian likelihood for each cluster that gives the most likely value of galaxy richness for each cluster. This method is known as Redmapper (Rykoff et al., 2014, 2016). Redmapper directly models the background galaxies, the galaxy red sequence and the number of member galaxies in a aperture whose size depends on the galaxy richness. by making use of the red sequence it can use the photometric data of each galaxy to decide whether the object is truly part of the cluster. Additionally field galaxies are usually found in the blue cloud. For the complete statistical sample, this also allows Redmapper to self-calibrate its photometric redshifts, as it models the red sequence as a function of redshift simultaneously. Redmapper has been applied to the SDSS data (Rykoff et al., 2014), DES verification data (Rykoff et al., 2016) and DES year 3 data (Pereira, 2021). The SDSS and DES year 3 have also used the RedMapper samples to constrain cosmology. In both cases, the samples need an additional calibration step to constrain the mass-observable scaling relation. For the cosmology from the SDSS sample, this has been done using both X-ray calibration (Kirby et al., 2019) and calibration using weak lensing (Fumagalli et al., 2024). A lensing approach to the calibration is also used for the DES year 3 results (Pereira, 2021). The cosmological constraints found using these samples are in agreement with the results from the Planck survey, though they are not constraining enough in general to find discrepancies at the level of the current S_8 tension.

1

X-ray selection X-ray selected cluster samples are usually based on a flux cut. In this case the survey instrument sweeps a certain region of the sky, and all objects that fall above that flux cut are included in the cluster sample. The X-ray photons that are detected originate from the very high temperature intracluster medium (ICM). One of the advantages of using X-rays is that the X-ray emission can also be directly used for calibration of the mass-observable scaling relation. Via the assumption of hydrostatic equilibrium, the mass can be directly estimated from the temperature and density profiles (see e.g. Sarazin, 1988). While real clusters not being in hydrostatic equilibrium does lead to a biased estimate of the mass (see e.g. Hoekstra et al., 2015; Eckert et al., 2016; Smith et al., 2016), this is an effect that can be accounted for when inferring cosmological parameters. Besides making a statistical inference, we can use lensing measurements, to directly infer the value of the hydrostatic bias from observations.

There are many X-ray surveys that have counted the number of clusters on the sky. The survey area and depth have improved by orders of magnitude. The earlier surveys like ROSAT (Rosati et al., 1998; Ebeling et al., 1998) had of the order of a hundred objects. Currently surveys with that number of objects, like the HSC-XXL survey Eckert et al. (2016); Akino et al. (2022) are able to go much deeper, and also probe lower mass objects. The eROSITA equatorial Final Depth Survey (eFEDS), goes very deep and includes around 550 clusters (Liu et al., 2022). The biggest current X-ray survey originates from the eROSITA all sky survey (Bulbul et al., 2024) that contains over twelve thousand individual objects. With these huge samples, cosmological constraints from cluster cosmology are able to get much tighter.

Nearly all of these surveys have also been used to constrain cosmology. However, the methods used to constrain the mass-observable scaling relation have changed over time. For the cluster cosmology with ROSAT, hydrostatic equilibrium was assumed to connect masses to temperature, which was then connected to a luminosity using observed scaling relations (Borgani et al., 2001). Similar scaling relations are used by Garrel et al. (2022), however, they are either calibrated using their own sample, or left as a free parameter. Both eFEDS (Chiu et al., 2023) and eROSITA (Ghirardini et al., 2024) make use of overlap with weak lensing surveys to calibrate their mass-observable scaling relations. At this point the large X-ray survey samples are able to put good constraints on the cosmological parameters, close to the level of Planck CMB for $\Omega_{\rm m}$ and σ_8.

SZ Selection The last method we will discuss is detecting clusters via their Sunyaev & Zeldovich (1972) signal. As CMB photon propagate the universe, they might interact with free, high energy, electrons trough inverse Compton scattering (referred to as the SZ effect). This effect scales directly with the electron pressure, which is very high at the centers of galaxy clusters. The SZ effect leaves a distinct spectral distortion in the CMB, which can be used to detect clusters.

In particular, by making use of a matched filter, clusters are detected by finding the exact spectral distortion among CMB maps at different frequencies. Because clusters have similar pressure profiles (Arnaud et al., 2010), they have a clear expected spacial and frequency signal. By using the expected signal in a matched filter approach, sources can be extracted, even when the signal is very faint (Melin et al., 2006, 2012). The SZ effect does not depend on redshift, even in projection, making it also sensitive to picking up clusters at high-z.

Whenever the CMB is detected at multiple frequencies, it is also possible to try and detect clusters via the SZ signal. The largest SZ cluster catalogues originate from the Planck satellite (Planck Collaboration et al., 2014, 2016b), the South Pole Telescope (SPT) (Bleem et al., 2024) and the Atacama cosmology telescope (Hilton et al., 2018, 2021). With these advances, the number of SZ detected sources has gone from a hundred with Planck, to thousand(s) with SPT and ACT. From these cluster samples it is possible to constrain cosmology. The results from the Planck SZ clusters (Planck Collaboration et al., 2016a) find a slight discrepancy in cosmological parameters with the primary CMB. This discrepancy is not found with the SPT clusters (Bocquet et al., 2024). The Planck cluster cosmology made use of a scaling relation calibrated using hydrostatic masses, and their analysis acounts for the hydrostatic bias, while the SPT cluster cosmology made use of a weak lensing calibration.

While SZ is a very clean and powerful method for the selection of clusters, it does have several drawbacks. One of the effects is source confusion. Also, when calculating the SZ signal it is important to take into account foregrounds as their spectral features might be wrongfully identified as clusters. Foregrounds include dust in clusters (Melin et al., 2018) and the cosmic infrared background (Zubeldia et al., 2023). Because the signal is already quite weak, these foreground effects add additional difficulty when selecting galaxy clusters.

1.3.2 Selection effects

In order to obtain clean samples with each selection method, it is important to have a good idea of systematics and biases introduced by the selection. Observationally, there have been many comparisons between samples with different selection methods, and clear differences have been found.

Lovisari et al. (2017) use X-ray observations to investigate the morphology of SZ selected clusters and find that they are more disturbed than X-ray selected sources. X-ray selected clusters are instead found to have a larger fraction of cool-core clusters. Both Rossetti et al. (2017) and Andrade-Santos et al. (2017) find an increase in cool core fraction for X-ray samples when compared to SZ samples. The difference is attributed to an increase in the central density for cool core clusters, which disproportionately boosts the X-ray flux compared with the SZ signal. However, Chon & Böhringer (2017) suggest that the differences

1

between SZ and X-ray selected samples might not originate from the different selection techniques, but instead they originate from differences between volume and flux limited samples. They do this by comparing a flux and volume limited sample with X-rays. In a comparison between X-ray clusters in eFEDS and optically selected clusters with HSC, Ota et al. (2023) find differences in the fraction of disturbed objects and the slope of the luminosity temperature relation. Marini et al. (2024) use the Magneticum simulations to investigate these effects with mock X-ray surveys. They find that the lowest mass galaxy groups that are detected have a bias towards being more gas-rich.

As long as these effects don't directly influence the mass-observable scaling relation, they will only have a moderate effect on cluster cosmology. However, some of these secondary observables, like the cluster gas fraction, are important for the calibration of hydrodynamic subgrid physics (Semboloni et al., 2011, 2013; McCarthy et al., 2017; Debackere et al., 2021; Giri & Schneider, 2021; Salcido et al., 2023b), and hence using biased samples might lead to unrealistic calibration, hampering our ability to use these constraints to make predictions for baryonic effects on cosmology.

In order to investigate potential selection biases for cosmology, one can compare scaling relations found in the same field for multiple selection methods. Willis et al. (2021) highlight how some of the morphological selection criteria can be influenced by the XMM point spread function and lead to sources being undetected in X-rays when comparing the XMM Newton XXL X-ray selected survey and HSC optically selected survey. When comparing SDSS optically selected clusters with the X-ray selected XMM cluster survey, Giles et al. (2022) find differences in the scatter of the $L_X - \lambda$ relation between the samples. Furthermore, the fits of the $T_X - L_X$ relation are found to be sensitive to the selection method. As scaling relations directly affect the cluster counts, this indicates that different selections might lead to using a biased mass-observable scaling relation. Grandis et al. (2021) indicate that biases due to contamination by smaller haloes might start playing a big role for eROSITA and SPT.

1.3.3 Models for cluster cosmology

If the observations have a well defined selection, the important final step is to create a robust model to compare the observed counts with. This comes down to two components, the HMF, which tells us the expected number of haloes, and an observable-mass scaling relation with its scatter, which relates our selection observable to a mass. By integrating over the observed volume, and accounting for the selection function, these two quantities can be integrated over to predict

number counts. The integral to solve is given by

$$N(X_{\rm C}) = \int_0^{z_{\rm max}} \int_0^{A_{\rm sky}} \int_{M_{\rm min}}^{M_{\rm max}} \phi(M,z,\boldsymbol{\theta}) \chi(M,z,X_{\rm C}) \frac{{\rm dV}}{{\rm d}\Omega{\rm dz}}(\boldsymbol{\theta}) {\rm dM d}\Omega {\rm dz}, \tag{1.6}$$

where $\phi(M,z,\boldsymbol{\theta})$ is the HMF; $\chi(M,z,X_{\rm C})$ is given by the selection function and the scaling relation and gives the probability of a cluster with mass M and at redshift z to be observed for a given cut $X_{\rm C}$; $\frac{{\rm dV}}{{\rm d}\Omega{\rm dz}}(\boldsymbol{\theta})$ gives the differential comoving volume; and finally $\boldsymbol{\theta}$ indicates the cosmological parameter vector. This is the equation is in simplest form. Additionally, the HMF will depend on astrophysical parameters and survey specific selection effects will lead to further modifications of ξ. Getting the modelling of the cluster population and selection effects correctly is key in inferring unbiased cosmological parameters.

To make predictions for the HMF we need to turn to models that can describe it as a function of the cosmological parameters and cosmology. The HMF has a characteristic shape, starting as a power-law, with an exponential turn-off. The HMF of the FLAMINGO simulations is shown in Figure 1.8. A widespread method to predict the how the HMF changes with cosmology was introduced by Jenkins et al. (2001). By the addition of an empirical fitting formula to the extended Press & Schechter (1974) formalism, the HMF can be computed from predictions of the non-linear power-spectrum made by Boltzmann solvers. This requires calibrating the empirical formula to cosmological simulations. The fitting formulae used where updated by Tinker et al. (2008, 2010), which included extensions to be able to predict the HMF for many different spherical overdensity definitions. However, we know that the HMF is affected by astrophysical effects (Velliscig et al., 2014; Bocquet et al., 2016; Schaye et al., 2023). To model these effects, the Jenkins et al. (2001) method was further updated and expanded to hydrodynamical simulations by Bocquet et al. (2016) by making use of the Magneticum simulation (Hirschmann et al., 2014). They also show that there are large differences in the inferred cosmological parameters when introducing hydro, and also between different HMF models. Just as for other cosmological observables, there are now also emulators of the HMF that interpolate the predictions of a large suite of dark matter only simulations. One example of such an emulator is the MiraTitanHMFEmulator (Bocquet et al., 2020), based on the Mira Titan simulations (Heitmann et al., 2016).

For the scaling relation, the standard approach is to assume a power law relation between mass and the observable with lognormal scatter. These relations are then truncated by the selection function to lead to the observed cluster counts. The proper calibration of the scaling relation and its scatter is very important for unbiased results (Mantz, 2019). This also includes potential redshift evolution of the relations. One method is to assume self-similar scaling (Kaiser, 1986, 1991). Most recent cluster cosmology inferences use the overlap between the survey volume and existing weak lensing surveys to calibrate the scaling relation

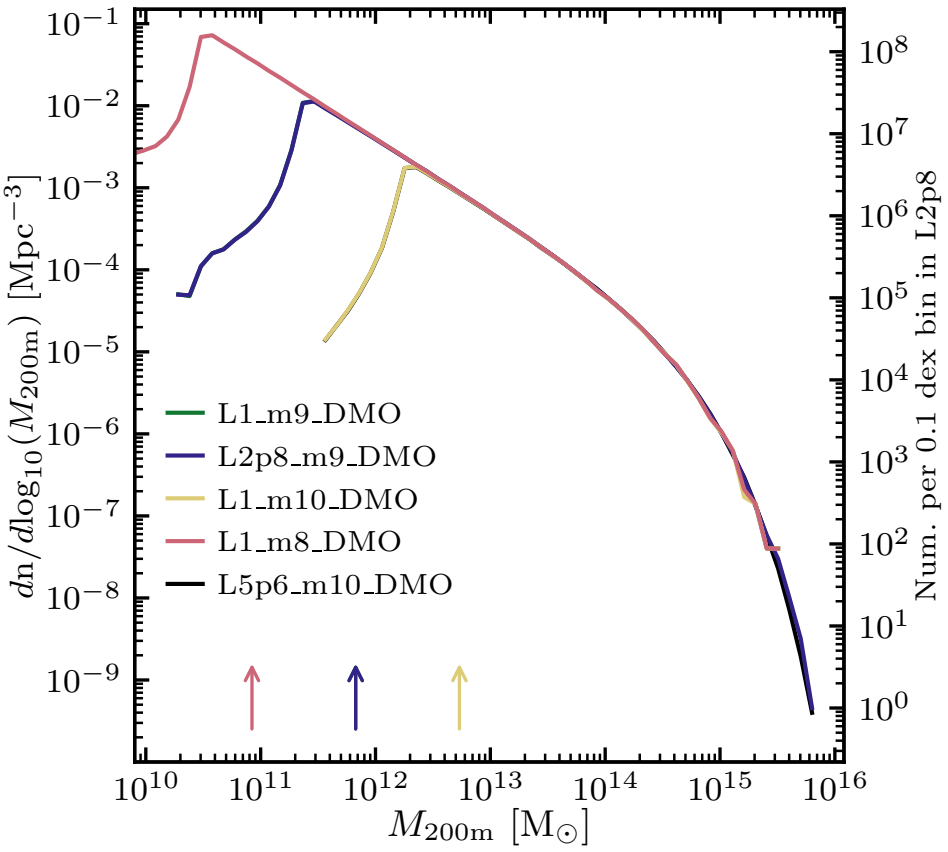


Figure 1.8: The HMF from dark matter only simulation in the FLAMINGO suite of simulations. The number density of objects per logarithmic mass bin is shown as a function of M_{200m}. The different lines indicate which FLAMINGO simulation the line is taken from. On the right y-axis, the number of objects in that bin for the 2p8 simulation is shown. Figure taken from Schaye et al. (2023).

(Chiu et al., 2023; Bocquet et al., 2024; Ghirardini et al., 2024). This way the scaling relation can be constrained self-consistently within the observed volume and simultaneously fit along with the cosmology.

1.3.4 Future surveys

With the release of the first eROSITA catalogue (Bulbul et al., 2024) and the most recent SPT catalogue (Bleem et al., 2024) there has been a huge increase in the available samples sizes. With upcoming surveys like Euclid and LSST, the sample sizes of optical surveys are also likely to increase by orders of magnitude (Artis et al., 2022). Furthermore, eROSITA has not yet reached its final depth, so increases in the X-ray sample are expected soon. For SZ there are multiple upcoming instruments, for example the Simons Observatory (Ade et al., 2019), which are forecasted to push the number of observed SZ clusters to over ten thousand.

Not only will the largest numbers allow us to obtain ever tighter observational constraints, they will also allow for a much more detailed look at the formation and evolution of these objects. This way we will be able to learn more about both cosmology and galaxy clusters in general. One of the main paths forward from a modelling perspective is to carefully forward model simulation results. As the observations get more detailed, the systematics introduced by the instrument, noise and projection will make it much harder to do apples to apples comparisons if they are not properly forward modelled. Careful forward modelling will also further help us understand selection effects, cluster properties, and shortcomings of our simulations.

1.4 This thesis

This thesis will cover a range of topics relating to the setting up of subgrid physics in hydrodynamical simulations, cluster cosmology selection and modelling, and the large scale, cosmic web, force and tidal field.

- In Chapter 2, the calibration strategy for the subgrid physics of the FLAMINGO hydrodynamical simulation suite is described. By making use of Gaussian process emulation, we are able to set up emulators using a training set of simulations where the subgrid physics is systematically varied. By making emulators for the stellar mass function and the cluster gas mass fraction, the subgrid physics can be fit to observations. Furthermore, we fit the emulators to systematic shifts in the observations to obtain constrained subgrid variations that can inform us about the uncertainty in the baryonic modelling.

- In Chapter 3, we use the FLAMINGO simulations to investigate three different selection methods, X-ray, SZ and galaxy richness. We find that depend-

1

ing on the redshift and mass of the sample of interest, different selection methods are more or less biased with respect to a mass selected sample. Overall, SZ selection seems to provide the most unbiased selection for all masses and redshifts. Richness is relatively unbiased up to $z = 1$, but it suffers the most from contamination by smaller haloes. We also find that contamination to a biased temperature and gas fraction for the smallest haloes in the sample when selecting based on SZ or X-rays.

- In Chapter 4, we assess how well standard cluster cosmology modelling is able to recreate the results of the FLAMINGO simulations. To investigate this, we create three surveys, similar to the Planck, SPT and future Simon's observatory SZ-selected cluster catalogues. We find that for future surveys, baryonic effects on the HMF will need to be accounted for. Furthermore, a single power-law mass-observable relation might lead to slight biases for future surveys. Additionally, we further confirm the fact that not all current HMF models in use have the required accuracy to reproduce the FLAMINGO simulation, which can lead to problems even at the sensitivity of Planck.

- Finally, in Chapter 5, we investigate the dynamical influence of the different components of the cosmic web. By making use of the NEXUS+ cosmic web identifier, we split both the cosmic force and tidal field into the parts that are caused by filaments, voids, walls and nodes. Our results reveal that the bulk of the motion in the Universe is caused by the gravitational force exerted by filaments. However, when it comes to the smaller-scale structures, the forces are caused by the voids. Nodes only affect their immediate environment. We thus find that filaments drive the dynamics of the cosmic web, while voids organise its finer details.

1.5 Outlook

There are multiple interesting avenues that are still open for investigation that would build upon the results presented in this thesis. As already highlighted previously, the main effort should go into building forward modelling pipelines for the cluster observables investigated in this thesis. Given the large range of outputs generated by the FLAMINGO simulation suite, which importantly includes full-sky lightcones for many observables, forward modelling should be relatively straightforward.

Another avenue that should be explored, is further application of emulators for simulations. We are working on using emulators to design the baryonic variations. In this way we can create a grid of training simulations that vary both cosmology and subgrid physics. By making use of emulators, these baryonic variations can be parameterised in terms of observable parameters, like for example

the gas and stellar mass fraction in clusters. This will not only allow us to gain a better understanding of the coupling between cosmology and baryons, but also to put observational priors on the baryonic parameters when needed.

The true power will come when combining these methods. If we can create forward modelled predictions for each of the simulations in the training set of a baryonic cosmology hypercube, we can directly emulate the predictions for upcoming telescopes. This might even allow us to emulate observations more and more directly, removing the need for intermediate modelling of the observed quantities. This will not only greatly reduce the systematics in the results obtained from comparisons between theory and observation, it will also allow us to understand better where our modelling is still lacking.

References

Abbott T. M. C., et al., 2022, Phys. Rev. D, 105, 023520
Ade P., et al., 2019, J. Cosmology Astropart. Phys., 2019, 056
Akino D., et al., 2022, PASJ, 74, 175
Andrade-Santos F., et al., 2017, ApJ, 843, 76
Anglés-Alcázar D., et al., 2021, ApJ, 917, 53
Angulo R. E., Zennaro M., Contreras S., Aricò G., Pellejero-Ibañez M., Stücker J., 2021, MNRAS, 507, 5869
Aragón-Calvo M. A., van de Weygaert R., Jones B. J. T., 2010, MNRAS, 408, 2163
Aricò G., Angulo R. E., Contreras S., Ondaro-Mallea L., Pellejero-Ibañez M., Zennaro M., 2021, MNRAS, 506, 4070
Arnaud M., Pratt G. W., Piffaretti R., Böhringer H., Croston J. H., Pointecouteau E., 2010, A&A, 517, A92
Artis E., Melin J. B., Bartlett J. G., Murray C., Euclid Consortium 2022, in mm Universe @ NIKA2 - Observing the mm Universe with the NIKA2 Camera. p. 00004 (arXiv:2111.05432), doi:10.1051/epjconf/202225700004
Asgari M., Mead A. J., Heymans C., 2023, The Open Journal of Astrophysics, 6, 39
Bahé Y. M., et al., 2017, MNRAS, 470, 4186
Bahé Y. M., et al., 2022, MNRAS, 516, 167
Barnes D. J., Kay S. T., Henson M. A., McCarthy I. G., Schaye J., Jenkins A., 2017a, MNRAS, 465, 213
Barnes D. J., et al., 2017b, MNRAS, 471, 1088
Behroozi P., Wechsler R. H., Hearin A. P., Conroy C., 2019, MNRAS, 488, 3143
Berlind A. A., Weinberg D. H., 2002, ApJ, 575, 587
Bermejo R., Wilding G., van de Weygaert R., Jones B. J. T., Vegter G., Efstathiou K., 2024, MNRAS, 529, 4325
Bigwood L., et al., 2024, arXiv e-prints, p. arXiv:2404.06098
Black W. K., Evrard A., 2022, MNRAS, 516, 1170
Bleem L. E., et al., 2022, ApJS, 258, 36
Bleem L. E., et al., 2024, The Open Journal of Astrophysics, 7, 13
Bocquet S., Saro A., Dolag K., Mohr J. J., 2016, MNRAS, 456, 2361
Bocquet S., Heitmann K., Habib S., Lawrence E., Uram T., Frontiere N., Pope A., Finkel H., 2020, ApJ, 901, 5
Bocquet S., et al., 2024, arXiv e-prints, p. arXiv:2401.02075
Bond J. R., Kofman L., Pogosyan D., 1996, Nature, 380, 603
Booth C. M., Schaye J., 2009, MNRAS, 398, 53
Borgani S., et al., 2001, ApJ, 561, 13
Borrow J., Schaller M., Bower R. G., Schaye J., 2022, MNRAS, 511, 2367
Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645
Braspenning J., et al., 2023, arXiv e-prints, p. arXiv:2312.08277
Brown S. T., Fattahi A., McCarthy I. G., Font A. S., Oman K. A., Riley A. H., 2024, arXiv e-prints, p. arXiv:2403.11692
Bulbul E., et al., 2024, A&A, 685, A106

Cautun M., van de Weygaert R., Jones B. J. T., 2013, MNRAS, 429, 1286
Cautun M., van de Weygaert R., Jones B. J. T., Frenk C. S., 2014, MNRAS, 441, 2923
Chaikin E., Schaye J., Schaller M., Benítez-Llambay A., Nobels F. S. J., Ploeckinger S., 2022, arXiv e-prints, p. arXiv:2211.04619
Chisari N. E., et al., 2019, The Open Journal of Astrophysics, 2, 4
Chiu I. N., Klein M., Mohr J., Bocquet S., 2023, MNRAS,
Chon G., Böhringer H., 2017, A&A, 606, L4
Clowe D., Bradač M., Gonzalez A. H., Markevitch M., Randall S. W., Jones C., Zaritsky D., 2006, ApJ, 648, L109
Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168
Costanzi M., et al., 2019, MNRAS, 488, 4779
Crain R. A., et al., 2015, MNRAS, 450, 1937
Cui W., et al., 2018, MNRAS, 480, 2898
DESI Collaboration et al., 2024, arXiv e-prints, p. arXiv:2404.03002
van Daalen M. P., Schaye J., Booth C. M., Dalla Vecchia C., 2011, MNRAS, 415, 3649
Dalla Vecchia C., Schaye J., 2008, MNRAS, 387, 1431
Dalla Vecchia C., Schaye J., 2012, MNRAS, 426, 140
Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, MNRAS, 486, 2827
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371
Dawson K. S., et al., 2013, AJ, 145, 10
Debackere S. N. B., Schaye J., Hoekstra H., 2020, MNRAS, 492, 2285
Debackere S. N. B., Schaye J., Hoekstra H., 2021, MNRAS, 505, 593
Debackere S. N. B., Hoekstra H., Schaye J., Heitmann K., Habib S., 2022, MNRAS, 515, 3383
Dicke R. H., Peebles P. J. E., Roll P. G., Wilkinson D. T., 1965, ApJ, 142, 414
Ebeling H., Edge A. C., Bohringer H., Allen S. W., Crawford C. S., Fabian A. C., Voges W., Huchra J. P., 1998, MNRAS, 301, 881
Eckert D., et al., 2016, A&A, 592, A12
Efstathiou G., Sutherland W. J., Maddox S. J., 1990, Nature, 348, 705
Einstein A., 1916, Annalen der Physik, 354, 769
Einstein A., 1917, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin, pp 142–152
Eke V. R., Cole S., Frenk C. S., 1996, MNRAS, 282, 263
Elahi P. J., Cañas R., Poulton R. J. J., Tobar R. J., Willis J. S., Lagos C. d. P., Power C., Robotham A. S. G., 2019, PASA, 36, e021
Elbers W., Frenk C. S., Jenkins A., Li B., Pascoli S., 2022, MNRAS, 516, 3821
Elbers W., et al., 2024, arXiv e-prints, p. arXiv:2403.12967
Euclid Collaboration et al., 2019, MNRAS, 484, 5509
Fattahi A., et al., 2016, MNRAS, 457, 844
Feldmann R., et al., 2023, MNRAS, 522, 3831
Ferland G. J., et al., 2017, Rev. Mex. Astron. Astrofis., 53, 385
Frenk C. S., White S. D. M., Efstathiou G., Davis M., 1985, Nature, 317, 595
Friedmann A., 1922, Zeitschrift fur Physik, 10, 377
Friedmann A., 1924, Zeitschrift fur Physik, 21, 326
Fumagalli A., Costanzi M., Saro A., Castro T., Borgani S., 2024, A&A, 682, A148
Ganeshaiah Veena P., Cautun M., Tempel E., van de Weygaert R., Frenk C. S., 2019, MNRAS, 487, 1607
Ganeshaiah Veena P., Cautun M., van de Weygaert R., Tempel E., Frenk C. S., 2021, MNRAS, 503, 2280
Garrel C., et al., 2022, A&A, 663, A3
Ghirardini V., et al., 2024, arXiv e-prints, p. arXiv:2402.08458
Giles P. A., et al., 2022, MNRAS, 516, 3878
Giri S. K., Schneider A., 2021, arXiv e-prints, p. arXiv:2108.08863

Grand R. J. J., et al., 2017, MNRAS, 467, 179
Grandis S., et al., 2021, MNRAS, 504, 1253
Guth A. H., 1981, Phys. Rev. D, 23, 347
Hadzhiyska B., et al., 2023, MNRAS, 525, 4367
Hahn O., Abel T., 2011, MNRAS, 415, 2101
Hahn O., Rampf C., Uhlemann C., 2021, MNRAS, 503, 426
Han J., Jing Y. P., Wang H., Wang W., 2012, MNRAS, 427, 2437
Heitmann K., Higdon D., Nakhleh C., Habib a., 2008, The Astrophysical Journal Letters, 646, L1
Heitmann K., Higdon D., White M., Habib S., Williams B. J., Lawrence E., Wagner C., 2009, ApJ, 705, 156
Heitmann K., et al., 2016, ApJ, 820, 108
Henden N. A., Puchwein E., Shen S., Sijacki D., 2018, MNRAS, 479, 5385
Heymans C., et al., 2021, A&A, 646, A140
Hilton M., et al., 2018, ApJS, 235, 20
Hilton M., et al., 2021, ApJS, 253, 3
Hirschmann M., Dolag K., Saro A., Bachmann L., Borgani S., Burkert A., 2014, MNRAS, 442, 2304
Hockney R. W., Eastwood J. W., 1981, Computer Simulation Using Particles
Hoekstra H., Herbonnet R., Muzzin A., Babul A., Mahdavi A., Viola M., Cacciato M., 2015, MNRAS, 449, 685
Hopkins P. F., 2015, MNRAS, 450, 53
Howlett C., Lewis A., Hall A., Challinor A., 2012, J. Cosmology Astropart. Phys., 1204, 027
Hubble E., 1929, Proceedings of the National Academy of Science, 15, 168
Huško F., Lacey C. G., Schaye J., Schaller M., Nobels F. S. J., 2022, MNRAS, 516, 3750
Icaza-Lizaola M., et al., 2020, MNRAS, 492, 4189
Ishiyama T., Fukushige T., Makino J., 2009, PASJ, 61, 1319
Ishiyama T., et al., 2021, MNRAS, 506, 4210
Jenkins A., 2013, MNRAS, 434, 2094
Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, MNRAS, 321, 372
Jones B., van de Weygaert R., 2009, Astrophysics and Space Science Proceedings, 8, 467
Kaiser N., 1986, MNRAS, 222, 323
Kaiser N., 1991, ApJ, 383, 104
Kaviraj S., et al., 2017, MNRAS, 467, 4739
Kennicutt Robert C. J., 1998, ApJ, 498, 541
Kirby M., et al., 2019, arXiv e-prints, p. arXiv:1910.13548
Komatsu E., et al., 2003, ApJS, 148, 119
Lemaitre G., 1931, MNRAS, 91, 483
Lesgourgues J., 2011a, arXiv e-prints, p. arXiv:1104.2932
Lesgourgues J., 2011b, arXiv e-prints, p. arXiv:1104.2934
Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473
Linde A. D., 1982, Physics Letters B, 108, 389
Liu A., et al., 2022, A&A, 661, A2
Lovisari L., et al., 2017, ApJ, 846, 51
Maksimova N. A., Garrison L. H., Eisenstein D. J., Hadzhiyska B., Bose S., Satterthwaite T. P., 2021, MNRAS, 508, 4017
Mantz A. B., 2019, MNRAS, 485, 4863
Marini I., et al., 2024, arXiv e-prints, p. arXiv:2404.12719
McCarthy I. G., Schaye J., Bird S., Le Brun A. M. C., 2017, MNRAS, 465, 2936
McCarthy I. G., Bird S., Schaye J., Harnois-Deraps J., Font A. S., van Waerbeke L., 2018, MNRAS, 476, 2999
McCarthy I. G., et al., 2023, MNRAS, 526, 5494
Mead A. J., Brieden S., Tröster T., Heymans C., 2021, MNRAS, 502, 1401
Melin J. B., Bartlett J. G., Delabrouille J., 2006, A&A, 459, 341

Melin J. B., et al., 2012, A&A, 548, A51
Melin J. B., Bartlett J. G., Cai Z. Y., De Zotti G., Delabrouille J., Roman M., Bonaldi A., 2018, A&A, 617, A75
Miyatake H., et al., 2023, arXiv e-prints, p. arXiv:2304.00704
Moran K. R., et al., 2022, MNRAS,
Morris M. D., Mitchell T. J., 1995, Journal of Statistical Planning and Inference, 43, 381
Moster B. P., Somerville R. S., Maulbetsch C., van den Bosch F. C., Macciò A. V., Naab T., Oser L., 2010, ApJ, 710, 903
Nelson D., Pillepich A., Ayromlou M., Lee W., Lehle K., Rohr E., Truong N., 2023, arXiv e-prints, p. arXiv:2311.06338
Ota N., et al., 2023, A&A, 669, A110
Pakmor R., et al., 2023, MNRAS, 524, 2539
Penzias A. A., Wilson R. W., 1965, ApJ, 142, 419
Pereira M. E. S., 2021, in American Astronomical Society Meeting Abstracts. p. 501.03
Perlmutter S., et al., 1999, ApJ, 517, 565
Pierre M., et al., 2016, A&A, 592, A1
Pillepich A., et al., 2018, MNRAS, 473, 4077
Planck Collaboration et al., 2014, A&A, 571, A21
Planck Collaboration et al., 2016a, A&A, 594, A24
Planck Collaboration et al., 2016b, A&A, 594, A27
Planck Collaboration et al., 2020a, A&A, 641, A6
Planck Collaboration et al., 2020b, A&A, 641, A8
Planck Collaboration et al., 2020c, A&A, 641, A9
Ploeckinger S., Schaye J., 2020, MNRAS, 497, 4857
Potter D., Stadel J., Teyssier R., 2017, Computational Astrophysics and Cosmology, 4, 2
Press W. H., Schechter P., 1974, ApJ, 187, 425
Robertson H. P., 1935, ApJ, 82, 284
Robertson H. P., 1936, ApJ, 83, 257
Rosati P., Della Ceca R., Norman C., Giacconi R., 1998, ApJ, 492, L21
Ross A. J., et al., 2017, MNRAS, 464, 1168
Rossetti M., Gastaldello F., Eckert D., Della Torre M., Pantiri G., Cazzoletti P., Molendi S., 2017, MNRAS, 468, 1917
Rubin V. C., Ford W. K. J., Thonnard N., 1980, ApJ, 238, 471
Rykoff E. S., et al., 2014, ApJ, 785, 104
Rykoff E. S., et al., 2016, ApJS, 224, 1
Sacks J., Welch W. J., Mitchell T. J., Wynn H. P., 1989, Statist. Sci., 4, 409
Salcido J., McCarthy I. G., Kwan J., Upadhye A., Font A. S., 2023a, MNRAS, 523, 2247
Salcido J., McCarthy I. G., Kwan J., Upadhye A., Font A. S., 2023b, MNRAS, 523, 2247
Sarazin C. L., 1988, X-ray emission from clusters of galaxies
Schaller M., et al., 2023, arXiv e-prints, p. arXiv:2305.13380
Schaye J., 2004, ApJ, 609, 667
Schaye J., Dalla Vecchia C., 2008, MNRAS, 383, 1210
Schaye J., et al., 2010, MNRAS, 402, 1536
Schaye J., et al., 2015, MNRAS, 446, 521
Schaye J., et al., 2023, MNRAS, 526, 4978
Semboloni E., Hoekstra H., Schaye J., van Daalen M. P., McCarthy I. G., 2011, MNRAS, 417, 2020
Semboloni E., Hoekstra H., Schaye J., 2013, MNRAS, 434, 148
Sievers J. L., et al., 2013, J. Cosmology Astropart. Phys., 2013, 060
Smith G. P., et al., 2016, MNRAS, 456, L74
Sousbie T., 2013, arXiv e-prints, p. arXiv:1302.6221
Springel V., 2011, in Alves J., Elmegreen B. G., Girart J. M., Trimble V., eds, Vol. 270, Computational Star Formation. pp 203–206, doi:10.1017/S1743921311000378
Springel V., et al., 2005a, Nature, 435, 629

Springel V., Di Matteo T., Hernquist L., 2005b, ApJ, 620, L79
Springel V., Pakmor R., Zier O., Reinecke M., 2021, MNRAS, 506, 2871
Stadel J. G., 2001, PhD thesis, University of Washington, Seattle
Stevens A. R. H., Sinha M., Rohl A., Sammons M. W., Hadzhiyska B., Hernández-Aguayo C., Hernquist L., 2023, arXiv e-prints, p. arXiv:2312.04137
Story K. T., et al., 2013, ApJ, 779, 86
Sunyaev R. A., Zeldovich Y. B., 1972, Comments on Astrophysics and Space Physics, 4, 173
Sutherland R. S., Dopita M. A., 1993, ApJS, 88, 253
Teyssier R., 2002, A&A, 385, 337
Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709
Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, ApJ, 724, 878
Velliscig M., van Daalen M. P., Schaye J., McCarthy I. G., Cacciato M., Le Brun A. M. C., Dalla Vecchia C., 2014, MNRAS, 442, 2641
Villaescusa-Navarro F., et al., 2021, ApJ, 915, 71
Walker A. G., 1937, Proceedings of the London Mathematical Society, 42, 90
Weinberger R., Springel V., Pakmor R., 2020, ApJS, 248, 32
White S. D. M., Frenk C. S., Davis M., Efstathiou G., 1987, ApJ, 313, 505
Wiersma R. P. C., Schaye J., Theuns T., Dalla Vecchia C., Tornatore L., 2009, MNRAS, 399, 574
Willis J. P., et al., 2021, MNRAS, 503, 5624
Yates R. M., Henriques B. M. B., Fu J., Kauffmann G., Thomas P. A., Guo Q., White S. D. M., Schady P., 2021, MNRAS, 503, 4474
Zel'Dovich Y. B., 1970, A&A, 500, 13
Zubeldia Í., Chluba J., Battye R., 2023, MNRAS, 522, 5123
Zwicky F., 1933, Helvetica Physica Acta, 6, 110
de Lapparent V., Geller M. J., Huchra J. P., 1986, ApJ, 302, L1
van Daalen M. P., McCarthy I. G., Schaye J., 2020, MNRAS, 491, 2424
van Weeren R. J., de Gasperin F., Akamatsu H., Brüggen M., Feretti L., Kang H., Stroe A., Zandanel F., 2019, Space Sci. Rev., 215, 16

2 FLAMINGO: Calibrating large cosmological hydrodynamical simulations with machine learning

Authors: **Roi Kugel**, Joop Schaye, Matthieu Schaller, John C. Helly, Joey Braspenning, Willem Elbers, Carlos S. Frenk, Ian G. McCarthy, Juliana Kwan, Jaime Salcido, Marcel P. van Daalen, Bert Vandenbroucke, Yannick M. Bahé, Josh Borrow, Evgenii Chaikin, Filip Huko, Adrian Jenkins, Cedric G. Lacey, Folkert S. J. Nobels, and Ian Vernon

Published in Monthly Notices of the Royal Astronomical Society, Volume 526, Issue 4, December 2023

Abstract

To fully take advantage of the data provided by large-scale structure surveys, we need to quantify the potential impact of baryonic effects, such as feedback from active galactic nuclei (AGN) and star formation, on cosmological observables. In simulations, feedback processes originate on scales that remain unresolved. Therefore, they need to be sourced via subgrid models that contain free parameters. We use machine learning to calibrate the AGN and stellar feedback models for the FLAMINGO cosmological hydrodynamical simulations. Using Gaussian process emulators trained on Latin hypercubes of 32 smaller-volume simulations, we model how the galaxy stellar mass function and cluster gas fractions change as a function of the subgrid parameters. The emulators are then fit to observational data, allowing for the inclusion of potential observational biases. We apply our method to the three different FLAMINGO resolutions, spanning a factor of 64 in particle mass, recovering the observed relations within the respective resolved mass ranges. We also use the emulators, which link changes in subgrid parameters to changes in observables, to find models that skirt or exceed the observationally allowed range for cluster gas fractions and the stellar mass function. Our method enables us to define model variations in terms of the data that they are calibrated to rather than the values of specific subgrid parameters. This approach is useful, because subgrid parameters are typically not directly linked to particular observables, and predictions for a specific observable are influenced by multiple subgrid parameters.

2.1 Introduction

2

The evolution of the large-scale distribution of matter in the Universe is highly sensitive to the underlying cosmological model. Current probes have given us our concordance cosmological model ΛCDM, which consists of a spatially flat universe, where dark energy and cold dark matter dominate the current energy density (for a review see Frieman et al., 2008).

The concordance model has been independently validated by a large array of probes. These include the cosmic microwave background (CMB) (e.g. Planck Collaboration et al., 2020), galaxy clustering and gravitational lensing (e.g. Abbott et al., 2022; Heymans et al., 2021), baryon acoustic oscillations (BAO) (e.g. Alam et al., 2021), and more (for a review see Turner, 2022). While all the probes broadly agree with the ΛCDM model, tensions remain between early universe probes, like the CMB, and late-time probes, like the distance ladder and weak lensing. For the H_0 and σ_8 parameters, the tension is at the level of a few standard deviations (e.g. Heymans et al., 2021; Abbott et al., 2022; Riess et al., 2022). Next generation surveys like *Euclid*[1] and LSST[2] will measure the matter power spectrum to per cent level accuracy (Euclid Collaboration et al., 2020). The results from these surveys will provide us with a stringent test of the concordance model, and show us whether these tensions will force us to modify the ΛCDM model.

Most of the modelling work for large-scale structure is done with collisionless N-body simulations (e.g. Heitmann et al., 2016a; DeRose et al., 2021; Euclid Collaboration et al., 2019). N-body simulations model the evolution of cold dark matter and can accurately predict the structure and clustering of dark matter haloes under the effect of gravity only. The dark part of the matter component is dominant in mass and hence, predictions from these simulations may provide stringent cosmological constraints. However, baryons change the distribution of dark matter through back reaction effects, but, with the exception of gravitational lensing, we are limited to observing the imprint of the distribution of dark matter on the baryonic matter. Most of the baryonic matter is found in the tenuous intergalactic medium (e.g. Nicastro et al., 2018; Macquart et al., 2020), which is very challenging to observe directly. Large-scale structure surveys use galaxies, which are located within dark matter haloes, to map the distribution of matter.

Sophisticated semi-analytical and semi-empirical models can make predictions for how galaxies evolve within their dark matter haloes (e.g. Cole et al., 2015; Lacey et al., 2016; Moster et al., 2018; Behroozi et al., 2019; Ayromlou et al., 2021). Baryonic effects can be simulated with halo models (e.g. Semboloni et al., 2011, 2013; Mead et al., 2015; Debackere et al., 2020; Acuto et al., 2021), added to N-body simulations by baryonification algorithms (e.g. Schneider & Teyssier,

[1] https://www.euclid-ec.org/
[2] https://www.lsst.org/

2015; Giri & Schneider, 2021; Aricò et al., 2021) or included as a parametric correction to the matter power spectrum (Van Daalen et al., 2020; Salcido et al., 2023). However, the most self-consistent way to model how the large-scale structure is coupled with baryons, is via large cosmological hydrodynamical simulations. Modern simulations like Magneticum (Hirschmann et al., 2014), EAGLE (Schaye et al., 2015; Crain et al., 2015), Horizon-AGN (Kaviraj et al., 2017), IllustrisTNG (Pillepich et al., 2018), BAHAMAS (McCarthy et al., 2017, 2018), SIMBA Davé et al. (2019) and MilleniumTNG (Pakmor et al., 2022) provide predictions for the interplay between galaxy formation and the large-scale structure. The results from hydrodynamical simulations can also inform the simpler parametric and analytic models.

One of the main difficulties for hydrodynamical simulations is the implementation and tuning of relevant astrophysical processes that originate on unresolved scales through subgrid physics. Processes like star formation and black hole growth occur on parsec scales, and are not resolved. The resulting feedback from stars and active galactic nuclei (AGN), do influence the distribution of matter on cosmological scales (Van Daalen et al., 2011, 2020; Debackere et al., 2020; Schneider et al., 2020). Therefore, we need to create simulations that model their effect on the resolved scales.

Subgrid physics models are characterised by a set of free parameters, in the sense that there is both uncertainty in the processes we try to model and uncertainty in how the models are affected by numerical limitations. An example of the latter is the impact of numerical over-cooling on galactic wind models (see Dalla Vecchia & Schaye, 2012). The numerical effects combined with the general non-linearity of galaxy formation makes it difficult to implement subgrid physics based solely on first principles. Instead, we have to calibrate the model by comparing it to a selection of observations, a partial forfeit of their predictive power. As argued by Schaye et al. (2015), this is a necessary sacrifice. By ensuring certain relations are reproduced, the simulation retains predictive power for other relations. Calibrating subgrid physics forces us to find a balance between how many observables one tries to match and how many of the results can be deemed predictions.

In this paper we discuss the calibration strategy used for the low-, intermediate- and high-resolution simulations of the FLAMINGO project (Full-hydro Large-scale structure simulations with All-sky Mapping for the Interpretation of Next Generation Observations; Schaye et al. 2023). The intermediate-resolution FLAMINGO model has the same resolution ($m_{\rm gas} = 1.07 \times 10^9$ M$_\odot$) as used for the BAHAMAS project (McCarthy et al., 2017, 2018), but in a volume of (2.8 Gpc)^3. This volume is over two orders of magnitude larger than BAHAMAS. Additionally, FLAMINGO includes a suite of feedback and cosmology variations in (1 Gpc)^3 volumes. This includes a high ($m_{\rm gas} = 1.34 \times 10^8$ M$_\odot$) and a low ($m_{\rm gas} = 8.56 \times 10^9$ M$_\odot$) resolution variation.

2

Our goal is to expand the large-scale structure science of the BAHAMAS project to larger volumes, different resolutions, and more cosmology and astrophysics variations with a new code and an improved subgrid physics model. The FLAMINGO simulation outputs also include on-the-fly full sky lightcones, both as particles and as maps, for a variety of observables. Similarly to BAHAMAS, we will calibrate to the observed present-day galaxy stellar mass function (SMF) and the gas fractions in groups and clusters of galaxies ($f_{\rm gas}$). We opt for the SMF to ensure we can reproduce galaxy clustering and lensing statistics if we use the correct cosmology. The gas fraction is used to ensure we have a realistic distribution of gas in and around clusters, which is not only important for cluster cosmology, but also for baryonic effects on the matter power spectrum (Semboloni et al., 2011; Schneider & Teyssier, 2015; Debackere et al., 2020; Van Daalen et al., 2020; Aricò et al., 2021; Salcido et al., 2023). While our fiducial models are calibrated to reproduce the data, we also calibrate the subgrid physics to the gas fraction and SMF data that has been shifted relative to the observed values. These feedback variations will enable future FLAMINGO projects to test the importance of astrophysical effects constrained by the uncertainties in the data.

For BAHAMAS, and also for simulations like EAGLE and IllustrisTNG, calibration was done by hand by varying the subgrid parameters within some reasonable range until the simulation lined up with the calibration targets. This approach works reasonably well in the context of galaxy formation, but it introduces biases into the parameter selection. For cosmology applications we require a more systematic and controlled approach. We want to be able to sample the parameter space with a Markov Chain Monte Carlo (MCMC) method and to find the posterior probabilities of each of the subgrid parameter values. This approach also allows us to take into account potential systematic effects in the data and/or simulations.

Because N-body simulations are too computationally expensive to be used directly in MCMC-like methods, we make use of machine learning, specifically emulation using Gaussian processes. While it is too expensive to run a new simulation for each MCMC step, we can train an emulator on a carefully sampled selection of input simulations. The emulator then gives us the predicted observable as a continuous function of the input parameters, which can be fed into any likelihood calculation code. Emulator-based methods have been used in combination with semi-analytic models of galaxy formation (Bower et al., 2010; Vernon et al., 2014; Rodrigues et al., 2017; Elliott et al., 2021) and have become particularly popular for cosmology. By training emulators on dark-matter-only simulations, their full non-linear matter power spectrum can be predicted with per cent level precision (e.g. Heitmann et al., 2009, 2016b; Euclid Collaboration et al., 2019; Angulo et al., 2021; Moran et al., 2022).

We directly emulate our calibration targets: the SMF and the gas fractions

in groups and clusters. This allows us to create a continuous simulation-based model that can be compared with observations. With the emulator we can use MCMC to directly fit the subgrid physics parameters to the observational data, while modelling statistical and systematic errors in both the simulations and the data. This procedure not only gives us a well-calibrated model, but also lets us determine the maximum variations allowed by the model. In this way our resulting simulations can provide upper and lower limits on the expected baryonic effects. More general machine learning techniques have been used to calibrate hydrodynamical simulations. Jo et al. (2023) calibrate to baryonic observables in the $(25\ \mathrm{Mpc})^3$ volumes of the CAMELS project (Villaescusa-Navarro et al., 2021) and Oh et al. (2022) apply a similar methodology to zooms of Milky Way haloes. However, these methods have not been applied to simulations of large cosmological volumes and they have not accounted for possible observational biases.

This paper is structured as follows. In Section 2.2 we describe the most relevant aspects of our simulation method and galaxy formation models. In Section 2.3 the reasoning for our calibration targets is explained, and we describe our compilation of data and how we include potential observational and simulation-originated biases in our analysis. In Section 2.4 we describe how we obtain the training data for the emulators. We also discuss how the emulators are trained and how we estimate the uncertainty in the predictions of the emulators. We describe our likelihoods and our fitting method in Section 2.5. In Section 2.6 we show the results of fitting the emulators at the three FLAMINGO resolutions. We also discuss how the emulators can be used to better understand subgrid physics using parameter sweeps and we use the emulator to find models that skirt or exceed the observational allowed range for the cluster gas fractions and the SMF. Finally, we summarise our method, strategy and results in Section 2.7. In this work, R_{500c} is defined as the radius within which the mean internal density is 500 times the critical density. The radius R_{500c} also defines M_{500c}, which is the mass inside R_{500c}.

2.2 Simulations

The simulation methods and galaxy formation model are described in detail in Schaye et al. (2023). Here we will provide a summary of the most relevant aspects. We describe in more detail the subgrid prescriptions that we calibrate in this work, namely those for stellar feedback (§2.2.1), the growth of supermassive black holes (§2.2.2), and AGN feedback (§2.2.3), and we will motivate the choice of priors for the subgrid parameters that are varied (these are listed in Table 2.2).

All simulations in this work use the open-source code Swift (Schaller et al., 2023). Swift is an N-body gravity and smooth particle hydrodynamics (SPH) solver that makes use of a fine-grained tasking framework and runs across multiple compute nodes using MPI. Gravity is solved using the Fast Multiple Method

Table 2.1: Numerical characteristics of the final Latin hypercubes of simulations. The columns list: the resolution qualifier, comoving box size, number of particles (there are initially equal numbers of dark matter and baryonic particles), initial baryonic particle mass, dark matter particle mass, comoving gravitational softening length, maximum physical gravitational softening length.

Resolution	L (cMpc)	N	m_g ($M_\odot$)	m_{DM} ($M_\odot$)	ϵ_{com} (ckpc)	ϵ_{prop} (pkpc)
Low [m10]	400	2×360^3	8.56×10^9	4.52×10^{10}	44.6	11.40
Intermediate [m9]	200	2×360^3	1.07×10^9	5.65×10^9	22.3	5.70
High [m8]	100	2×360^3	1.34×10^8	7.06×10^8	11.2	2.85

(Greengard & Rokhlin, 1987). We use the SPHENIX SPH scheme (Borrow et al., 2022b) with a Wendland (1995) C^2 kernel. Massive neutrinos are implemented into SWIFT via the δf method of Elbers et al. (2021).

Initial conditions are generated using a modified version of MONOFONIC (Hahn et al., 2021) that includes massive neutrinos. We use unperturbed initial conditions for the neutrino particles. We do not include large scale neutrino perturbations in the initial conditions, as these have a negligible effect in the small box sizes used for this work. We adopt the '3x2pt + all' cosmology from Abbott et al. (2022) ($\Omega_m = 0.306$, $\Omega_b = 0.0486$, $\sigma_8 = 0.807$, $H_0 = 68.1$, $n_s = 0.967$) with a minimal neutrino mass of 0.06 eV. The particle masses and gravitational softening lengths corresponding to the three different resolutions that we will consider are listed in Table 2.1.

For simulations with volumes as large as FLAMINGO, it is currently impossible to resolve all the processes that are important for galaxy formation. Therefore, we make use of subgrid models. FLAMINGO builds upon the models of OWLS (Schaye et al., 2010), used for Cosmo-OWLS (Le Brun et al., 2014), BAHAMAS (McCarthy et al., 2017), and EAGLE (Schaye et al., 2015), ported from the code GADGET (Springel, 2005) to SWIFT.

We use the radiative cooling tables from Ploeckinger & Schaye (2020), which are based on photo-ionisation models run with CLOUDY (Ferland et al., 2017) that include both the metagalactic and interstellar radiation fields, and that account for self-shielding, dust, and cosmic rays.

As we are unable to resolve the multiphase interstellar medium, we follow Schaye & Dalla Vecchia (2008) and impose a temperature floor. The pressure of gas with hydrogen number densities $n_H > 10^{-4}$ cm^{-3} and an overdensity greater than 100 is limited from below to $P/k_B = 800$ K $(n_H/10^{-4}$ cm$^{-3})^{4/3}$, where k_B is the Boltzmann constant.

During the simulation gas particles can be stochastically converted into star particles following the description of Schaye & Dalla Vecchia (2008). Particles

with total hydrogen number density[3] $n_{\rm H} > 10^{-1}$ cm^{-3}, an overdensity > 10 and within 0.3 dex of the temperature floor are stochastically allowed to convert into stars with a probability given by the particle's star formation rate,

$$\dot{m}_* = m_{\rm g} A (1\ {\rm M}_\odot {\rm pc}^{-2})^{-n} \left(\frac{\gamma}{G} f_{\rm g} P \right)^{(n-1)/2}, \tag{2.1}$$

where $m_{\rm g}$ is the gas particle mass, $\gamma = 5/3$ is the adiabatic index, and G is the gravitational constant. The star formation rate is derived such that self-gravitating discs reproduce the observed Kennicutt-Schmidt relation (Kennicutt Jr., 1998; Kennicutt Jr. et al., 2007). We assume the gas fraction, $f_{\rm g}$, is unity, $A = 1.515 \times 10^{-4}$ M$_\odot$ yr^{-1} pc^{-2}, and $n = 1.4$.

For the low-resolution simulation we were forced to relax the star formation parameters, as the default prescription was unable to form enough stars, even in large haloes and without stellar feedback. For low resolution, all particles with density $n_{\rm H} > 10^{-3}$ cm^{-3}, overdensity > 10 and temperature $T < 10^5$ K are star forming.

Each stellar particle is treated as a simple stellar population with a Chabrier (2003) initial mass function (IMF). Following Wiersma et al. (2009), we model stellar mass loss and track the abundances of the individual elements H, He, C, N, O, Ne, Mg, Si, and Fe. We also include type Ia supernova with rates taken from Schaye et al. (2015).

[3] Due to a bug, in the intermediate-resolution simulations gas particles with a metallicity equal to exactly zero were only allowed to form stars at densities higher than 10 cm^{-3}. This had little to no effect on any results at resolved stellar masses, but it did reduce the number of stars formed in the lowest-mass galaxies. Fixing this bug would potentially have allowed us to match the SMF to stellar masses corresponding to fewer than 10 particles.

Table 2.2: Priors and best-fitting values for the subgrid parameters for each of the three simulation resolutions. Low-resolution simulations do not include stellar feedback. The rows titled 'Median+CL' give the median and the 16th and 84th percentile confidence level (CL) obtained from the posterior of the fits. The rows titled 'best-fitting' list the maximum likelihood value from the fitting, which is our fiducial value. The last row 'Log' indicates whether the parameter is sampled logarithmically. The best-fitting values for the jet model are listed in Table 2.8 and the priors for the jet model are listed in Table 2.9.

Resolution	Parameter	$f_{\rm SN}$	$\Delta v_{\rm SN}$	$\log_{10}\Delta T_{\rm AGN}$ [K]	$\beta_{\rm BH}$
High-res [m8]	Prior	$[0.2, 0.9]$	$[80, 400]$	$[7.7, 8.9]$	$[0.0, 0.9]$
	Median+CL	$0.56^{0.15}_{-0.12}$	169^{+87}_{-65}	$8.03^{+0.13}_{-0.14}$	$0.23^{0.20}_{-0.15}$
	best-fitting	0.524	259	8.07	0.038
Intermediate-res [m9]	Prior	$[0, 0.5]$	$[200, 800]$	$[7.5, 8.5]$	$[0.1, 0.9]$
	Median+CL	$0.20^{+0.11}_{-0.09}$	479^{+167}_{-197}	$7.84^{+0.18}_{0.20}$	$0.55^{+0.15}_{-0.16}$
	best-fitting	0.238	562	7.95	0.514
Low-res [m10]	Prior	-	-	$[7, 9.5]$	$[0, 3]$
	Median+CL	-	-	$8.26^{+0.15}_{-0.15}$	$0.50^{+0.17}_{-0.16}$
	best-fitting	-	-	8.29	0.373
	Log	No	Yes	Yes	No

2.2.1 Stellar feedback

Although we will often refer to stellar feedback as supernova feedback, it may also represent other sources of energy released by massive stars that are unresolved by our simulations such as stellar winds, radiation pressure or cosmic rays.

Stellar feedback is implemented kinetically. The energy budget is normalised to the expected kinetic energy from core collapse supernovae, assuming that each star with a mass between 8 and 100 $M_\odot$ injects 10^{51} erg of kinetic energy into its surrounding medium. A fraction $f_{\rm SN}$ of this energy is assumed to be coupled to the ISM on scales resolved by the simulation and is used to kick neighbouring gas particles with a target velocity $\Delta v_{\rm SN}$. We use the method of Chaikin et al. (2022a)[4] to inject the kinetic energy in a statistically isotropic manner while ensuring that both momentum and energy are conserved. Note that if the relative velocities between the star and gas particles are nonzero, energy conservation results in differences between the actual and target kick velocities.

Following Dalla Vecchia & Schaye (2008) and Richings & Schaye (2016), we inject the kinetic energy probabilistically during each time step after the star particle has formed. The probability that a star particle kicks a given SPH neighbour is

$$p_{\rm kick}(f_{\rm SN}, \Delta v_{\rm SN}, m_{\rm ngb}, t, \Delta t) = 2\frac{f_{\rm SN}\Delta E_{\rm SNII}(t, \Delta t)}{m_{\rm ngb}\Delta v_{\rm SN}^2}, \tag{2.2}$$

where $\Delta E_{\rm SN}$ denotes the amount of energy released by the star particle of age t during a time step Δt and $m_{\rm ngb}$ is the total gas mass in the star particle's SPH kernel. The feedback efficiency, $f_{\rm SN}$, and the target kick velocity $\Delta v_{\rm SN}$ are the two stellar feedback parameters that are varied during the calibration.

The effect of stellar feedback generally scales with $f_{\rm SN}$, which sets the amount of energy that is injected. Based on the calibration of BAHAMAS (McCarthy et al., 2017) and after some experimentation with runs in which we varied only one parameter, we settled on prior ranges of $0.2 - 0.9$ and $0 - 0.5$ for high- and intermediate-resolution, respectively. The low-resolution simulations do not require any stellar feedback at all because of the strong suppression of star formation due to the limited resolution and because galaxies in the regime where stellar feedback dominates (stellar mass $M_* \ll 10^{11}\,M_\odot$) are only sampled by $\lesssim 10$ stellar particles.

If the kick velocity is too small, then stellar feedback ceases to be effective because of excessive radiative losses caused by the too-low post-shock temperatures (the well-known numerical over-cooling problem, see Dalla Vecchia & Schaye, 2012) and/or because the velocities are small compared to the escape velocities.

[4] There is one difference w.r.t. the method described by the authors. In the case where a particle would be kicked twice in a single time step, which we do not allow, we put the unused kick energy in a thermal dump, instead of adding it back to the star's feedback energy reservoir.

The lower limits for $\Delta v_{\rm SN}$ are 80 and 200 km s^{-1} for the high- and intermediate-resolution simulations, respectively. Our additional tests showed that for lower velocities the kicks stopped having a significant effect.

If the kick velocity is too large, then the feedback becomes poorly sampled, thus limiting its effectiveness. Our aim is to calibrate the SMF down to masses corresponding to just a few stellar particles. The expectation value for the number of kicks imparted by a single stellar particle is given by Chaikin et al. (2022a)

$$\langle N_{\rm kicks,\ SN}\rangle = 1.85\left(\frac{f_{\rm SN}}{0.25}\right)\left(\frac{\Delta v_{\rm SN}}{400\ {\rm km\ s^{-1}}}\right)^{-2}, \tag{2.3}$$

where we assumed the stellar and gas particles to have the same mass. Based on the above considerations and some small test runs, we limit the maximum kick velocity to 400 and 800 km s^{-1} for the high- and intermediate-resolution simulations, respectively. This implies $\langle N_{\rm kicks,\ SN}\rangle \approx 2$ and $\langle N_{\rm kicks,\ SN}\rangle \approx 0.4$ for high- and intermediate-resolution respectively. There should be at least four kicks for objects with 10 stellar particles at each resolution.

2.2.2 Black hole growth

Following Di Matteo et al. (2008) and Booth & Schaye (2009) we seed haloes with black holes (BHs) during the simulation. Starting at $z = 19$ we run a friends of friends group finder every time the expansion factor increases by a factor 1.00751. We seed a BH in every group that is above a certain mass threshold and that does not already have a BH. We seed BHs in haloes above a mass of $2.757 \times 10^{11}\ {\rm M}_\odot (m_{\rm g}/1.07 \times 10^9\ {\rm M}_\odot)$, corresponding to roughly fifty dark matter particles at each resolution. Because the Bondi & Hoyle (1944) accretion rate is proportional to the square of the BH mass, an increase in initial mass can cause BHs to grow much earlier. We use a BH seed mass of $10^5\ {\rm M}_\odot$ for intermediate and high resolution, and of $10^7\ {\rm M}_\odot$ for low resolution. The seed mass had to be increased for low resolution, since the rapid growth phase of the BHs corresponds to unresolved galaxy masses (see e.g. Bower et al., 2017; McAlpine et al., 2018).

As we do not properly resolve dynamical friction at our resolution, BHs are repositioned by hand to the minimum of the gravitational potential following the method of Bahé et al. (2022)[5]. For BH mergers we also follow the prescription by Bahé et al. (2022).

[5]The exclusion of the BH from the calculation of the gravitational potential used for repositioning was only done for high and low resolution, as we only became aware of its importance later. This significantly strengthened the quenching of star formation in galaxies with large stellar masses for our high resolution simulations.

Besides merging with other BHs, BHs grow via accretion of gas, which is assumed to occur at a modified Bondi-Hoyle rate,

$$\dot{m}_{\rm accr} = \alpha \frac{4\pi G c^2 m_{\rm BH}^2 \rho}{\left(c_{\rm s}^2 + v_{\rm BH}^2\right)^{3/2}}, \tag{2.4}$$

where $m_{\rm BH}$ is the BH mass, $c_{\rm s}$ is the sound speed of the gas, ρ is the gas density, c is the speed of light and $v_{\rm BH}$ is the velocity of the BH with respect to its environment. The factor α is a boost factor that is added because we do not resolve the Bondi radius and because we lack the resolution to model the phase structure of the ISM. We use the parametrization of Booth & Schaye (2009),

$$\alpha = \max\left[\left(\frac{n_{\rm H}}{n_{\rm H,*}}\right)^{\beta_{\rm BH}}, 1\right], \tag{2.5}$$

where $n_{\rm H,*} = 0.1$ cm^{-3}, which corresponds to the density threshold for star formation in the intermediate- and high-resolution simulations (we use the same value for all resolutions). The logarithmic density slope $\beta_{\rm BH}$ is a free parameter that we vary during the calibration. After some experimentation using simulations where only a single parameter is varied between runs, we settled on priors of $0-0.9$, $0.1-0.9$ and $0-3$ for high , intermediate and low resolution, respectively.

The gas accretion rate is capped at the Eddington (1913) rate. Following Bahé et al. (2022), the BH is allowed to 'nibble' on neighbouring gas particles until the gas particles only have half of their original mass remaining.

2.2.3 AGN feedback

In all but two of the simulations AGN feedback energy is injected into the medium surrounding the BH in thermal form using the prescription from Booth & Schaye (2009). The model used in the remaining simulations is based on jet feedback and is described in §2.2.3.

While accreting gas, the BH adds a fraction $\epsilon_{\rm r}\epsilon_{\rm f} = 0.015$ of the accreted rest mass energy to an internal feedback energy reservoir, where $\epsilon_{\rm r} = 0.1$ is the assumed radiative efficiency and $\epsilon_{\rm f} = 0.15$ is the assumed AGN feedback efficiency, i.e. the fraction of the radiated energy that is coupled to the gas surrounding the BH. Once enough energy is available to increase the temperature of $n_{\rm heat}$ gas particles by $\Delta T_{\rm AGN}$, this energy is injected into the neighbouring gas particles. The energy injected in a single event is proportional to $n_{\rm heat}\Delta T_{\rm AGN}$, where $\Delta T_{\rm AGN}$ is the increase in temperature that is applied to $n_{\rm heat}$ neighbours. We find that it is the product $n_{\rm heat}\Delta T_{\rm AGN}$ that is most important for regulating how much gas is expelled from clusters, and that $\Delta T_{\rm AGN}$ and $n_{\rm heat}$ are largely degenerate. We therefore fix $n_{\rm heat}$ to one and use $\Delta T_{\rm AGN}$ as a free parameter that is varied in the

calibration. Following the findings by Chaikin et al. (2022b), we inject the thermal energy into the nearest neighbour of the BH, which gives results that are nearly indistinguishable from a statistically isotropic approach.

To choose the prior for $\Delta T_{\rm AGN}$ we take a similar approach as for the stellar feedback kick velocity. However, instead of avoiding velocities that are too low to have an effect, we now have to make sure that feedback raises the temperature to a value sufficiently high to avoid catastrophic numerical over-cooling. The sampling issue is also slightly different than for stellar feedback. While stellar feedback is limited to young stars, BHs can inject energy throughout their lives and hence the time sampling of these events becomes important. If the time between AGN feedback events becomes too long, then the BHs will be unable to self-regulate. If BHs cannot regulate their growth, then this can lead to an unrealistic mass distribution of both the BHs and their host galaxies. To summarise, we have two main considerations:

1. What is the $\Delta T_{\rm AGN}$ below which radiative losses are already severe at injection for the densities at which stars form?

2. What is the $\Delta T_{\rm AGN}$ above which the time between AGN events becomes longer than the BH growth time?

Dalla Vecchia & Schaye (2012) demonstrated that the density above which thermal feedback becomes ineffective can be predicted based on the ratio of the radiative cooling time, which depends on the density and temperature, and the sound crossing time across a resolution element, which depends on the numerical resolution. According to their equation 18, feedback becomes inefficient for densities exceeding

$$n_{{\rm H},t_c} = 0.25\ {\rm cm}^{-3} \left(\frac{\Delta T_{\rm AGN}}{10^{7.5}\ {\rm K}}\right)^{3/2} \left(\frac{m_{\rm g}}{1.09\times10^9\ {\rm M}_\odot}\right)^{-1/2}. \tag{2.6}$$

Comparing this to our threshold for star formation ($n_{\rm H} = 10^{-1}\,{\rm cm}^{-3}$ for intermediate/high resolution and $10^{-3}\,{\rm cm}^{-3}$ for low resolution), yields minimum values of $\log_{10}\Delta T_{\rm AGN}/{\rm K} = 6.9$, 7.2, and 6.2 for the high, intermediate, and low resolution, respectively. However, the above equation assumes radiative losses to be dominated by Bremsstrahlung and Dalla Vecchia & Schaye (2012) showed that it underestimates the radiative losses for $\Delta T_{\rm AGN} < 10^7\,{\rm K}$. For this reason we do not consider values below $10^7\,{\rm K}$. On the other hand, since we inject the energy at the end of the time step, the feedback can do work during a single time step even if the temperature is too low to avoid overcooling, which means that somewhat lower values than implied by the above equation (but still higher than $10^7\,{\rm K}$) may still be of interest.

If we define $\Delta m_{\rm BH}$ to be the gas mass that must be accreted for the BH to have sufficient energy to heat a single gas particle, then the ratio of the time between

AGN feedback events and the time of BH growth is given by (Booth & Schaye, 2009),

$$\frac{t_{\rm AGN}}{t_{\rm BH}} = \frac{\Delta m_{\rm BH}/\dot{m}_{\rm BH}}{m_{\rm BH}/\dot{m}_{\rm BH}} \tag{2.7}$$

$$= \frac{m_{\rm g} k_{\rm B}(1-\epsilon_{\rm r})}{(\gamma-1)\mu m_{\rm H}\epsilon_{\rm f}\epsilon_{\rm r}c^2}\frac{n_{\rm heat}\Delta T_{\rm AGN}}{m_{\rm BH}} \tag{2.8}$$

$$\approx 0.98\left(\frac{1-\epsilon_r}{0.9}\right)\left(\frac{m_{\rm g}}{1.09\times10^9{\rm M}_\odot}\right)\left(\frac{\epsilon_{\rm f}\epsilon_{\rm r}}{0.015}\right)^{-1}\times \left(\frac{n_{\rm heat}\Delta T_{\rm AGN}}{10^{8.5}\ {\rm K}}\right)\left(\frac{m_{\rm BH}}{10^7\,{\rm M}_\odot}\right)^{-1}, \tag{2.9}$$

where $\gamma = 5/3$ is the ratio of specific heats and $\mu = 0.6$ is the mean particle mass in units of the proton mass $m_{\rm H}$. Given that we expect to need AGN feedback to quench star formation in galaxies with stellar mass $M_* \gtrsim 10^{11}\,{\rm M}_\odot$ and that in this mass range BHs are observed to have masses $M_{\rm BH} \sim 10^{-3} M_*$ (Häring & Rix, 2004), we need the BHs to become self-regulating when $M_{\rm BH} \ll 10^8\,{\rm M}_\odot$. The condition $t_{\rm AGN} < t_{\rm BH}$ then implies that for our $n_{\rm heat} = 1$ we require $\Delta T_{\rm AGN} \lesssim 10^{8.5}$ K for intermediate resolution, and values 8 times higher (lower) for high (low) resolution.

Based on the above considerations and some small test runs, we adopted the flat priors $\log_{10}\Delta T_{\rm AGN}/{\rm K} = 7.7-8.9$, $7.5-8.5$, and $7.0-9.5$ for high, intermediate and low resolution, respectively. For both intermediate and high resolution the prior ranges are smaller than what is possible based on our considerations. From our test runs we found that these ranges bracket a sufficiently large range in the observables we are interested in and the smaller ranges lead to slightly better sampling of the parameter space around the best-fitting model. For low resolution the prior extends to (unnecessarily) high values, but we will see that the best-fitting value is actually similar to those for the other resolutions. We can afford a larger prior range for the low resolution simulations as we are only sampling two parameters.

Jet feedback

In addition to the fully thermal AGN feedback scheme described above, we also calibrate a kinetic AGN feedback variation. The model used for kinetic AGN feedback is based on the spin-driven jet feedback model described by Huško et al. (2022), implemented into SWIFT. In this model energy is injected by kicking two particles on opposite sides of the BH, according to its angular momentum vector. The angular momentum of the BH is calculated in a subgrid model for an accretion disc that is based on general relativistic magneto-hydrodynamics simulations of single BHs in the low accretion regime (< 0.01 Eddington). For more

details see Huško et al. (2022). The spin from black holes that remains after mergers is computed according to the description by Rezzolla et al. (2008).

Due to the relatively low resolutions used for FLAMINGO, we make some simplifications to the complete model. As we intend for the jet model to be maximally different from the thermal feedback mode, we do not switch from kinetic to thermal feedback at high Eddington rates, and instead use the kinetic feedback at all accretion rates. Instead of using the efficiencies based on the subgrid accretion model, we fix the jet efficiency to $\epsilon = 0.015$. This efficiency is equal to the combined coupling and radiative efficiency, $\epsilon_{\rm f}\epsilon_{\rm r}$, for the thermal mode feedback. This implies that for each unit of mass accreted by the BH, the same amount of energy becomes available in the jet model as for the fiducial thermal model. While we do not use a spin-dependent feedback efficiency, we do still use the subgrid model to track the angular momentum vector of the BH and use it to select the direction in which gas particles are kicked. The BH accretion model is identical to that described in §2.2.2, and for calibration of the jet model we vary the boost factor $\beta_{\rm BH}$.

When the BH has accreted enough mass, two neighbouring gas particles are kicked with a total kinetic energy equal to

$$E_{\rm jet} = 2 \times \frac{1}{2} m_{\rm g} v_{\rm jet}^2, \tag{2.10}$$

where $v_{\rm jet}$ is the target jet velocity (we use the term target because it is the energy that is fixed, similarly to the supernova kicks, see §2.2.1), which is a free parameter that we calibrate. The jet velocity plays a role similar to $\Delta T_{\rm AGN}$ for the case of thermal feedback. As the energy is injected in kinetic form, the model is less affected by thermal losses, but picking velocities that are too low will make the gas unable to escape to large distances (see Huško et al., 2022). For very high values we again run into sampling issues. Based on these considerations and some initial tests, we use flat priors over the range of $v_{\rm jet}/({\rm km\ s^{-1}}) = 10^{2.7} - 10^{3.5}$, corresponding in energy to $\Delta T_{\rm AGN}/{\rm K} \approx 10^{7.1} - 10^{8.7}$. We only calibrate this model at intermediate resolution.

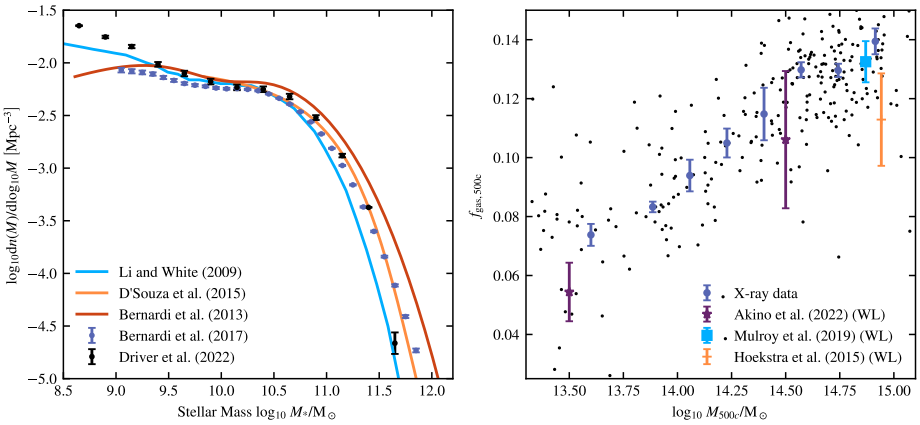


Figure 2.1: Compilation of observational data used for calibration. On the left we plot the SMF. On the right we plot the cluster gas fraction versus total mass, both measured at R_{500c}. Where available we display the 1σ measurement errors, which do not include intrinsic scatter. The X-ray data are binned from a compilation of available data, see §2.3.2, except the lowest mass point, which is obtained from a fit by Lovisari et al. (2015). We show the individual clusters as black dots. Note that the X-ray data are plotted without any correction for the hydrostatic mass bias. For this work we use the Driver et al. (2022) data for the SMF, and the X-ray and Akino et al. (2022) data for the gas fractions.

Table 2.3: Mass ranges in $M_\odot$ used for each observable when fitting the emulator to data. The values are rounded because the exact ranges vary with the values of the observational bias factors.

Observable	SMF M_* lower limit	SMF M_* upper limit	$f_{\mathrm{gas,500c}}\ M_{500c}$ lower limit	$f_{\mathrm{gas,500c}}\ M_{500c}$ upper limit
High-res [m8]	$10^{8.67}$	$10^{11.50}$	$10^{13.50}$	$10^{13.73}$
Intermediate-res [m9]	$10^{9.92}$	$10^{11.50}$	$10^{13.50}$	$10^{14.36}$
Low-res [m10]	$10^{11.17}$	$10^{11.50}$	$10^{13.50}$	$10^{14.53}$

2.3 Observational data and biases

Before we can start to calibrate our simulations, we need to have observational data to compare with our simulations. We calibrate to the galaxy stellar mass function (SMF) and the gas fractions in groups and clusters ($f_{\rm gas,500c}(M_{500c})$).

One of the goals of the FLAMINGO simulations is to predict galaxy clustering and cross correlations between galaxies and other tracers of the matter distribution. The SMF allows us to constrain the stellar content of haloes as a function of their mass. This is not only crucial for the prediction of observations using galaxies, the stellar mass also directly affects the distribution of dark matter in haloes, and the orbits of subhaloes. Although matching the SMF does not ensure that each halo contains the correct stellar mass, it suggests the relation is at least statistically plausible provided the model assumes the correct cosmology.

Besides galaxy clustering, we also wish to use FLAMINGO to investigate other cosmological observables tracing the distribution of matter, such as X-ray emission, the Sunyaev-Zeldovich (SZ) effect and lensing maps. From studies by Semboloni et al. (2013), Van Daalen et al. (2020) and Salcido et al. (2023) we know that the gas fractions in clusters have a large impact on the matter power spectra on scales relevant for e.g. cosmic shear. By calibrating to the observed gas fractions, we can also make robust predictions for the distribution of gas expelled from group/cluster cores.

We calibrate to the same observables as were used for the BAHAMAS simulation (McCarthy et al., 2017, 2018). In this section we will discuss the data that we considered and the observational biases that we account for.

2.3.1 The galaxy stellar mass function

Constraining the SMF has been the goal of a large number of studies, many of which are based on the SDSS (Li & White, 2009; D'Souza et al., 2015; Bernardi et al., 2013, 2017) or the more recent GAMA survey (Baldry et al., 2012; Wright et al., 2017; Driver et al., 2022). A compilation of these data sets is shown in the left panel of Fig. 2.1. It is clear that there are substantial systematic differences between some of the different groups that have tried to measure the SMF, particularly at the low- and high-mass ends. However, some of the most significant outliers are older results. While there are still discrepancies at the high-mass end, the results from the three most recent studies, D'Souza et al. (2015); Bernardi et al. (2017); Driver et al. (2022), are in reasonable agreement over a large part of the mass range. Instead of trying to combine different data sets, we limit the fitted mass range to $M_* < 10^{11.5}\,{\rm M}_\odot$ and we choose to use the most recent GAMA result from Driver et al. (2022) at $z = 0$. Not only is this the most recent study, it also provides a useful prior for possible biasing due to cosmic variance. The upper mass limit also decreases the possible bias we get due to our choice of sim-

ulation aperture (see §2.4.2 and Appendix 2.A for more details). We always set a simulation-resolution dependent lower mass limit on the mass range we use for fitting. The mass ranges we use can be found in Table 2.3.

Fitting the SMFs from simulations to observations requires special care. There are some important differences/sources of uncertainty that need to be taken into account:

1. Observations suffer from random errors in measuring the mass. while simulations have no mass measurement errors (at least for a fixed definition of a galaxy, i.e. for a given subhalo finder). Simulations do suffer from randomness errors (see Borrow et al., 2022a), as discussed by these authors, this issue is negligible for our analysis because we consider large ensembles of galaxies..

2. Observations possibly suffer from systematic errors, which may originate from spectral energy distribution fitting, corrections for dust extinction, surface brightness profile fitting, and/or selection effects.

3. Observations may suffer from cosmic variance.

Before discussing how we take each of these effects into account, we note that the uncertainty in the stellar IMF is not directly relevant because the observational analysis and the simulations use the same IMF. The observed SMF also depends on the assumed cosmology, but this is close enough to the one used in the simulations to have a negligible effect on the comparison.

Random errors on the observed stellar mass

Symmetric observational scatter in the measured stellar mass will cause a systematic shift in the inferred SMF. Because there are more galaxies in lower mass bins, it is more likely for galaxies to scatter to a higher mass bin than to a lower mass bin. This is especially important at the high-mass end, where the SMF is steep. This effect is known as Eddington (1913) bias. We account for it by adding scatter to the simulation masses. We adopt the lognormal scatter from Behroozi et al. (2019), which has a redshift-dependent standard deviation of

$$\sigma(\log_{10} M_*) = \min(0.070 + 0.071z, 0.3)\ \text{dex}, \tag{2.11}$$

where we sample the lognormal distribution for each galaxy. This then adds an Eddington-like bias to the simulation results, consistent with observations.

Systematic errors in the observed stellar mass

There are systematic discrepancies between the different observations. The reason for this is mostly found in the stellar population synthesis and dust correction

models used, as the observed luminosity functions agree better between different studies than the mass functions. However, at the FLAMINGO resolution, the stellar masses can be predicted much more accurately than the star formation histories, current-day star formation rates and dust extinction rates. Therefore, calibration to the SMF is preferable over a direct comparison with the luminosity function.

To account for potential systematic shifts in the observed stellar masses, we include a stellar mass bias parameter

$$\log_{10}(M_{*,\mathrm{obs}}) \rightarrow \log_{10}(M_{*,\mathrm{obs}}) + \log_{10} b_*, \tag{2.12}$$

where the bias b_* is assumed to be independent of mass. Note that the sign is defined such that a positive stellar mass bias implies the observations underestimate the true stellar mass. We use a lognormal prior to constrain the bias parameter. The prior is taken from Behroozi et al. (2019) (their eq. 25) and is based on the existing tensions between observed time-integrated star formation rates and observed SMFs,

$$\log_{10} b_* = \mathcal{N}(0, 0.14), \tag{2.13}$$

where $\mathcal{N}(\mu, \sigma)$ is a normal distribution with mean μ and standard deviation σ.

We adopt a mass-independent bias. While a mass-dependent bias might have improved the agreement between the data and the simulations, the mass dependence is unknown and therefore there is no obvious parametrization of the mass dependence. This implies the new free parameters would have no clear priors. Additionally, we note that our decision not to fit above a stellar mass of $10^{11.5}$ $\mathrm{M}_\odot$ has a similar effect as switching to a much higher stellar mass bias above this mass.

Cosmic variance

Driver & Robotham (2010) showed that the error on the SMF due to cosmic variance can be 5–10 per cent for surveys like GAMA and the SDSS, depending on the volume considered. Cosmic variance can bias the number density measurements, because the survey may consist of slightly over- or under-dense regions. For our mass range we assume that this effect is independent of mass (S. P. Driver, private communication). To account for cosmic variance, we allow the observed number densities to shift up and down slightly,

$$f_{\mathrm{obs}} \rightarrow f_{\mathrm{obs}} + \log_{10}(b_{\mathrm{cv}}). \tag{2.14}$$

Note that the sign is defined such that a positive cosmic variance bias implies the observations underestimate the number density of galaxies. We constrain this bias parameter with a Gaussian prior taken from Driver et al. (2022). They

estimate the error due to cosmic variance to be about 6 per cent, so our prior is given by

$$b_{\mathrm{cv}} = \mathcal{N}(1, 0.06). \tag{2.15}$$

Table 2.4: Overview of the cluster gas mass fraction data used for this work. The first column lists the reference from which the data were obtained, the second column lists the number of objects, where 'fit' indicates that the main result is a fitted relation between M_{500c} and $f_{\mathrm{gas},500c}$, the third column shows how the total mass was measured (HSE: X-ray data assuming hydrostatic equilibrium; WL: weak gravitational lensing), and the final column contains comments on the selection method.

Reference	N	Type	Selection
Vikhlinin et al. (2006)	10	HSE	Nearby, relaxed, ambiguous X-ray limit
Maughan et al. (2008)	114	HSE	NED Cross-match, $z > 0.1$
Rasmussen & Ponman (2009)	15	HSE	Bright groups
Sun et al. (2009)	23	HSE	$0.015 < z < 0.13$, resolved temperature profiles
Pratt et al. (2010)	31	HSE	X-ray flux limited, $z < 0.2$
Lin et al. (2012)	94	HSE	Infrared magnitude limited
Laganá et al. (2013)	126	HSE	Crossmatch between Maughan et al. (2008) and SDSS; X-ray flux limit
Sanderson et al. (2013)	5	HSE	Optical magnitude limit, $\sigma \leq 500c$ km s^{-1}
Gonzalez et al. (2013)	15	HSE	Optical magnitude limit, $0.03 < z < 0.13$
Lovisari et al. (2015)	20	HSE	X-ray flux limited
Hoekstra et al. (2015)	50	WL	X-ray flux limited
Pearson et al. (2017)	8	HSE	GAMA r-band selection, $N > 12$, $z < 0.12$
Mulroy et al. (2019)	fit	WL	X-ray luminosity limit
Lovisari et al. (2020)	120	HSE	tSZ-selected from Planck data.
Akino et al. (2022)	fit	WL	C1 - X-ray selected, C2 no clear selection.

2.3.2 The cluster gas mass fractions

Data for the cluster gas mass fractions, $f_{\rm gas,500c}$, come in two varieties. They are either obtained purely from X-ray observations, or from a combination of X-ray and weak gravitational lensing observations where the latter are used to measure the total cluster mass. For the X-ray only data, the density and temperature profiles fitted to the observations are used to measure the total mass assuming the gas is in hydrostatic equilibrium (HSE). In both cases the gas mass is obtained by integrating the density profile measured from X-ray observations out to the measured value of R_{500c}. Table 2.4 summarises all the different sets of data that we use.

As was the case for the SMF, there are biases that we need to account for when we compare observations with simulations. There are four distinct issues that we take into account:

1. At the low-mass end selection effects become important, because at fixed halo mass objects with a higher gas content will tend to emit more X-ray radiation. Any X-ray selected sample may therefore have gas fractions that are biased high, particularly at low masses.

2. The measurement of total mass from X-ray data under the assumption of HSE is well documented to be biased low (e.g. Hoekstra et al., 2015; Eckert et al., 2016; Smith et al., 2016).

3. For the weak lensing data, we make use of the fits of the relation between gas fraction and mass provided by the authors. The fits are preferred to individual measurements as the fits account for the selection function of the sample. However, for our purposes the fits need to be sampled at particular masses. This needs to be done in a way that limits the covariance between the samples and that is representative of the data used (i.e. no extrapolation).

4. As clusters are rare objects they are usually observed over a large redshift range. Furthermore, because weak lensing is most efficient when the lens is halfway between the observer and the background galaxies, weak lensing observations tend to probe higher redshifts than X-ray data. Clusters evolve over time, so we need to make sure the simulation samples are representative for the observational samples we compare them with.

For the cluster gas fractions the largest mass we can fit for is limited by the box size of each simulation. The upper mass limit used for fitting therefore changes with resolution (as we use a different box size for each resolution). The upper limits can be found in Table 2.3.

Table 2.5: Compilation of cluster X-ray gas fraction data used for calibration. These values are for the DESYR3 cosmology ($h = 0.681$, $\Omega_{\rm m} = 0.298$). The values are obtained by taking the median of the X-ray data described in Table 2.4 in eight logarithmically spaced bins between $10^{13.8}$ and $10^{15.0}$ $\rm M_\odot$. The errors are the absolute difference between the 16th or 84th percentile and the median (whichever is largest), obtained by bootstrap resampling the median.

M_{500c} ($\log_{10} \rm M_\odot$)	$f_{\rm gas,500c}$
13.89	0.083 ± 0.002
14.06	0.094 ± 0.003
14.23	0.105 ± 0.005
14.40	0.115 ± 0.008
14.57	0.130 ± 0.002
14.74	0.130 ± 0.002
14.91	0.139 ± 0.003

X-ray data

The first set of gas fraction data we describe is the X-ray (or HSE) data. For each data set we store M_{500c} and $f_{\rm gas,500c}$, with asymmetric errors where available, and correct the data to the FLAMINGO cosmology ($M_{500c} \propto h^{-1}$, $f_{\rm gas,500c} \propto h^{-1.5}$). The combined data set has 581 objects but contains duplicates. For each object that appears more than once we calculate a new data point by taking an unweighted mean of the different measurements. The mean is taken in both M_{500c} and $f_{\rm gas,500c}$. Because the duplicates are often based on (in part) the same data, the errors will not be independent and we combine them via

$$\sigma^2 = \frac{1}{N} \sum_i^N \sigma_i^2, \tag{2.16}$$

where N is the number of times a single object appears in the set. This leaves us with 533 objects. Note that we do not use the errors for the re-binning, as we make use of bootstrap re-sampling to compute the errors.

We need to consider redshift evolution. The emulators will be trained on simulation snapshots corresponding to a single redshift. Imposing a redshift cut of $z < 0.25$ causes the median redshift of the X-ray sample to become 0.1, thus allowing us to compare with simulation snapshots at $z = 0.1$. The redshift cut reduces the sample to 310 objects. The individual masses and gas fractions are shown as black dots in Fig. 2.1.

We combine the X-ray measurements by computing the median gas fraction in eight logarithmically spaced hydrostatic mass bins between $10^{13.8}$ and $10^{15.0}$ $\rm M_\odot$. For each bin, the error on the median is obtained by taking the difference between

the median and the 16th–84th percentiles obtained from bootstrap resampling the objects. This gives us asymmetric errors around the median. As our likelihood uses symmetric errors, we use only the greater of the positive and negative errors. The tabulated data points can be found in Table 2.5.

2

Furthermore, selection effects are expected to be most prevalent at lower halo masses. The median observed gas fraction as a function of mass shows a clear trend-break at $M_{500c,\mathrm{HSE}} \approx 10^{13.8}\,\mathrm{M}_\odot$. Below this mass the gas fractions no longer decrease, but instead plateau, a behaviour that deviates from what is expected for an unbiased sample (e.g. McCarthy et al., 2017). To deal with this we impose a mass cut at a hydrostatic mass of $M_{500c,\mathrm{HSE}} > 10^{13.8}\,\mathrm{M}_\odot$, but add the fit from Lovisari et al. (2015) at their median mass ($4\times10^{13}\,\mathrm{M}_\odot$) as a separate data point.

We account for hydrostatic mass bias by adding a constant bias term to the HSE masses,

$$\log_{10} M_{500c} = \log_{10} M_{500c,\mathrm{HSE}} - \log_{10}(b_{\mathrm{HSE}}). \tag{2.17}$$

Note that values $b_{\mathrm{HSE}} < 1$ imply that the hydrostatic mass estimate underestimates the true mass. We neglect the effect of hydrostatic bias on the gas fraction because it is comparatively small (McCarthy et al., 2017). This is because both the total and gas mass increase with increasing R_{500c}. The measured gas fraction will differ only at the level of the change in cumulative gas fraction between the true and biased R_{500c}. This is expected to cause only mild changes in the gas fraction (see e.g. fig. 6 of Velliscig et al., 2014). Before calculating the median that we compare with the simulation we thus adjust all the observed HSE masses. By combining both X-ray and weak lensing observations, we can constrain the hydrostatic bias. However, we found that our compilation of data on its own is not constraining enough without the use of a prior. To define our prior, we take the values 0.72 ± 0.08 from Eckert et al. (2016) and 0.76 ± 0.06 from Hoekstra et al. (2015) and combine the two to obtain the Gaussian prior

$$b_{\mathrm{HSE}} = \mathcal{N}(0.74, 0.10). \tag{2.18}$$

Eckert et al. (2016) and Hoekstra et al. (2015) estimate the hydrostatic mass bias by directly comparing the masses they obtain from weak lensing and from X-rays.

Weak lensing data

We complement the X-ray data with the latest HSC-XXL weak gravitational lensing data from Akino et al. (2022). Higher-mass data from Mulroy et al. (2019) and Hoekstra et al. (2015) are available and plotted in Fig. 2.1, but the box size used for our calibration runs is too small to make use of them. To compare with the weak lensing data, we make use of the power-law fits to the relation between the gas fraction and mass given by the authors. These fits take selection effects into account. Because the power-law fits have two free parameters, sampling them at more than two masses would result in strong covariance between the sampled

points. We therefore use the fit to create two data points that are spaced equally far from the pivot used by the authors. This gives us $f_{\mathrm{gas},500c}(M_{500c} = 10^{13.5}\,\mathrm{M}_\odot) = 0.054 \pm 0.010$ and $f_{\mathrm{gas},500c}(M_{500c} = 10^{14.5}\,\mathrm{M}_\odot) = 0.106 \pm 0.023$. Due to the limited box size, we use only the lower, $M_{500c} = 10^{13.5}\,\mathrm{M}_\odot$, point for fitting high- and intermediate-resolution simulations. For low resolution we are able to include the second $M_{500c} = 10^{14.5}\,\mathrm{M}_\odot$ point.

The median redshift of the HSC-XXL sample is $z = 0.3$. We therefore construct a separate emulator for $f_{\mathrm{gas},500c}$ at $z = 0.3$, which we use to fit the weak lensing data. The fits make use of self-similar scaling to move the different clusters to the same redshift, so we could have corrected them to the redshift $z = 0.1$ used for the X-ray data. However, we prefer to use a redshift close to that of the actual sample, to minimize the size of the correction. Akino et al. (2022) give both the weak lensing inferred and the true M_{500c}, as they correct for the expected bias on the weak lensing inferred M_{500c}. We make use of their calibrated true M_{500c} masses.

2.4 Emulator construction

Cosmological hydrodynamical simulations are too expensive to be run for each step in an MCMC chain used to evaluate likelihoods. In order to use simulation outputs in MCMC methods, we therefore make use of emulators trained on a set of simulations. Emulators are used to interpolate results in the parameter space between training simulations. They are able to predict the output of the simulations as a continuous function of the input parameters, in a fraction of the original computation time. This method has previously been applied to the matter power spectrum (e.g. Heitmann et al., 2009, 2016b; Euclid Collaboration et al., 2019; Angulo et al., 2021) and to baryonic observables (e.g. Oh et al., 2022; Jo et al., 2023). By using emulators, we can interpolate between the results of a set of training simulations and obtain a fully continuous prediction of how the simulation responds to changes in subgrid parameters.

2.4.1 Training sets

The first step in setting up the emulator is to create a training set. In our training set we want to vary those subgrid parameters that we know are important for the calibration. As discussed in Section 2.2, for the intermediate- and high-resolution simulations we vary the following four parameters: the stellar feedback efficiency, f_{SN}, the target kick velocity for stellar feedback, Δv_{SN}, the power-law slope of the density dependence of the black hole accretion boost factor, β_{BH}, and the AGN heating temperature, ΔT_{AGN} (v_{jet}, the target kick velocity for AGN feedback in the jet model). For the low-resolution simulations we do not require stellar feedback and therefore vary only the last two parameters. The ranges over

which the parameters are varied are motivated in Section 2.2 and listed in Table 2.2 (Table 2.9 for the jet model).

To optimise the parameter space, we make use of a Latin hypercube, first proposed by McKay et al. (1979). To set up a Latin hypercube with $N_{\rm sims}$ nodes, we start with an ordered list of $N_{\rm sims}$ independent samples along every dimension of the hypercube, where the number of dimensions equals the number of subgrid parameters that are varied. These samples are then combined and shuffled to create a set of $N_{\rm sims}$ points $\boldsymbol{\theta}$ that are distributed uniformly within the hypercube, where in our case $\boldsymbol{\theta} = (f_{\rm SN}, \log_{10}\Delta v_{\rm SN}, \beta_{\rm BH}, \log_{10}\Delta T_{\rm AGN})$ for intermediate and high resolution, and $\boldsymbol{\theta} = (\beta_{\rm BH}, \log_{10}\Delta T_{\rm AGN})$ for low resolution. Our criterion for optimising the sampling is the 'maximin' approach, which maximises the minimum distance that sampled points are away from each other. An in depth explanation of how the method works is provided by Heitmann et al. (2009). We apply to each sample a random shift of at most half the average spacing between samples. We then run the $N_{\rm sims}$ simulations corresponding to the nodes of the Latin hypercube.

We use the public package SWIFTEMULATOR[6] (Kugel & Borrow, 2022), built on the package GEORGE (Ambikasaran et al., 2015a), to set up the Latin hypercube as well as to train and test the emulators. SWIFTEMULATOR streamlines the emulation process for results obtained from SWIFT runs. Within SWIFTEMULATOR we use the Latin hypercube generator from PYDOE (Baudin et al., 2012).

We use $N_{\rm sims} = 32$. The sampling of parameter space provided by the Latin hypercube used for intermediate resolution is shown in Fig. 2.2. The box sizes used for the training are $(100\ {\rm Mpc})^3$, $(200\ {\rm Mpc})^3$ and $(400\ {\rm Mpc})^3$ for high, intermediate, and low resolution, respectively. The volume is a compromise between computational cost and the maximum mass for which we train the emulator. Each run cost ~ 800, ~ 1300 and ~ 1600 cpu hours for low, intermediate and high resolution respectively. Using single simulations with an eight times larger volume at each resolution and with the results of Schaye et al. (2023), we have verified that these box sizes are sufficiently large for box size effects to be negligible with respect to the production runs.

2.4.2 Obtaining the required simulation output

From our simulation we take three snapshots at $z = 0$, 0.1 and 0.3. For each snapshot we find haloes and subhaloes using VELOCIRAPTOR (Elahi et al., 2019b; Cañas et al., 2019). After an initial friends of friends group search it uses the full 6-D phase space information to disentangle the central and satellite subhaloes.

One of the difficulties of comparing with data, is that we have to choose how to define the edge of simulated galaxies. Observed cluster gas mass fractions are measured within R_{500c}. For the stellar masses needed to compute the

[6] https://swiftemulator.readthedocs.io/en/latest/

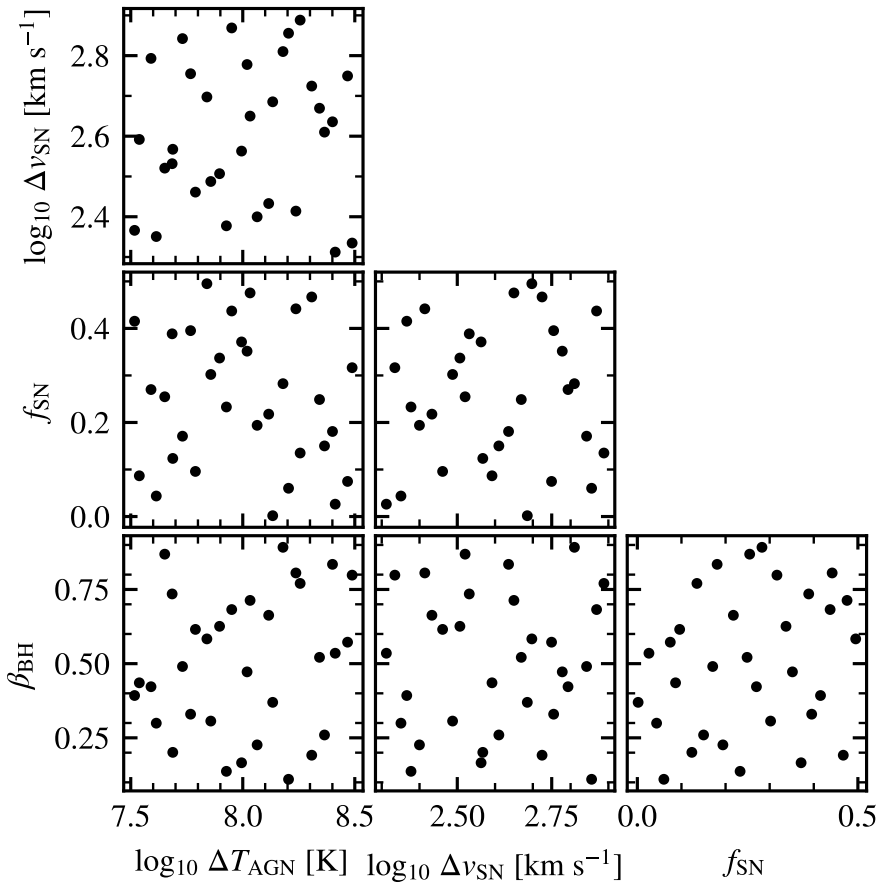


Figure 2.2: The sampling of parameters in the 32-node Latin hypercube used to train the emulator for the intermediate-resolution simulations.

SMF, the situation is less clear. Ideally, we would create mock observations, fit them with Sérsic profiles and integrate these to obtain stellar masses, which is the procedure adopted by observational studies. This was recently done for the EAGLE simulation by De Graaff et al. (2022). However, the resolution of the FLAMINGO simulations is too limited to mimic the observational strategy. As shown by Schaye et al. (2023), FLAMINGO significantly overestimates the sizes of low-intermediate mass galaxies, which means we cannot create realistic virtual galaxy observations. Based on the findings of De Graaff et al. (2022), we choose to calibrate the SMF using a 3D aperture with a radius of 50 kpc for the simulations. A comparison between different choices of aperture can be found in Appendix 2.A, where we show that the aperture becomes only important above a stellar mass of $\approx 10^{11}\,\mathrm{M}_\odot$.

Before computing the galaxy SMF, we first add random errors to the simulation stellar masses as described in §2.3.1. The SMF is then sampled in 25 logarithmically spaced mass bins between $10^9\,\mathrm{M}_\odot$ and $2\times10^{12}\,\mathrm{M}_\odot$ for intermediate- and low-resolution simulations, and 40 bins between $10^8\,\mathrm{M}_\odot$ and $2\times10^{13}\,\mathrm{M}_\odot$ for high-resolution simulations. We choose to use a finer binning than is available for the observational data to allow the emulator to capture the finer features of the predicted SMF. Tests with different binning strategies show this had no effect on the results. We have enough galaxies across the fitted mass range for the Poisson errors to still be very small even with finer binning. The uncertainty we provide to the emulator is the Poisson error for each bin.

For the gas fraction we instead opt for an adaptive binning strategy. While the simulation volumes used for the calibration are large enough to constrain the SMF over the adopted mass range, at the high cluster mass end, we always run out of clusters before we run out of data to compare with. For all resolutions we use 20 bins between M_{500c} of 10^{13} and $10^{15}\,\mathrm{M}_\odot$ although we never manage to make use of this entire range. As the higher mass bins start to run out of objects, we allow the highest mass bin to stretch to include a sufficient number of objects. We require each bin to contain at least ten objects. We also limit the stretching of the bin to half the original bin width. The uncertainties we provide to the emulator are based on the 16th–84th percentiles. As the emulator only takes symmetrical errors, we take mean of the absolute difference between the median and 16th percentile and the difference between the median and 84th percentile. For both the SMF and the cluster gas fraction we discard any empty bins.

2.4.3 Training using Gaussian processes

After measuring the SMF and cluster gas fraction for each node of the hypercube, we can train an emulator for each observable. Because each individual node of the Latin hypercube requires a cosmological hydro simulation, we are operating in a regime where we have a limited number of samples. We also know a priori that

the observables we want to emulate (i.e., the galaxy number density and group and cluster gas fractions) vary smoothly with mass and with the values of the subgrid parameters. Both these properties are in the regime in which Gaussian processes give excellent predictive power with respect to the input data (see e.g. Rasmussen et al., 2004; Rasmussen & Williams, 2006).

We set up a different Gaussian process for each relation we emulate. We combine the mass (either stellar or M_{500c}) and subgrid parameters into a single input data vector $\boldsymbol{x} = (\log_{10} M, \boldsymbol{\theta})$, from which the emulator then predicts the dependent quantity, which is either the number density of galaxies, $f(M_*)$, or the gas fraction, $f_{\mathrm{gas,500c}}$. Each emulator thus has $N+1$ parameters, where N is the number of subgrid parameters that are varied. In order to limit the dynamic range, we transformed many of the inputs to log-space. This includes the masses (aperture stellar mass or M_{500c}), the values of the SMF and the two subgrid parameters that are sampled in log-space (Δv_{SN} and ΔT_{AGN}). This is an important step as it greatly increases the smoothness of the emulated relations, making it much easier for the emulator to give accurate predictions. As the input relations are smooth over the range we are interested in, we do not require any other transformations of the input. We feed the data directly into the Gaussian process. We use a squared exponential kernel

$$k(\boldsymbol{x}, \boldsymbol{x}') = \exp\left(-\frac{(\boldsymbol{x}-\boldsymbol{x}')^T \boldsymbol{\Theta}^{-1} (\boldsymbol{x}-\boldsymbol{x}')}{2}\right), \tag{2.19}$$

where $\boldsymbol{\Theta}$ represents a diagonal matrix containing the hyperparameters that set the scale for each input parameter, and $\boldsymbol{x}$ and $\boldsymbol{x}'$ are two positions in parameter space. The hyperparameters are optimised based on maximising the marginal likelihood (see Rasmussen & Williams, 2006). As we train a separate Gaussian process for each relation, we also have a separate set of hyperparameters for each relation. We have verified the posteriors of the hyperparameters to ensure that the values we use are well converged.

2.4.4 Error estimation

It is important to verify that the emulator is able to give accurate results before we use it to find best-fitting subgrid and bias parameters. Moreover, we need to quantify the accuracy of the emulator because we will account for emulation errors when fitting to data. The best way to measure the uncertainty in the emulator predictions is to perform test simulations that span the emulated parameter space. However, this implies that we would need to run many additional simulations. To save time, we choose instead to measure the uncertainty by making use of k-fold cross-validation, which we will refer to as cross-checks.

We create N_{sims} new data sets, where N_{sims} is the number of nodes in our Latin hypercube (32 in our case). For each of these data sets we take out one simulation

and retrain the emulator on the reduced set of $N_{\text{sims}} - 1$ samples. We then test how accurately the emulator is able to predict the simulation that was left out. We do this by taking the ratio between the result from the run that was left out, and the prediction of the emulator for the parameter values of the left-out run. This gives us a value for each mass bin in the training data. We combine the ratios for all mass bins and N_{sims} emulators into a single list and compute the standard deviation, $\sigma_{\text{crosscheck}}$. The error on the emulator prediction, σ_{emu}, is then given by

$$\sigma_{\text{emu}} = |\sigma_{\text{crosscheck}} f(M, \grave{})|, \tag{2.20}$$

where $f(M, \grave{})$ is the value predicted by the emulator for mass M and at parameter values $\grave{}$. The result of the cross checks for the Latin hypercube of intermediate-resolution simulations can be seen in Fig. 2.3. It is important to note that cross checks are a conservative method to estimate the uncertainty. The input for cross-checks is uniformly sampled, implying that a significant fraction of the test points is located near the boundaries of the parameter space, where a Gaussian process is naturally less accurate.

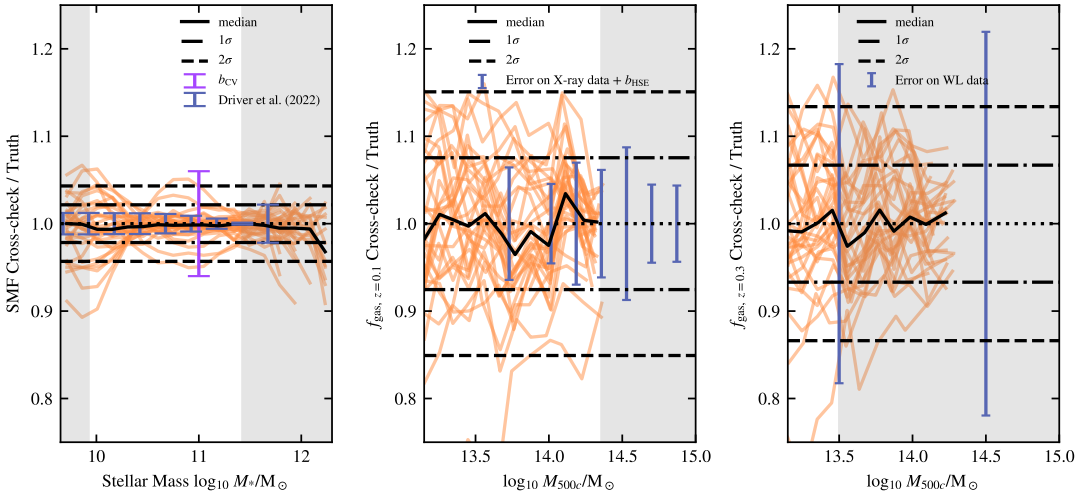


Figure 2.3: Performance of the emulator on cross checks (see §2.4.4) for the redshift $z = 0$ SMF (left panel), the $z = 0.1$ X-ray cluster gas fractions (middle panel), and the $z = 0.3$ weak lensing cluster gas fractions (right panel) at intermediate [m9] resolution. Each of the 32 red lines corresponds to the case where a single simulation from the 32-node Latin Hypercube has been omitted from the training set. The curves show the ratio of the emulator prediction for the parameter values of the omitted simulation to the actual simulation values. The solid black line shows the median as a function of mass. The horizontal dash-dotted and dashed lines indicate, respectively, the 1 and 2 σ mean errors on the emulator. The horizontal dotted lines indicate the one-to-one lines, i.e. zero errors. The grey bands indicate the regions that are not used for fitting in Section 2.5. In each panel we also indicate the observational errors. For the SMF we show the error due to cosmic variance and the errors on the data by Driver et al. (2022), for the $z = 0.1$ gas fractions we combine the error from the X-ray data with the error due to hydrostatic bias and for the z=0.3 gas fraction we show the error on the weak lensing data by Akino et al. (2022). The emulator predictions are accurate enough to predict to simulation output within the observed constraints

Table 2.6: Accuracy of the emulators, $\sigma_{\text{crosscheck}}$, for the three different simulation resolutions and the jet model AGN variation, in percentages. The values are obtained by taking the standard deviation of the ratio between the result from the simulation omitted from the Latin hypercube and the prediction from the emulator trained on all but that simulation.

Calibration target	High	Intermediate	Low	Jet
$\log_{10}$ SMF	2.7	2.2	1.5	1.9
$f_{\text{gas},z=0.1}$	8.9	7.5	4.8	7.1
$f_{\text{gas},z=0.3}$	7.9	6.7	4.2	6.1

From Fig. 2.3 it is clear that our emulators do not suffer from significant systematic errors for our three calibration targets, the $z = 0$ SMF, $z = 0.1$ X-ray cluster gas fractions, and $z = 0.3$ weak lensing cluster gas fractions. There are no significant trends with mass, and the medians ratio is centered close to one, which corresponds to an error of zero.

It is clear that the emulator for the SMF is more accurate than the emulators for the gas fractions. This is a reflection of the way we constrain the input simulations. In the case of the SMF, the errors on the input are Poisson errors, which are quite small for our simulation volumes in the mass range we are interested in. The f_{gas} errors are based on the 16th–84th percentiles of the simulated gas fractions in each mass bin, which can be larger than the 5 per cent accuracy that the emulator attains.

The emulator accuracy for all resolutions can be found in Table 2.6. The emulators become more accurate going to lower resolution. There are several possible reasons for this trend. First, we used larger box sizes for the lower-resolution simulations, so the uncertainty intrinsic to the simulation is smaller at fixed mass. Second, we used a slightly larger parameter range for high resolution than for intermediate resolution, while for low resolution we only used two parameters, greatly reducing the sampled space.

The obtained accuracy is sufficient, as it is higher than the observational scatter/uncertainty. Any deviations between the model and the data at the level of the emulator error would still be consistent with the observational constraints, especially as we allow for observational biases in our analysis.

2.5 Using the emulator for parameter estimation

To use the emulator as the model that we compare with observational data, we need a way to optimise the subgrid parameters $\boldsymbol{\theta}$ (see Section 2.2) and, optionally, the observational bias factors $\log_{10} b_*$, b_{CV}, and b_{HSE} (see Section 2.3).

For parameter optimisation we use the Markov chain Monte Carlo (MCMC)

package EMCEE (Foreman-Mackey et al., 2013). We use the ensemble sampler, which we give our posterior likelihood. For every fit we have done using MCMC, we have varied the number of walkers and steps to ensure the resulting values are converged. We discard the first 500 steps of each chain to avoid systematic errors due to the burn-in phase.

To evaluate the goodness of fit of an emulator prediction to the observations, we first define the log likelihood for a single observed mass bin. For the SMF this is given by

$$\ln \mathcal{P}_{\rm SMF}(M_{*,\rm obs}, b_{\rm cv}, b_*, \boldsymbol{\theta}) \equiv - \frac{[f_{\rm obs}(M_{*,\rm obs}) + \log_{10} b_{\rm CV} - f_{\rm emu}(b_* M_{*,\rm obs}, \boldsymbol{\theta})]^2}{\sigma^2_{\rm obs}(M_{*,\rm obs}) + \sigma^2_{\rm emu}(b_* M_{*,\rm obs}, \boldsymbol{\theta})}, \quad (2.21)$$

Here $f(M_*)$ is the SMF,

$$f(M_*) \equiv \log_{10}\left(\frac{{\rm d}n}{{\rm d}\log_{10}(M_*)}\right), \quad (2.22)$$

the subscripts indicate whether the quantity is observed ('obs') or emulated ('emu'), $\boldsymbol{\theta}$ is a vector containing the values of the varied subgrid parameters, and σ is the error on f. For $\sigma_{\rm emu}$ this refers to the error on the emulator from cross-checks, equation 2.20. The expression also accounts for observational bias factors due to cosmic variance, $b_{\rm CV}$, and the conversion of direct observables into stellar mass, b_*, that were discussed in §2.3.1. For cluster gas fractions measured from X-ray observations the log likelihood is defined as

$$\ln \mathcal{P}_{\rm gas}(M_{500c,\rm obs}, b_{\rm HSE}, \boldsymbol{\theta}) \equiv - \frac{\left[f_{\rm gas,500c,obs}(M_{500c,\rm obs}) - f_{\rm gas,500c,emu}(b^{-1}_{\rm HSE} M_{500c,\rm obs}, \boldsymbol{\theta})\right]^2}{\sigma^2_{\rm obs}(M_{500c,\rm obs}) + \sigma^2_{\rm emu}(b^{-1}_{\rm HSE} M_{500c,\rm obs}, \boldsymbol{\theta})}, \quad (2.23)$$

where $b_{\rm HSE}$ is an observational bias factor due to the assumption of hydrostatic equilibrium that was discussed in §2.3.2. For gas fractions measured from weak lensing plus X-ray observations the log likelihood definition is identical except that we assume the masses are unbiased, implying $b_{\rm HSE} = 1$ (see e.g. Becker & Kravtsov, 2011; Bahé et al., 2012). Note that for the likelihood of both the SMF and the cluster gas fraction we include a variance term to account for the error on the emulator prediction. This is added to avoid situations where we over-fit with respect to the uncertainty from the emulator alone.

The likelihood for the observational data is a combination of the likelihoods

of the individual mass bins of the three data sets

$$\ln \mathcal{P}_{\rm likelihood}(b_{\rm cv}, b_*, b_{\rm HSE}, \boldsymbol{\theta}) = \frac{1}{N_{\rm SMF}} \sum_i^{N_{\rm SMF}} \ln \mathcal{P}_{\rm SMF}(M_{*,{\rm obs},i}, b_{\rm cv}, b_*, \boldsymbol{\theta}) + \frac{1}{2}\left[\frac{1}{N_{\rm HSE}} \sum_j^{N_{\rm HSE}} \ln \mathcal{P}_{\rm gas,X\text{-}ray}(M_{500c,{\rm obs},j}, b_{\rm HSE}, \boldsymbol{\theta}) + \frac{1}{N_{\rm WL}} \sum_k^{N_{\rm WL}} \ln \mathcal{P}_{\rm gas,WL}(M_{500c,{\rm obs},k}, \boldsymbol{\theta})\right], \quad (2.24)$$

where $N_{\rm SMF}$, $N_{\rm HSE}$ and $N_{\rm WL}$ are the number of (re-binned) observational data points (i.e. mass bins) for the SMF, the X-ray cluster gas fraction and the weak lensing cluster gas fraction, respectively. The values of N depend on the fitted mass ranges (Table 2.3) and vary with resolution. We normalise each likelihood by the number of data points to ensure each separate likelihood is not directly dependent on the number of bins used. Furthermore, we average the likelihoods from the two types of cluster gas fraction data to ensure that the cluster gas fraction and SMF data carry equal weight. In an unweighted fit, the SMF would drive the results, because it is much better constrained. As the baryon fractions are the main driver of the baryonic suppression of the matter power spectrum (see e.g. Van Daalen et al., 2011, 2020; Debackere et al., 2020; Schneider et al., 2020; Salcido et al., 2023), we choose to give the gas fractions equal weight in our analysis.

We then combine the different likelihoods into a single posterior,

$$\log \mathcal{P}_{\rm posterior} = \log \mathcal{P}_{\rm likelihood} + \log \mathcal{P}_{\rm prior}, \quad (2.25)$$

where the total prior is

$$\log \mathcal{P}_{\rm prior} = \log \mathcal{P}_{\rm bias}(b_*) + \log \mathcal{P}_{\rm bias}(b_{\rm cv}) + \log \mathcal{P}_{\rm bias}(b_{\rm HSE}) + \log \mathcal{P}_{\rm subgrid}(\boldsymbol{\theta}), \quad (2.26)$$

$\mathcal{P}_{\rm bias}$ are our priors for the observational bias factors, and $\mathcal{P}_{\rm subgrid}$ is our combined prior for the subgrid parameters in $\boldsymbol{\theta}$ that we wish to calibrate. For the subgrid parameters, we use flat priors that do not extend beyond the ranges used for the Latin hypercube (see Table 2.2) in order to avoid extrapolations. The priors on the bias factors were discussed in Section 2.3.

We also calculate the reduced χ^2 for some of our models. We define the re-

duced χ^2 as

$$\chi^2_\nu = \Bigg[\sum_i^{N_{\rm SMF}} \log \mathcal{P}_{\rm SMF}(M_{*,{\rm obs},i}, b_{\rm cv}, b_*, \boldsymbol{\theta}) + \sum_j^{N_{\rm HSE}} \log \mathcal{P}_{\rm gas,X\text{-}ray}(M_{500c,{\rm obs},j}, b_{\rm HSE}, \boldsymbol{\theta}) + \sum_k^{N_{\rm WL}} \log \mathcal{P}_{\rm gas,WL}(M_{500c,{\rm obs},k}, \boldsymbol{\theta}) \Bigg] / (N_{\rm SMF} + N_{\rm HSE} + N_{\rm WL} - N_\theta), \quad (2.27)$$

where N_θ is the number of sub-grid and bias parameters used for the fit.

2.6 Results

In this section we will describe the main results from our calibration approach. We use the emulators to perform parameter sweeps in §2.6.1, then we discuss the fitting results, first at intermediate resolution in §2.6.2 and then at the other resolutions in §2.6.3, and finally we discuss how we use the emulator to set up two AGN feedback variations in §2.6.4.

2.6.1 Parameter sweeps

Emulators can be used to investigate the effect of individual parameters via parameter sweeps, where the emulator predicts the effect of varying a single parameter over the range used for the Latin hypercube, while keeping all other parameters fixed to their best-fitting values. Parameter sweeps can give valuable insight into the importance of particular physical processes and prevent calibration through emulation from becoming a black box. The result of the subgrid parameter sweeps for our intermediate resolution runs are shown in Fig. 2.4. Looking at the response of the calibration targets, it is clear that the different parameters have distinct effects, indicating that the fits will not have any strong degeneracies between the varied subgrid parameters.

Increasing the slope of the black hole accretion rate boost factor suppresses the high-mass end of the SMF, but has almost no effect on the low-mass end and the cluster gas fractions. Increasing the AGN temperature jump leads to a mild reduction of the high-mass SMF, but a strong decrease of the cluster gas fractions. The effects of increasing the stellar feedback energy and kick velocity are more similar. In both cases the stellar masses are decreased, leading to a mass-dependent stretching of the SMF towards lower masses. Depending on the galaxy mass, the SMF can either increase or decrease, though the effect is small for the high-mass end. Cluster gas fractions decrease when either of the stellar

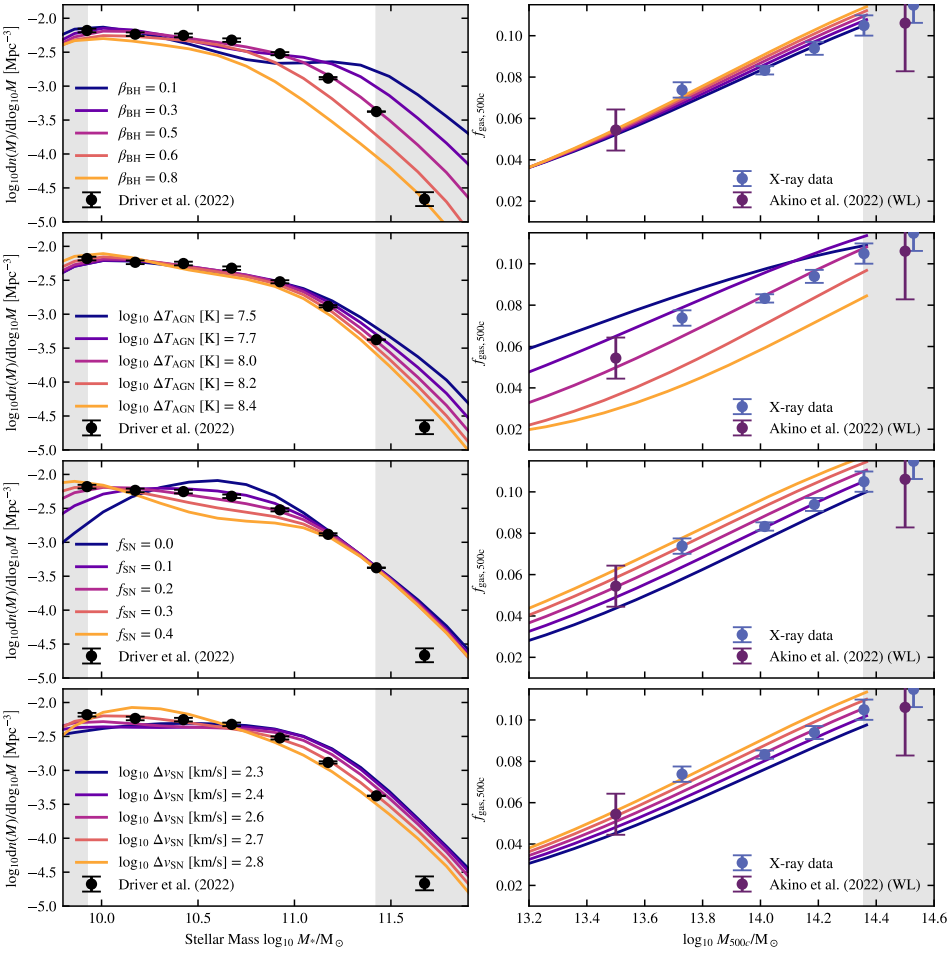


Figure 2.4: Subgrid parameter sweeps using the emulator trained on our 32-node Latin hypercube of $(200\ \mathrm{Mpc})^3$ intermediate-resolution simulations. The parameter sweeps are centred on the best-fitting parameters (see §2.6.2). The left and right columns show the galaxy stellar mass function and cluster gas fractions, respectively. In each row a single subgrid parameter is varied across the allowed range. From top to bottom we vary the slope of the black hole accretion rate boost factor slope, the AGN heating temperature, the stellar feedback energy, and the stellar feedback kick velocity. The grey regions indicate the mass ranges that are excluded for fitting (see also Table 2.3). Parameter sweeps help gain insight into how changes in subgrid model parameters map onto observables.

Table 2.7: Results from the fitting for the observational bias factors. The second column shows the median and 16th and 84th percentiles, the third column lists the maximum likelihood value which we denote as the best-fitting.

Bias	Median+CL	best-fitting
Stellar mass $\log_{10} b_*$	$0.06^{+0.11}_{-0.11}$	0.026
Cosmic variance $b_{\rm CV}$	$0.98^{+0.06}_{-0.06}$	0.995
Hydrostatic equilibrium $b_{\rm HSE}$	$0.74^{+0.09}_{-0.09}$	0.743

feedback parameters increases, presumably because the stronger stellar feedback suppresses black hole growth and hence AGN feedback (Bower et al., 2017).

2.6.2 The best-fitting intermediate-resolution model

The best-fitting (i.e. maximum likelihood) values of the subgrid and observational bias parameters can be found in Tables 2.2 and 2.7, respectively. These tables also list the medians and $16-84$ per cent confidence levels of the posterior distributions.

The posteriors for the subgrid and bias parameters resulting from fitting the emulator predictions for intermediate-resolution simulations to the data are shown in Fig. 2.5. The first thing to note is that the maximum likelihood model (solid, red circle) lies comfortably within the 68 per cent confidence intervals (inner contour) for each parameter and that it does not lie close to an edge of the parameter space. The chosen parameter ranges, i.e. the imposed priors, are thus sufficiently large for the models to bracket the target data and they do not drive the results.

It is also clear that there are no strong degeneracies between any of the subgrid parameters or between any of the bias parameters. The absence of strongly degenerate subgrid parameters is partially by construction, because we chose to fix some of the parameters that would otherwise have caused the results to become degenerate (e.g. $n_{\rm heat}$ and $\Delta T_{\rm AGN}$, see §2.2.3). There is, however, significant degeneracy between the slope of the density dependence of the black hole accretion boost factor ($\beta_{\rm BH}$) and the stellar mass bias (b_*). These two parameters are anti-correlated. Increasing the bias shifts the observed SMF towards higher masses, which means the black hole boost factor needs to decrease to allow more stars to form in high-mass galaxies, whose growth is controlled by AGN feedback.

The best-fitting values for the galaxy mass and cosmic variance biases are $\log_{10} b_* = 0.026$ and $b_{\rm CV} = 0.995$, respectively. The fitted hydrostatic bias, $b_{\rm HSE} = 0.743$, enables the model cluster gas fractions to agree simultaneously with the Akino et al. (2022) weak lensing data and the compilation of X-ray data. For all the bias values we find posteriors that are in agreement with the priors, so we

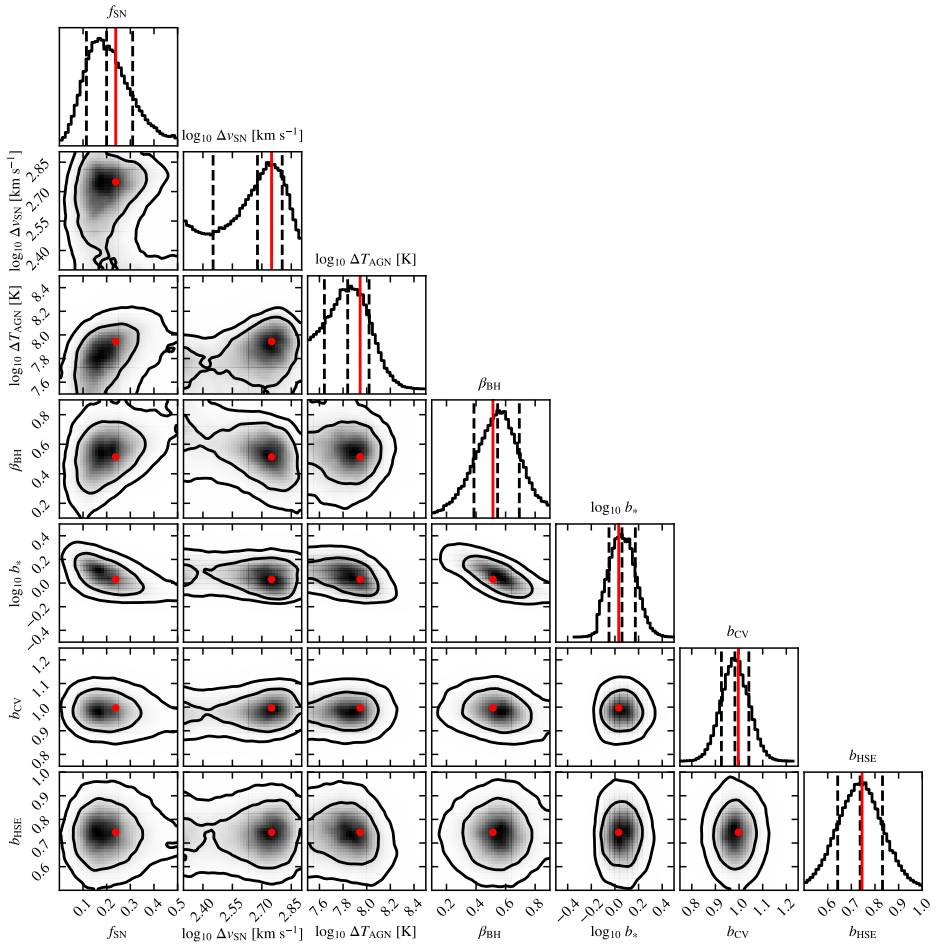


Figure 2.5: The posterior distributions of the model parameters resulting from fitting the emulator to the observed SMF and cluster gas fractions for intermediate-resolution simulations. The parameters shown are the stellar feedback energy, f_{SN}, the stellar feedback kick velocity, Δv_{SN}, the AGN feedback temperature jump, ΔT_{AGN}, the logarithmic slope of the density dependence of the black hole accretion rate boost factor, β_{BH}, the stellar mass bias, b_{M_*}, the hydrostatic mass bias, b_{HSE}, and the cosmic variance bias, b_{CV}. The four subgrid parameters are described in Section 2.2 and the three observational bias factors are discussed in Section 2.3. The black contours show the 68 and 95 per cent confidence levels. The panels along the diagonal show the one dimensional probability density for each parameter. In these plots the three vertical lines indicate the 16th, 50th and 84th percentiles. The solid, red circles indicate the maximum likelihood values, which were used for the fiducial model. Each panel is centered on the centers of the priors given in Table 2.2. The posteriors show that we can find a single solution that fits the simulations to the observational data.

conclude that our fitting does not put any significant additional constraints on the bias parameters.

The best-fitting emulator predictions for intermediate resolution are compared with the data in the middle row of Fig. 2.6, which also shows the result of a $(200\ \mathrm{Mpc})^3$ simulation run with the best-fitting subgrid parameter values (i.e. our fiducial model). The left and right panels show the SMF and cluster gas fractions, respectively. The gas fractions are shown for both the redshift of the X-ray data, $z = 0.1$ (light blue line and dark blue data points), and the redshift of the weak lensing data, $z = 0.3$ (purple line and dark purple data points). Grey regions and dotted line styles indicate mass ranges that were excluded from the fit. The ranges can be found in Table 2.3. Note that the fitted bias factors have been used to shift the data. We obtain good agreement with the fitted observations with a reduced $\chi^2_\nu = 1.23$ for the combined fit to the SMF and the cluster gas fractions. The good agreement between the blue and the red lines demonstrates that the emulator was able to predict accurately what the fiducial simulation would look like in the fitted mass range.

Remarkably, the simulations fit the SMF down to galaxy masses corresponding to slightly fewer than ten stellar particles. Comparing the predicted gas fractions at $z = 0.1$ and 0.3, we see there is very little evolution. The model overshoots the gas fractions for cluster masses between $M_{500c} \approx 10^{13.8}\,\mathrm{M}_\odot$ and $\approx 10^{14.5}\,\mathrm{M}_\odot$, by about 1σ. We emphasize, however, that our observational error bars are about a factor of five smaller than the observed object-to-object scatter. Unfortunately, a box size of $(200\ \mathrm{Mpc})^3$ (or even $(400\ \mathrm{Mpc})^3$) is not large enough to constrain the gas fractions in haloes with $M_{500c} \geq 10^{15}\,\mathrm{M}_\odot$. Performing the same analysis in a larger volume would potentially allow the emulator to train up to the range where the M_{500c}-f_{gas} relation starts to flatten.

2.6.3 The best-fitting subgrid high- and low-resolution models

Although we use the simulation-based emulator to fit for the observational biases, the biases refer to observational effects and should thus be the same for all models. We therefore do not vary them between the different simulation resolutions. We use the intermediate-resolution simulations to fit the biases, because their resolution and box size enable us to fit a substantial mass range for both the SMF and the cluster gas fractions (see Fig. 2.6). For the other resolutions we keep the observational biases fixed to the values listed in Table 2.7. In this way we ensure that a direct comparison can be made between the three different resolutions[7].

[7]The Driver et al. (2022) data points at $M_{\star,\mathrm{obs}} \leq 10^{10}\ \mathrm{M}_\odot$ were updated after we had already finished the $(2.8\ \mathrm{Gpc})^3$ intermediate-resolution FLAMINGO simulation. To be able to use the updated data for the calibration of the high-resolution simulations, which resolve the SMF down to masses for which the data were updated, we re-fit the observational biases at intermediate resolution while keep-

Fixing the observational biases to the values found for intermediate resolution leaves only four parameters to fit for high resolution. For low resolution we only have two parameters to vary because we turn off stellar feedback as these simulations do not resolve the masses below which stellar feedback dominates (see §2.2.1). The best-fitting parameter values for each resolution can be found in Table 2.2. Corner plots of the posterior distributions for the subgrid parameters are shown in Appendix 2.B. A comparison of the best-fitting emulator prediction, the data and runs using the predicted best-fitting subgrid parameter values is shown in the top and bottom rows of Fig. 2.6 for $(100\ \mathrm{Mpc})^3$ high- and $(400\ \mathrm{Mpc})^3$ low-resolution volumes respectively.

At high resolution there is again excellent agreement between the emulator prediction and the observed data, with reduced $\chi^2_\nu = 1.15$. The high-resolution simulation resolves the largest range of stellar mass in the SMF, from $\approx 10^{8.6}\,\mathrm{M}_\odot$ to $\approx 10^{11.5}\,\mathrm{M}_\odot$. There is a dip around a mass of $10^{10.2}\ \mathrm{M}_\odot$ and a slight bump around the knee of the mass function. but the maximum deviation from the data is less than 5 per cent. It seems that the emulator was unable to predict the dip, and the best-fitting simulation falls outside of the predicted errors. Comparing the predicted errors between the different resolutions, it is clear that the high-resolution simulation has the largest predicted error. This is due to it using the smallest box size. This causes the emulator prediction to be too "smooth" when compared with simulation results. The deviation at the dip is less than the 1 σ uncertainty due to cosmic variance. The small box size $(100\ \mathrm{Mpc})^3$ used for calibration at high resolution, limits the mass range that can be used to fit the gas fractions to halo masses lower than $6\times10^{13}\ \mathrm{M}_\odot$. This leaves only two data points to compare to. The agreement in the fitted range is however very good.

Comparing the best-fitting subgrid parameter values for the high-resolution model to those for intermediate resolution (Table 2.2), we see that the stellar feedback requires about twice as much energy and about half as high a kick velocity. This reflects the need for stronger stellar feedback when higher gas densities are resolved and the fact that feedback can be efficient down to smaller wind velocities in the lower-mass haloes that remained unresolved at intermediate resolution. While the AGN heating temperatures are very similar, the high-resolution simulations require a much smaller slope of the black hole accretion rate boost factor, $\beta_{\rm BH} = 0.038$ (where zero corresponds to no boost) versus $\beta_{\rm BH} = 0.514$ at intermediate resolution. Since the high-resolution simulation can resolve higher gas densities, and hence higher black hole accretion rates, we do not need to boost

ing the subgrid parameters constant. The stellar mass bias changed from $\log_{10} b_* = 0.031$ to 0.026, the cosmic variance bias changed from $b_{\rm CV} = 1.014$ to 0.995 and the HSE bias from $b_{\rm HSE} = 0.745$ to 0.743. The bias values changed by a negligible amount with respect to the 16th–84th percentile confidence levels, for both b_* and $b_{\rm HSE}$ the change is less than 3 per cent of the 16th–84th percentile range. For $b_{\rm CV}$ the change is ~ 15 per cent of the 16th–84th percentile range. The values we report in Table 2.7 use the most up-to-date Driver et al. (2022) data.

the accretion rate as much.

At low resolution the agreement with the data is also very good, with reduced $\chi^2_\nu = 0.95$. Now it is the stellar mass range that is very limited, $M_* \approx 10^{11.17}\,\mathrm{M}_\odot$ to $M_* \approx 10^{11.5}\,\mathrm{M}_\odot$, which includes only two data points. The larger box size of $(400\ \mathrm{Mpc})^3$ allows for the use of the two Akino et al. (2022) weak lensing data points as well as five X-ray data points for fitting the cluster gas fractions. However, the high-mass plateau of the gas fractions remains out of reach for this box size. The comparison of the best-fitting subgrid parameter values of the low-resolution model to those of the higher-resolution simulations (Table 2.2) is difficult to interpret because the low-resolution model requires a much lower threshold density for star formation, a much higher black hole seed mass, and does not include any stellar feedback.

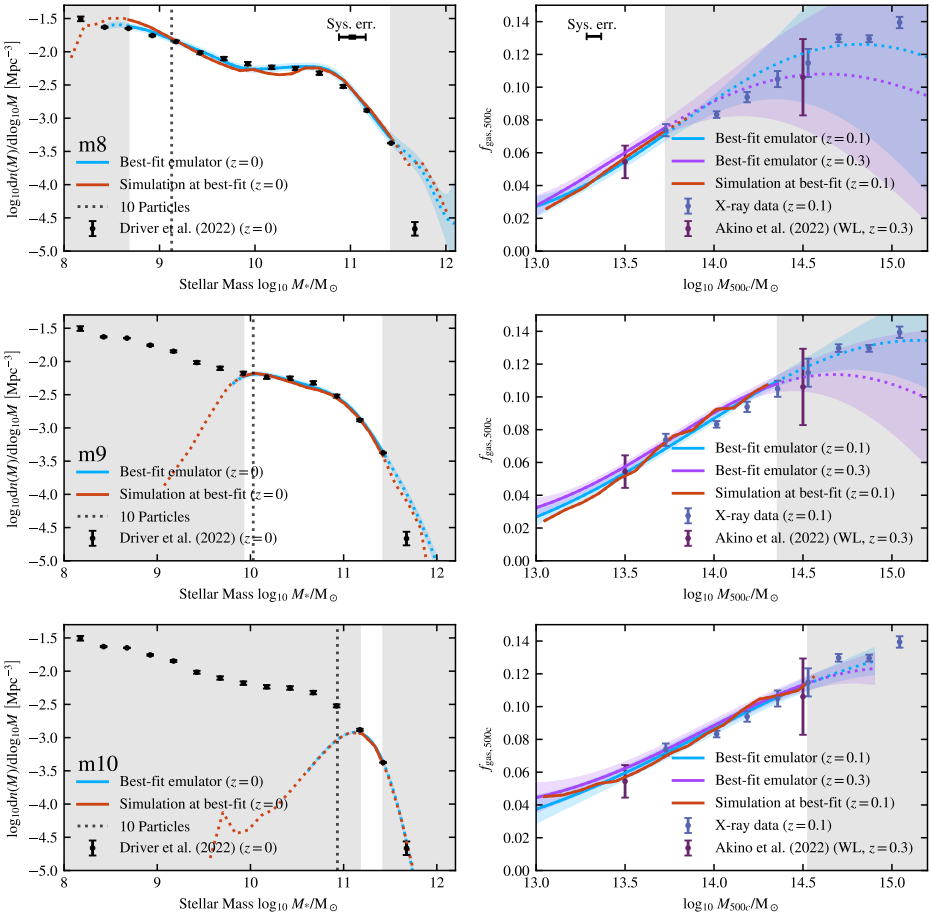


Figure 2.6: Comparison of the best-fitting models to the observed galaxy stellar mass function (SMF; left column) at $z = 0$ and observed cluster gas fractions (right column). The top, middle and bottom rows show results for high-, intermediate- and low-resolution simulations, respectively. The observations are plotted as points with error bars (black: Driver et al. (2022) SMF at $z = 0$, dark blue: compilation of X-ray data at $z = 0.1$, dark magenta: Akino et al. (2022) weak lensing data at $z = 0.3$). Each panel shows the best-fitting emulator prediction as a blue curve, the emulator uncertainty as a blue shaded region, and the result from a simulation using the best-fitting subgrid parameter values in a $(100\ \mathrm{Mpc})^3$, $(200\ \mathrm{Mpc})^3$, and $(400\ \mathrm{Mpc})^3$ volume for high, intermediate, and low resolution, respectively, as a red curve. For $f_{\mathrm{gas},500c}$ we only plot the best-fitting simulation result at $z = 0.1$ in red, and leave out the result at $z = 0.3$ to avoid clutter. For the cluster gas fractions, besides showing in blue the $z = 0.1$ emulator that should be compared with the dark blue X-ray data, we also show the $z = 0.3$ emulator, in magenta, that is used to fit the dark magenta Akino et al. (2022) weak lensing data. The grey regions indicate the mass ranges that are excluded from the fitting, see also Table 2.3. The model predictions are shown using dotted lines in these excluded ranges. The vertical dotted line in the left panels indicates a mass corresponding to ten stellar particles. The SMF and X-ray gas fraction data have been shifted by the best-fitting observational bias factors (see Table 2.7), which are however negligible for the SMF. The SMF from the best-fitting simulation includes Eddington bias (see §2.3.1) in line with how the emulator is trained. The systematic errors given by the priors on the bias parameters are shown as points with error bars in the top panels. At each resolution we obtain excellent agreement between the emulator, a simulation with the best-fitting parameters, and the observational data.

As we obtain a good fit to the same data for each of the three resolutions, we conclude that we have good 'weak convergence' between the three resolutions, using the terminology of Schaye et al. (2015). The FLAMINGO suite includes high-, intermediate-, and low-resolution simulations that were run with our fiducial subgrid parameter values in volumes with side lengths of 1, 2.8, and 1 Gpc, respectively. For a comparison of these models with other data, we refer to Schaye et al. (2023).

2.6.4 Feedback variations

One of the goals of FLAMINGO is to investigate the impact of feedback on cosmological observables. In this section we show how we use emulators to calibrate simulations to produce gas fractions or SMFs that have been shifted away

from their fiducial, observed values. We focus mostly on changes to the gas fractions, as previous work has shown that baryon fractions in groups and clusters anti-correlate with the baryonic suppression of the matter power spectrum on the scales relevant for current and next generation surveys (e.g. Semboloni et al., 2013; Van Daalen et al., 2020; Debackere et al., 2020; Salcido et al., 2023). For clusters, the gas fractions dominate over the stellar fraction when computing the baryon fractions (the stellar mass content of haloes becomes important at smaller scales). While most of our variations use our fiducial thermal AGN feedback model, we will also calibrate a model that uses kinetic, jet-like AGN feedback.

To quantify the effect of reasonable changes in the astrophysics, we include a set of feedback variations in the simulation suite. These simulations should at least bracket the uncertainty in the cluster gas fraction data, while fitting the SMF data. Previous works created variations of subgrid physics based directly on the values of certain subgrid parameters. For example, the BAHAMAS project (McCarthy et al., 2018) varied the AGN heating temperature by ±0.2 dex, which resulted in very small changes to the SMF and cluster gas fractions that roughly bracketed the observational uncertainty. To arrive at the values of the subgrid parameters for our runs, we make use of the emulators and we will allow all fitted subgrid parameters to vary. Our variations are based on systematically shifting of the data, based on their uncertainties, making the variations less reliant on the sub-grid model used. We also include models with gas fractions that are probably ruled out observationally, because we anticipate these will be useful to gain insight into the effect of baryonic feedback on other cosmological observables.

The variations are run at intermediate resolution. We use the fiducial values of the observational bias factors listed in Table 2.7. For the gas fraction variations, the SMF data are kept the same except for one variation, where we systematically reduce all observed stellar masses. The $f_{\rm gas}$ data are shifted up by 2σ and down by 2, 4 and 8σ for the fgas+2σ, -2σ, -4σ and -8σ models respectively, where σ is the error obtained from bootstrapping for the X-ray data, or the error on the fit for the weak lensing data from Akino et al. (2022), as discussed in §2.3.2. We systematically shift all the data by $N\sigma$ under the assumption that the errors in the gas fraction are mostly systematic and correlated. We shift in steps of 2 and 4σ instead of a smaller shift (for example 1σ) as the cluster-to-cluster scatter is much larger than the errors we found from bootstrapping (see §2.3.2). We also create a models that vary the SMF. As the baryonic suppression is sensitive to the total baryon fraction (see e.g. Salcido et al., 2023), we include these variations to investigate the effect of changes in the baryon fraction at a constant gas fraction, and to see the effect of changing the stellar fractions. For these variations, we systematically shift the SMF data to lower masses according to the 1σ given by the stellar mass bias (0.14 dex; §2.3.1). For the M*–1σ model we use the fiducial gas fractions and for the fgas–4σ + M* – 1σ model we simultaneously shift the X-ray and weak lensing gas fractions down by 4σ.

Table 2.8: best-fitting values for the subgrid parameters for the feedback variations at intermediate resolution. The columns list the name of the variation, the number of σ by which the observed $f_{\rm gas}$ data was shifted, and for each parameter the median and 16th to 84th percentile confidence level (CL), and the best-fitting (i.e. maximum likelihood) fiducial values. Note that for the jet AGN model the seventh and eighth columns show $v_{\rm jet}$ instead of the heating temperature, while for the other feedback variations they show $\Delta T_{\rm AGN}$.

		$f_{\rm SN}$		$\Delta v_{\rm SN}$ [km s^{-1}]		$\Delta T_{\rm AGN}$ [K] or $v_{\rm jet}$ [km s^{-1}]		$\beta_{\rm BH}$	
Variation	σ	Median+CL	best-fitting	Median+CL	best-fitting	Median+CL	best-fitting	Median+CL	best-fitting
fgas+2σ	+2	$0.22^{+0.09}_{-0.08}$	0.219	525^{+151}_{-186}	577	$10^{7.69^{+0.16}_{-0.13}}$	$10^{7.71}$	$0.58^{+0.10}_{-0.10}$	0.554
Fiducial	0	$0.20^{+0.11}_{-0.09}$	0.238	479^{+167}_{-197}	562	$10^{7.84^{+0.18}_{-0.20}}$	$10^{7.95}$	$0.55^{+0.15}_{-0.16}$	0.514
fgas-2σ	-2	$0.21^{+0.08}_{-0.07}$	0.206	478^{+149}_{-179}	552	$10^{8.03^{+0.14}_{-0.16}}$	$10^{8.08}$	$0.54^{+0.10}_{-0.09}$	0.497
fgas-4σ	-4	$0.20^{+0.08}_{-0.07}$	0.191	479^{+167}_{-162}	532	$10^{8.18^{+0.13}_{-0.13}}$	$10^{8.21}$	$0.51^{+0.09}_{-0.09}$	0.482
fgas-8σ	-8	$0.15^{+0.07}_{-0.06}$	0.145	417^{+156}_{-154}	483	$10^{8.36^{+0.09}_{-0.11}}$	$10^{8.40}$	$0.49^{+0.07}_{-0.08}$	0.462
M*$-\sigma$	0	$0.30^{+0.10}_{-0.10}$	0.322	537^{+124}_{-198}	608	$10^{7.98^{+0.14}_{-0.17}}$	$10^{8.06}$	$0.68^{+0.11}_{-0.10}$	0.626
M*$-\sigma$ + fgas-4σ	-4	$0.25^{+0.10}_{-0.08}$	0.261	490^{+127}_{-174}	557	$10^{8.25^{+0.13}_{-0.13}}$	$10^{8.27}$	$0.65^{+0.09}_{-0.09}$	0.620
Jet	0	$0.19^{+0.07}_{-0.06}$	0.166	562^{+196}_{-164}	477	977^{+311}_{-236}	836	$0.54^{+0.10}_{-0.12}$	0.597
Jet + fgas-4σ	-4	$0.18^{+0.08}_{-0.06}$	0.176	524^{+200}_{-162}	527	1949^{+238}_{-251}	1995	$0.44^{+0.07}_{-0.08}$	0.439

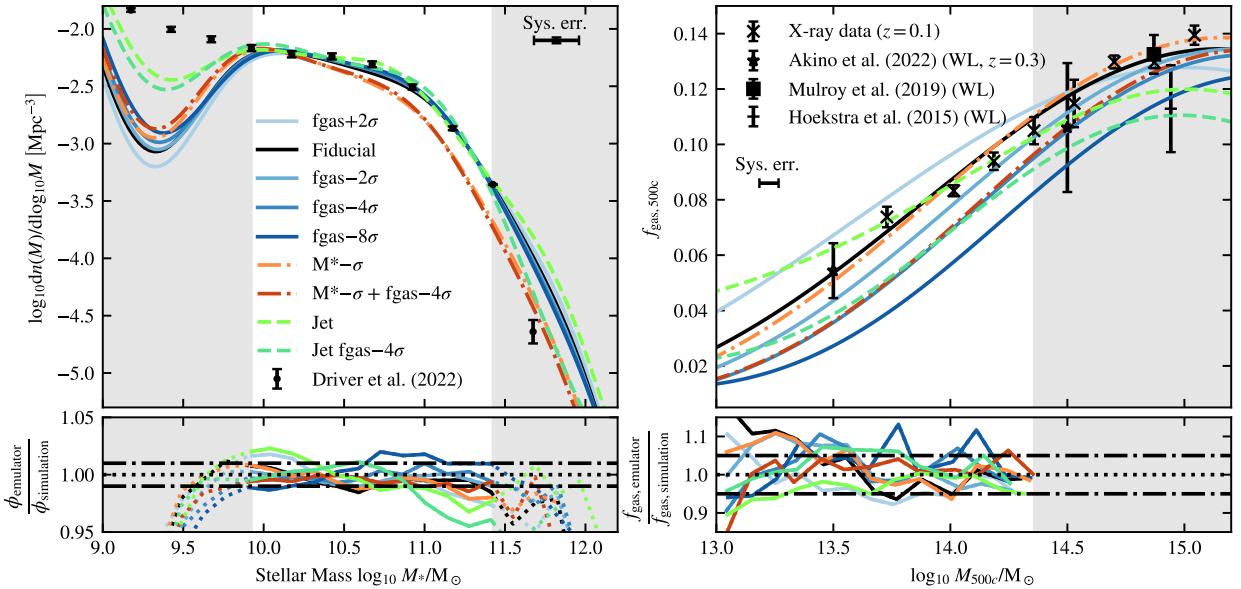


Figure 2.7: Top left and right panels: The emulator predictions for the SMF and gas fractions, respectively, for the feedback variations and the fiducial model (different colors, as indicated in the legend). The observations are shown as black points with error bars. In the top corners of the panels we indicate the assumed systematic errors in the data from the priors on the fitted biases. The bottom panels show the ratio of the emulator prediction and a $(200\ \mathrm{Mpc})^3$ simulation run with the same parameters. In both panels the black dotted line indicates a ratio of one. For the SMF ($f_{\mathrm{gas},500c}$), the black dot-dashed lines indicate deviations of 1 per cent (5 per cent). We only show the cluster gas fraction emulator prediction at $z = 0.1$ and leave out the $z = 0.3$ gas fraction results to avoid clutter. The excluded mass range for fitting is indicated by the grey regions (see also Table 2.3.) We use the emulators to make a direct mapping between our subgrid physics models and systematic shifts in the observations, based on the observational errors.

The best-fitting subgrid parameter values for the feedback variations can be found in Table 2.8. The changes in the subgrid parameters with respect to the fiducial model are small. As expected, the AGN subgrid parameters bracket the fiducial values, with the fgas−2σ model having a slightly higher AGN feedback temperature. As could already be seen in Fig. 2.4, the gas fraction is very sensitive to $\Delta T_{\rm AGN}$, which varies by only 0.37 dex between the fgas+2σ and −2σ models, in good agreement with BAHAMAS. The fgas−4σ and −8σ models follow this trend. Changes in the gas fractions are driven mainly by changes in $\Delta T_{\rm AGN}$. Going from the fgas−4σ to the M*−1σ + fgas−4σ model, the biggest change is seen in $f_{\rm SNII}$ and $\beta_{\rm BH}$, as expected from Fig. 2.4. The increase in the BH accretion boost factor is required to compensate for the removal of gas by the increased supernova energy.

The feedback models are compared with the fiducial model and the calibration data in Fig. 2.7. In the top two panels we show the emulator predictions for the SMF and the gas fractions for each of the variations. Within the fitted mass ranges there is excellent agreement for the SMF between all the different cluster gas fraction variations. There is good agreement between $f_{\rm gas}$ for the $f_{\rm gas}-4\sigma$ and the SMF−1σ + $f_{\rm gas}-4\sigma$ variations. In the bottom panels we compare the emulator predictions to the results of (200 Mpc)^3 simulations run with the best-fitting parameters. For the SMF, we see that the emulator predictions are accurate at around the per cent level, with only the jet model fgas−4σ deviating by ≈ 5 per cent. For $f_{\rm gas}$, all predictions are accurate to ≈ 10 per cent, and most predictions are accurate to within ≈ 5 per cent. The accuracy is slightly better than the expected emulator accuracy from cross-checks (see Table 2.6). We conclude that by allowing for small adjustments to four subgrid parameters, we are able to vary specific observables while keeping others constant.

In addition to the parameter variations, we also calibrate a different implementation of AGN feedback. As described in §2.2.3 this model uses kinetic bipolar kicks instead of thermal injections to distribute AGN feedback energy around accreting BHs[8]. As the subgrid model differs fundamentally from the fiducial model, we run a new Latin hypercube with 32 intermediate-resolution simulations in (200 Mpc)^3 volumes. The subgrid parameter ranges for this hypercube can be found in Table 2.9. To construct the emulator, we again follow the prescription of Section 2.4 and we again verify its accuracy using cross-checks (see Table 2.6). The goal is to have a simulation with a different implementation of AGN feedback calibrated to the same observables as the fiducial implementation. We therefore use the same fitting limits, methods and likelihoods as for the fidu-

[8] Due to a bug, the calibration of the jet models was done using a version of the model where the jets are launched along the z-direction of the simulation box, instead of along the spin axis of the black hole. We have verified that this leads to small differences, in agreement with the results reported by Huško et al. (2023), who showed that the directionality of the jets has little effect. When using the correct implementation, the agreement with the emulator of the SMF becomes slightly better for both runs that use jets; and for fgas, the agreement only worsens outside the range used for calibration.

cial intermediate-resolution model. For the jet model we fit to both the fiducial data and to the perturbed data used to calibrate the $f_{\rm gas}-4\sigma$ model. The resulting medians and best-fitting values can be found in Table 2.8.

The jet models are shown as the green lines in Fig. 2.7. They show some differences from the fiducial thermal AGN feedback models. The jet models fit the knee of the SMF slightly better by having slightly more galaxies with $M_* \approx 10^{10.7}\,{\rm M}_\odot$. The difference at the very low-mass end of the SMF, below the fitted range, is due to the fact that the bug in the threshold of star formation for zero metallicity gas (see footnote 3) was fixed for the jet models. The $f_{\rm gas}-4\sigma$ jet model also has a significant reduction in the number of galaxies with masses above our fitting limit, thus yielding a SMF with a steeper high-mass cut off. However, the bottom panel suggests that this is at least partially explained by the fact that the emulator under-predicts the number density by a few per cent. Compared with the thermal AGN models fit to the same data, the jet models predict higher gas fractions in groups ($\rm M_{500c} \sim 10^{13}\,M_\odot$), where there is, however, no observational data. From the bottom panels we can see that for $f_{\rm gas}$ the accuracy of the jet emulator does not differ significantly from the emulator for the thermal AGN feedback models.

2.7 Conclusions

In order to fully exploit the large-scale structure data that will become available with surveys like *Euclid* and LSST, we need to acquire a deeper understanding of how baryonic effects, like AGN and stellar feedback, impact the matter distribution. The most self-consistent way of experimenting with these effects is through the use of cosmological hydrodynamical simulations. The FLAMINGO project provides such simulations in volumes sufficiently large to study the evolution of large-scale structure and massive galaxy clusters for different numerical resolutions, cosmologies and astrophysical models.

As feedback processes originate on unresolved scales, we have to add them via subgrid prescriptions. However, because these subgrid models are theoretically not well constrained, they need to be calibrated to reproduce a relevant set of observables. Previous simulation projects like EAGLE (Schaye et al., 2015; Crain et al., 2015), IllustrisTNG (Pillepich et al., 2018), BAHAMAS McCarthy et al. (2017, 2018) and SIMBA Davé et al. (2019) achieved good agreement with data by varying subgrid parameters by hand until the simulation lined up with the target observations. However, for cosmology a more robust and objective calibration method is desirable, particularly if it can also be used to predict the effect of subgrid variations that have not been simulated directly.

To create a robust method of calibration, we make use of machine learning, specifically Gaussian process emulators. Instead of emulating the effects of changes in the cosmological parameters, which is becoming a common ap-

plication of machine learning in cosmology, we emulate the observables that we want to match to observations as a function of a set of subgrid parameters. For three different numerical resolutions, which span a factor of 64 in particle mass, we train an emulator on 32 input simulations where we vary the four most impactful subgrid parameters, two of which relate to stellar feedback and two of which relate to AGN feedback (Section 2.2). In addition, we train an emulator for another intermediate-resolution implementation of AGN feedback, which uses jets (i.e., directed kinetic feedback) instead of injecting the feedback energy thermally. At each resolution we run simulations with 360^3 gas particles, implying a $(100\ \mathrm{Mpc})^3$, $(200\ \mathrm{Mpc})^3$ and $(400\ \mathrm{Mpc})^3$ volume for FLAMINGO high [m8], intermediate [m9] and low [m10] resolution, respectively. We then use MCMC to fit the emulator to carefully selected observational data. We repeat the same procedure for each resolution, and only change the fitted mass ranges to account for resolution and box size limitations. Additionally, we have created a set of subgrid physics implementations based on fitting the emulators to the data after systematically shifting it by $N\sigma$.

We calibrate to the observed low-redshift galaxy stellar mass function (SMF) from the GAMA survey and a compilation of group and cluster gas fraction measurements based on X-ray and weak lensing data. A novel aspect of our approach is that we also fit for possible observational biases (i.e., systematic errors). We account for biases in the stellar mass and the cluster mass inferred from X-ray data under the assumption of hydrostatic equilibrium, as well as for the effect of cosmic variance on the SMF. In addition, we account for the effect of random errors in the observed stellar mass on the SMF (i.e., Eddington bias) by randomly perturbing the simulated stellar masses(Section 2.3). The observational biases are only fit during the calibration of the intermediate-resolution simulations and the best-fitting values are then also applied to the other resolutions.

Our main conclusions are:

- By carefully setting up the subgrid parameter space, we were able to train emulators that are more accurate than the target observational constraints (Fig. 2.3).

- The emulator framework enables simultaneously fitting for subgrid parameters and observational biases. For FLAMINGO, the posteriors found for the biases are driven by and in agreement with the priors. We find a negligible value for the stellar mass and cosmic variance error, and a hydrostatic bias of $b_{\mathrm{HSE}} = 0.743$.

- Emulators can be used to make parameter sweeps, i.e. plots showing how the trained relation depends on the value of a single subgrid parameter (Fig. 2.4). As the emulators give the continuous response of the trained relation to changes in subgrid parameters, emulators can be used to gain a

deeper understanding of how the observable relations are affected by the subgrid models.

- The parameter space that we explore is devoid of major degeneracies between the subgrid parameters. The emulator+MCMC framework finds a single best-fitting solution (Fig. 2.5). We note that this is partially by construction, as parameters that had major degeneracies were omitted from the parameter space (see Section 2.2). For future work it might be interesting to see if these degeneracies can be solved by fitting the model to additional observational data.

- At each resolution we find excellent agreement between the best-fitting model and the calibration data (Fig. 2.6).

- The emulator framework can be used to map observational uncertainties onto changes in subgrid parameters. By fitting the emulator to variations in gas fractions and the SMF, we produce a set of simulations for which specific observables are varied while keeping others constant (Fig. 2.7). As the model variations are directly tied to observations, the resulting simulations can be used to quantify the effect of uncertainties in the calibration data on the predictions for other observables.

- We used the emulator framework to calibrate a different implementation of the model, which we did for kinetic AGN feedback (in contrast with the thermal AGN feedback used our fiducial model; Fig. 2.7). By making different models match the same calibration observations, the simulations can be used to quantify the uncertainty in predictions for other observables due to uncertainties in the underlying physics.

We have used Gaussian process emulators to create a close link between subgrid models and observations. By creating a robust statistical framework for calibration, future hydrodynamical simulations will be able to use available and upcoming data to constrain the subgrid physics and to quantify the uncertainty in the predictions of simulations that remains after the models have been constrained to fit particular sets of data. In this work we have focused on calibrating simulations using different resolutions, and a single variation of the implementation of AGN feedback. For future work the same framework could be used to get agreement between different simulation codes and subgrid models for specific observables. In this way we could improve our understanding of the degeneracies between different methods and the uncertainties in their predictions.

In the companion paper Schaye et al. (2023) we present the large-volume FLAMINGO simulations that use the calibrated parameter values that we ob-

tained here. More information on and visualisations of the FLAMINGO simulations can be found on the website.[9]

2

Acknowledgements

This work is partly funded by Vici grant 639.043.409, Veni grant 639.041.751 and research programme Athena 184.034.002 from the Dutch Research Council (NWO). This work used the DiRAC@Durham facility managed by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). The equipment was funded by BEIS capital funding via STFC capital grants ST/K00042X/1, ST/P002293/1, ST/R002371/1 and ST/S002502/1, Durham University and STFC operations grant ST/R000832/1. DiRAC is part of the National e-Infrastructure.EC is supported by the funding from the European Union' s Horizon 2020 research and innovation programme under the Marie Skodowska-Curie grant agreement No 860744 (BiD4BESt). We gratefully acknowledge financial support from the Swiss National Science Foundation (SNSF) under funding reference 200021_213076. F. H. would like to acknowledge support from the Science Technology Facilities Council through a CDT studentship (ST/P006744/1). I.V. gratefully acknowledges UKRI (EP/W011956/1) and Wellcome (218261/Z/19/Z) funding. ARJ, CGL, JCH and CSF acknowledge the STFC consolidated grants ST/T000244/1 and ST/X001075/1. This research project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 769130). The research in this paper made use of the SWIFT open-source simulation code (http://www.swiftsim.com, Schaller et al. (2018a)) version 0.9.0.

[9] https://flamingo.strw.leidenuniv.nl/

References

Abbott T. M. C., et al., 2022, Phys. Rev. D, 105, 023520

Acuto A., McCarthy I. G., Kwan J., Salcido J., Stafford S. G., Font A. S., 2021, MNRAS, 508, 3519

Akino D., et al., 2022, PASJ, 74, 175

Alam S., et al., 2021, Phys. Rev. D, 103, 083533

Ambikasaran S., Foreman-Mackey D., Greengard L., Hogg D. W., O'Neil M., 2015a, IEEE Transactions on Pattern Analysis and Machine Intelligence, 38, 252

Ambikasaran S., Foreman-Mackey D., Greengard L., Hogg D. W., O'Neil M., 2015b, IEEE Transactions on Pattern Analysis and Machine Intelligence, 38, 252

Angulo R. E., Zennaro M., Contreras S., Aricò G., Pellejero-Ibañez M., Stücker J., 2021, MNRAS, 507, 5869

Aricò G., Angulo R. E., Contreras S., Ondaro-Mallea L., Pellejero-Ibañez M., Zennaro M., 2021, MNRAS, 506, 4070

Ayromlou M., Nelson D., Yates R. M., Kauffmann G., Renneby M., White S. D. M., 2021, MNRAS, 502, 1051

Bahé Y. M., McCarthy I. G., King L. J., 2012, MNRAS, 421, 1073

Bahé Y. M., et al., 2022, MNRAS, 516, 167

Baldry I. K., et al., 2012, MNRAS, 421, 621

Baudin M., Christopoulou M., Collette Y., Martinez J.-M., 2012, pyDOE: The experimental design package for Python, https://github.com/tisimst/pyDOE

Becker M. R., Kravtsov A. V., 2011, ApJ, 740, 25

Behroozi P., Wechsler R. H., Hearin A. P., Conroy C., 2019, MNRAS, 488, 3143

Bernardi M., Meert A., Sheth R. K., Vikram V., Huertas-Company M., Mei S., Shankar F., 2013, MNRAS, 436, 697

Bernardi M., Meert A., Sheth R. K., Fischer J. L., Huertas-Company M., Maraston C., Shankar F., Vikram V., 2017, MNRAS, 467, 2217

Bondi H., Hoyle F., 1944, MNRAS, 104, 273

Booth C. M., Schaye J., 2009, MNRAS, 398, 53

Borrow J., Borrisov A., 2020, The Journal of Open Source Software, 5, 2430

Borrow J., Schaller M., Bahe Y. M., Schaye J., Ludlow A. D., Ploeckinger S., Nobels F. S. J., Altamura E., 2022a, arXiv e-prints, p. arXiv:2211.08442

Borrow J., Schaller M., Bower R. G., Schaye J., 2022b, MNRAS, 511, 2367

Bower R. G., Vernon I., Goldstein M., Benson A. J., Lacey C. G., Baugh C. M., Cole S., Frenk C. S., 2010, MNRAS, 407, 2017

Bower R. G., Schaye J., Frenk C. S., Theuns T., Schaller M., Crain R. A., McAlpine S., 2017, MNRAS, 465, 32

Cañas R., Elahi P. J., Welker C., del P Lagos C., Power C., Dubois Y., Pichon C., 2019, MNRAS, 482, 2039

Chabrier G., 2003, PASP, 115, 763

Chaikin E., Schaye J., Schaller M., Benítez-Llambay A., Nobels F. S. J., Ploeckinger S., 2022a, arXiv e-prints, p. arXiv:2211.04619

Chaikin E., Schaye J., Schaller M., Bahé Y. M., Nobels F. S. J., Ploeckinger S., 2022b, MNRAS, 514, 249

Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2015, GALFORM: Galactic modeling (ascl:1510.005)

Crain R. A., et al., 2015, MNRAS, 450, 1937
D'Souza R., Vegetti S., Kauffmann G., 2015, MNRAS, 454, 4027
van Daalen M. P., Schaye J., Booth C. M., Dalla Vecchia C., 2011, MNRAS, 415, 3649
van Daalen M. P., McCarthy I. G., Schaye J., 2020, MNRAS, 491, 2424
Dalla Vecchia C., Schaye J., 2008, MNRAS, 387, 1431
Dalla Vecchia C., Schaye J., 2012, MNRAS, 426, 140
Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, MNRAS, 486, 2827
DeRose J., et al., 2021, arXiv e-prints, p. arXiv:2105.13547
Debackere S. N. B., Schaye J., Hoekstra H., 2020, MNRAS, 492, 2285
Di Matteo T., Colberg J., Springel V., Hernquist L., Sijacki D., 2008, ApJ, 676, 33
Driver S. P., Robotham A. S. G., 2010, MNRAS, 407, 2131
Driver S. P., et al., 2022, MNRAS, 513, 439
Eckert D., et al., 2016, A&A, 592, A12
Eddington A. S., 1913, MNRAS, 73, 359
Elahi P. J., Poulton R., Canas R., 2019a, The Astrophysics Source Code Library, p. ascl:1911.020
Elahi P. J., Cañas R., Poulton R. J. J., Tobar R. J., Willis J. S., Lagos C. d. P., Power C., Robotham A. S. G., 2019b, PASA, 36, e021
Elbers W., Frenk C. S., Jenkins A., Li B., Pascoli S., 2021, MNRAS, 507, 2614
Elliott E. J., Baugh C. M., Lacey C. G., 2021, MNRAS, 506, 4011
Euclid Collaboration et al., 2019, MNRAS, 484, 5509
Euclid Collaboration et al., 2020, A&A, 642, A191
Ferland G. J., et al., 2017, Rev. Mex. Astron. Astrofis., 53, 385
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Frieman J. A., Turner M. S., Huterer D., 2008, ARA&A, 46, 385
GPy since 2012, GPy: A Gaussian process framework in Python, http://github.com/SheffieldML/GPy
Gardner J., Pleiss G., Weinberger K. Q., Bindel D., Wilson A. G., 2018, GPyTorch: Blackbox Matrix-Matrix Gaussian Process Inference with GPU Acceleration. Curran Associates, Inc., https://proceedings.neurips.cc/paper/2018/file/27e8e17134dd7083b050476733207ea1-Paper.pdf
Giri S. K., Schneider A., 2021, arXiv e-prints, p. arXiv:2108.08863
Gonzalez A. H., Sivanandam S., Zabludoff A. I., Zaritsky D., 2013, ApJ, 778, 14
de Graaff A., Trayford J., Franx M., Schaller M., Schaye J., van der Wel A., 2022, MNRAS, 511, 2544
Greengard L., Rokhlin V., 1987, Journal of Computational Physics, 73, 325
Hahn O., Rampf C., Uhlemann C., 2021, MNRAS, 503, 426
Häring N., Rix H.-W., 2004, ApJ, 604, L89
Heitmann K., Higdon D., White M., Habib S., Williams B. J., Lawrence E., Wagner C., 2009, ApJ, 705, 156
Heitmann K., et al., 2016a, ApJ, 820, 108
Heitmann K., et al., 2016b, ApJ, 820, 108
Heymans C., et al., 2021, A&A, 646, A140
Hirschmann M., Dolag K., Saro A., Bachmann L., Borgani S., Burkert A., 2014, MNRAS, 442, 2304
Hoekstra H., Herbonnet R., Muzzin A., Babul A., Mahdavi A., Viola M., Cacciato M., 2015, MNRAS, 449, 685
Huško F., Lacey C. G., Schaye J., Schaller M., Nobels F. S. J., 2022, MNRAS,
Huško F., Lacey C. G., Schaye J., Nobels F. S. J., Schaller M., 2023, arXiv e-prints, p. arXiv:2307.01409
Jo Y., et al., 2023, ApJ, 944, 67
Kaviraj S., et al., 2017, MNRAS, 467, 4739
Kennicutt Jr. R. C., 1998, ApJ, 498, 541
Kennicutt Jr. R. C., et al., 2007, ApJ, 671, 333
Kugel R., Borrow J., 2022, Journal of Open Source Software, 7, 4240
Lacey C. G., et al., 2016, Monthly Notices of the Royal Astronomical Society, 462, 3854
Laganá T. F., Martinet N., Durret F., Lima Neto G. B., Maughan B., Zhang Y. Y., 2013, A&A, 555, A66
Le Brun A. M. C., McCarthy I. G., Schaye J., Ponman T. J., 2014, MNRAS, 441, 1270

Li C., White S. D. M., 2009, MNRAS, 398, 2177
Lin Y.-T., Stanford S. A., Eisenhardt P. R. M., Vikhlinin A., Maughan B. J., Kravtsov A., 2012, ApJ, 745, L3
Lovisari L., Reiprich T. H., Schellenberger G., 2015, A&A, 573, A118
Lovisari L., et al., 2020, ApJ, 892, 102
Macquart J. P., et al., 2020, Nature, 581, 391
Maughan B. J., Jones C., Forman W., Van Speybroeck L., 2008, ApJS, 174, 117
McAlpine S., Bower R. G., Rosario D. J., Crain R. A., Schaye J., Theuns T., 2018, MNRAS, 481, 3118
McCarthy I. G., Schaye J., Bird S., Le Brun A. M. C., 2017, MNRAS, 465, 2936
McCarthy I. G., Bird S., Schaye J., Harnois-Deraps J., Font A. S., van Waerbeke L., 2018, MNRAS, 476, 2999
McKay M. D., Beckman R. J., Conover W. J., 1979, Technometrics, 21, 239
Mead A. J., Peacock J. A., Heymans C., Joudaki S., Heavens A. F., 2015, MNRAS, 454, 1958
Moran K. R., et al., 2022, MNRAS,
Moster B. P., Naab T., White S. D. M., 2018, MNRAS, 477, 1822
Mulroy S. L., et al., 2019, MNRAS, 484, 60
Nicastro F., et al., 2018, Nature, 558, 406
Oh B. K., An H., Shin E.-j., Kim J.-h., Hong S. E., 2022, MNRAS, 515, 693
Pakmor R., et al., 2022, arXiv e-prints, p. arXiv:2210.10060
Pearson R. J., et al., 2017, MNRAS, 469, 3489
Pillepich A., et al., 2018, MNRAS, 473, 4077
Planck Collaboration et al., 2020, A&A, 641, A6
Ploeckinger S., Schaye J., 2020, MNRAS, 497, 4857
Pratt G. W., et al., 2010, A&A, 511, A85
Rasmussen J., Ponman T. J., 2009, MNRAS, 399, 239
Rasmussen C. E., Williams C. K. I., 2006, Gaussian Processes for Machine Learning. MIT Press Ltd
Rasmussen C., Bousquet O., Luxburg U., Rätsch G., 2004, Advanced Lectures on Machine Learning: ML Summer Schools 2003, Canberra, Australia, February 2 - 14, 2003, Tübingen, Germany, August 4 - 16, 2003, Revised Lectures, 63-71 (2004), 3176
Rezzolla L., Barausse E., Dorband E. N., Pollney D., Reisswig C., Seiler J., Husa S., 2008, Phys. Rev. D, 78, 044002
Richings A. J., Schaye J., 2016, MNRAS, 458, 270
Riess A. G., et al., 2022, ApJ, 934, L7
Rodrigues L. F. S., Vernon I., Bower R. G., 2017, MNRAS, 466, 2418
Salcido J., McCarthy I. G., Kwan J., Upadhye A., Font A. S., 2023, Monthly Notices of the Royal Astronomical Society, 523, 2247
Sanderson A. J. R., O'Sullivan E., Ponman T. J., Gonzalez A. H., Sivanandam S., Zabludoff A. I., Zaritsky D., 2013, MNRAS, 429, 3288
Schaller M., Gonnet P., Draper P. W., Chalk A. B. G., Bower R. G., Willis J., Hausammann L., 2018a, SWIFT: SPH With Inter-dependent Fine-grained Tasking (ascl:1805.020)
Schaller M., Gonnet P., Draper P. W., Chalk A. B. G., Bower R. G., Willis J., Hausammann L., 2018b, The Astrophysics Source Code Library, p. ascl:1805.020
Schaller M., et al., 2023, arXiv e-prints, p. arXiv:2305.13380
Schaye J., Dalla Vecchia C., 2008, MNRAS, 383, 1210
Schaye J., et al., 2010, MNRAS, 402, 1536
Schaye J., et al., 2015, MNRAS, 446, 521
Schaye J., et al., 2023, arXiv e-prints, p. arXiv:2306.04024
Schneider A., Teyssier R., 2015, J. Cosmology Astropart. Phys., 2015, 049
Schneider A., Stoira N., Refregier A., Weiss A. J., Knabenhans M., Stadel J., Teyssier R., 2020, J. Cosmology Astropart. Phys., 2020, 019
Semboloni E., Hoekstra H., Schaye J., van Daalen M. P., McCarthy I. G., 2011, MNRAS, 417, 2020
Semboloni E., Hoekstra H., Schaye J., 2013, MNRAS, 434, 148
Smith G. P., et al., 2016, MNRAS, 456, L74

Springel V., 2005, MNRAS, 364, 1105
Sun M., Voit G. M., Donahue M., Jones C., Forman W., Vikhlinin A., 2009, ApJ, 693, 1142
Turner M. S., 2022, arXiv e-prints, p. arXiv:2201.04741
Velliscig M., van Daalen M. P., Schaye J., McCarthy I. G., Cacciato M., Le Brun A. M. C., Dalla Vecchia C., 2014, MNRAS, 442, 2641
Vernon I., Goldstein M., Bower R., 2014, Statistical Science, 29, 81
Vikhlinin A., Kravtsov A., Forman W., Jones C., Markevitch M., Murray S. S., Van Speybroeck L., 2006, ApJ, 640, 691
Villaescusa-Navarro F., et al., 2021, ApJ, 915, 71
Vogelsberger M., et al., 2014, Monthly Notices of the Royal Astronomical Society, 444, 1518
Wendland H., 1995, Advances in Computational Mathematics, 4, 389
Wiersma R. P. C., Schaye J., Theuns T., Dalla Vecchia C., Tornatore L., 2009, MNRAS, 399, 574
Wright A. H., et al., 2017, MNRAS, 470, 283

Appendix

2.A Different apertures

Fig. 2.8 compares the SMF results for different choices of 3D apertures with radii of 30, 50 (our fiducial aperture) and 100 kpc. For each non-fiducial aperture we retrain the emulator on the SMFs obtained with the different aperture. The new emulator, based on a different aperture, is then evaluated at the fiducial subgrid parameter values. We do not refit the SMF for each aperture, because we wish to quantify the effect of the aperture size on the SMF predicted by a given simulation. The choice of aperture only has an impact at the largest stellar masses (see also Schaye et al. 2015). For our analysis this implies that the main effect of an increase in aperture would be a slight increase of the slope of the density dependence of the AGN accretion rate boost factor. However, for the fitted mass range this effect is relatively small. The effect of using a mass measurement method more similar to that used by observers may be larger (e.g. De Graaff et al., 2022), but such a comparison is not feasible at the resolution of our simulations.

2.B Posteriors for high- and low-resolution

The posteriors for low resolution are shown in Fig. 2.9. There is a degeneracy between the two parameters. Both parameters are sampled well within our chosen ranges. Even though the range for the heating temperature is much wider than for the other resolutions, we find that the best-fitting value is in the range where AGN feedback is well sampled, and does not suffer from catastrophic numerical overcooling (see ğ2.2.3).

The posteriors for the high-resolution simulation are shown in Fig. 2.10. Similar to the intermediate-resolution posteriors we find a best-fitting model within the chosen parameter ranges. The best-fitting value for $\beta_{\rm BH}$ is quite close to the edge, partly due to a degeneracy between $\beta_{\rm BH}$ and $\Delta v_{\rm SN}$. The high-resolution posteriors are more degenerate than for intermediate-resolution. This is likely due to the fact that we fit a much broader range of the SMF, making it more important to get the balance between stellar and AGN feedback right. The posteriors show that there are some significant degeneracies in how this problem can be solved.

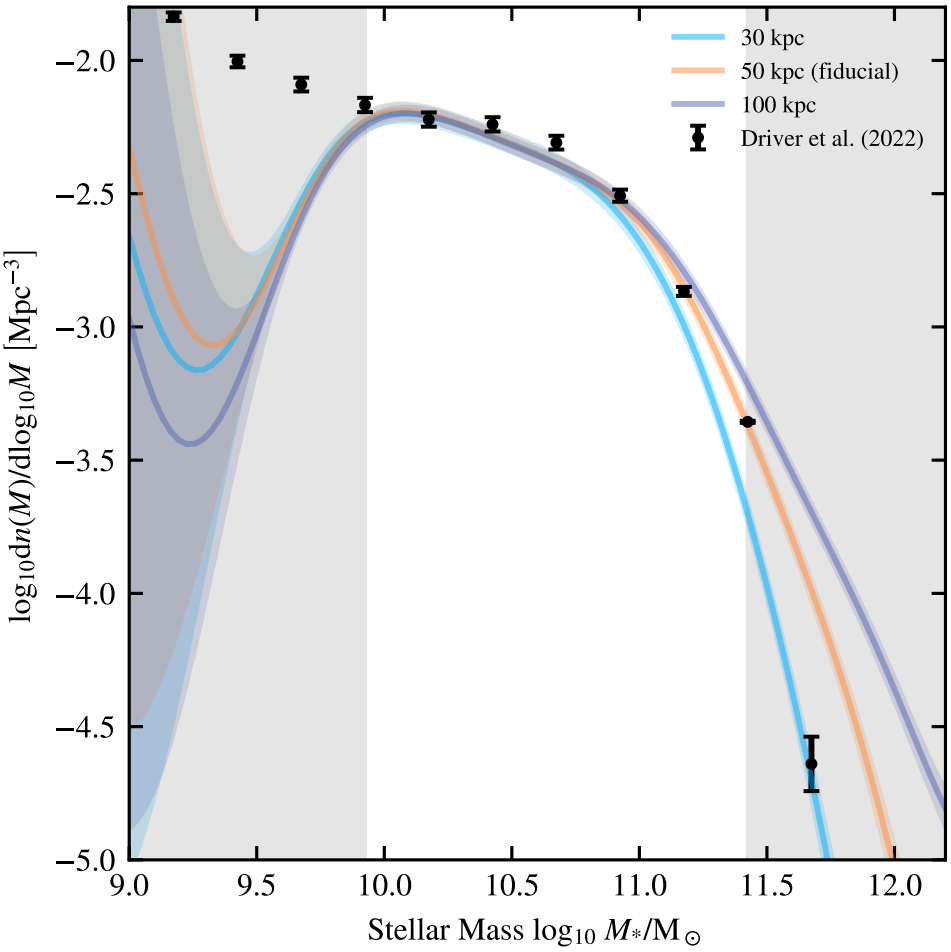


Figure 2.8: The effect on the SMF of choosing a different aperture when measuring stellar masses in the simulation. For each line we set up a new emulator based on the simulation results for the corresponding aperture. Each emulator is then used to predict the behaviour at the best-fitting parameter values for the fiducial 50 kpc aperture. Differences between the apertures start to occur above a stellar mass of 10^{11} $M_\odot$

Table 2.9: Subgrid parameter ranges for the Latin hypercube used to train the jet model emulators.

Parameter	Prior
$f_{\rm SN}$	[0.0,0.5]
$\Delta v_{\rm SN}$ [km s^{-1}]	[$10^{2.3}$,10^{3}]
$v_{\rm jet}$ [km s^{-1}]	[$10^{2.7}$,$10^{3.5}$]
$\beta_{\rm BH}$	[0.1,0.7]

Note that for both and high and low resolution we have fixed the biases to the values for for intermediate resolution, see §2.6.2.

2.C Parameter ranges for the AGN Jet model

The subgrid parameter ranges for the Latin hypercube that was used to train the emulators for the AGN jet model can be found in Table 2.9.

2.D swift-emulator: A Python package for emulation of simulated scaling relations - Roi kugel & Josh Borrow

The following section has been adapted from Journal of Open Source Software, 7(72), 4240, https://doi.org/10.21105/joss.04240

2.D.1 Summary

swift-emulator is a Python toolkit for using Gaussian processes machine learning to emulate scaling relations from cosmological simulations. swift-emulator focuses on implementing a clear, easy to use design and API to remove the barrier to entry for using emulator techniques. swift-emulator provides tools for every step: the design of the parameter sampling, the training of the Gaussian process model, and validating and analysing the trained emulators. By making these techniques easier to use, in particular in combination with the SWIFT code (Schaller et al., 2018b; Borrow & Borrisov, 2020), it will be possible use fitting methods (like MCMC) to calibrate and better understand theoretical simulation models.

2.D.2 Statement of need

One of the limits of doing cosmological (hydrodynamical) simulations is that any simulation is limited to only a single set of parameters, be these choices of cosmology, or the implemented physics (e.g., stellar feedback). These parameters

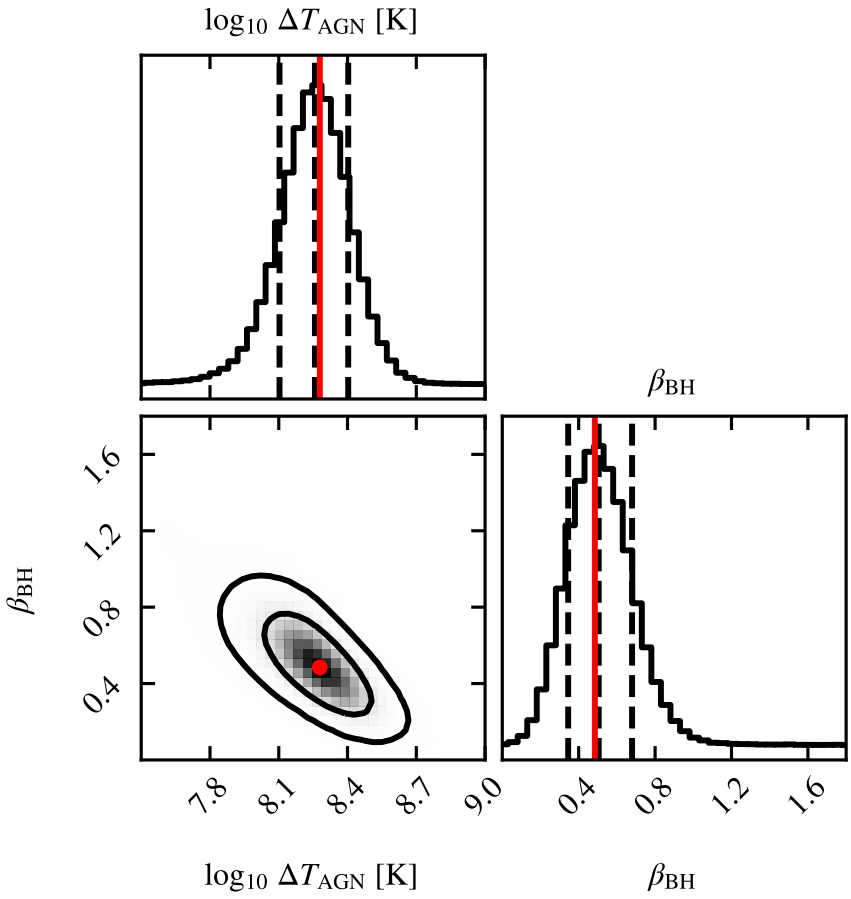


Figure 2.9: The posterior distributions of the model parameters resulting from fitting the emulator for low-resolution simulations to the observed SMF and cluster gas fractions. The parameters shown are the AGN feedback temperature jump $\Delta T_{\rm AGN}$ and the logarithmic slope of the density dependence of the black hole accretion rate boost factor, $\beta_{\rm BH}$. The two subgrid parameters are described in Section 2.2. The black contours show the 68 and 95 per cent confidence levels. The panels along the diagonal show the one dimensional probability density for each parameter. In these plots the three vertical lines indicate the 16th, 50th and 84th percentiles. The solid, red circle indicate the maximum likelihood values, which were used for the fiducial model. There is some degeneracy, but there is a clear single best-fitting solution.

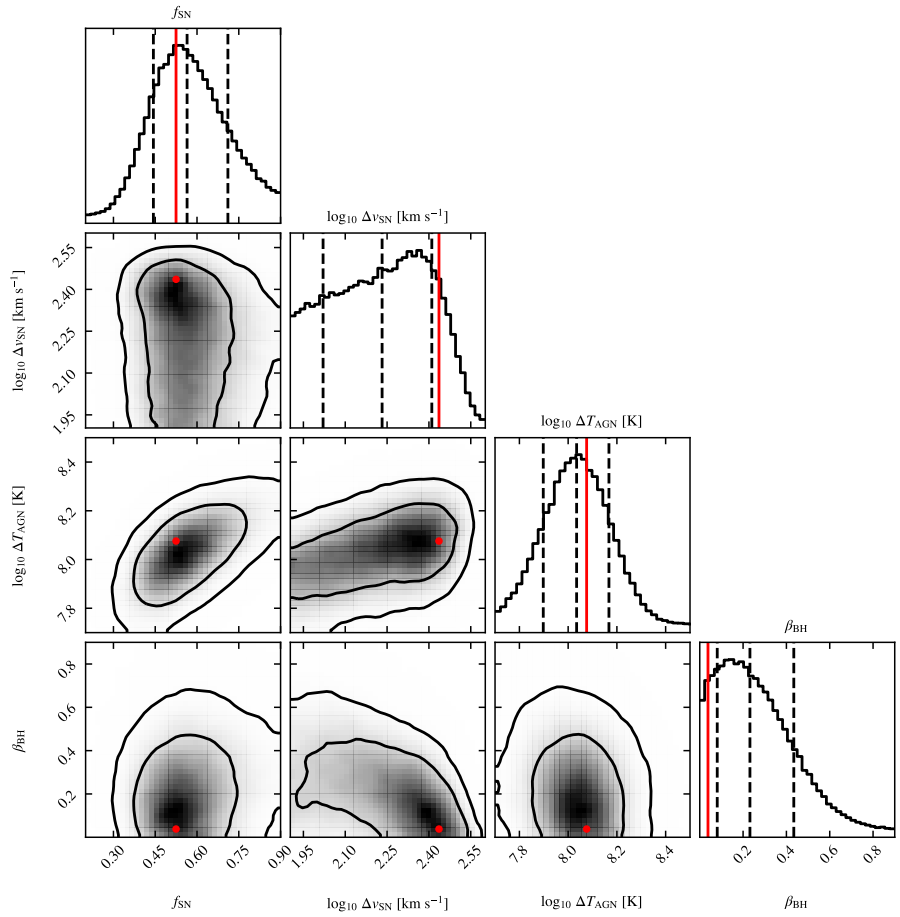


Figure 2.10: The posterior distributions of the model parameters resulting from fitting the emulator for high-resolution simulations to the observed SMF and cluster gas fractions. The parameters shown are the stellar feedback energy, f_{SN}, the stellar feedback kick velocity, Δv_{SN}, the AGN feedback temperature jump, ΔT_{AGN} and the logarithmic slope of the density dependence of the black hole accretion rate boost factor, β_{BH}. The four subgrid parameters are described in Section 2.2. The black contours show the 68 and 95 per cent confidence levels. The panels along the diagonal show the one dimensional probability density for each parameter. In these plots the three vertical lines indicate the 16th, 50th and 84th percentiles. The solid, red circles indicate the maximum likelihood values, which were used for the fiducial model. The results show some moderate degeneracies, but the individual parameters each have a clear peak close to the best-fitting values.

need to be tuned to calibrate against observational data. At odds with this, cosmological simulations are computationally expensive, with the cheapest viable runs costing thousands of CPU hours, and running up to tens of millions for the largest volumes at the highest resolutions. This makes the use of cosmological simulations in state-of-the-art fitting pipelines (e.g., MCMC), where tens of thousands to millions of evaluations of the model are required to explore the parameter space, computationally unfeasible. In order to get a statistical grip on the models of cosmology and galaxy formation, a better solution is needed.

This problem is a major limiting factor in "calibration" of the sub-resolution (subgrid) models that are often used. Works like Illustris (Vogelsberger et al., 2014), EAGLE (Crain et al., 2015), BAHAMAS (McCarthy et al., 2017), and Illustris-TNG (Pillepich et al., 2018) are able to "match" observed relations by eye, but a statistical ground for the chosen parameters is missing. This poses a significant problem for cosmology, where a deeper understanding of our subgrid models will be required to interpret results from upcoming surveys like LSST and EUCLID.

A solution here comes through the use of machine learning techniques. Training 'emulators' on a limited amount of simulations enables the evaluation of a fully continuous model based on changes in the underlying parameters. Instead of performing a new simulation for each required data-point, the emulator can predict the results a simulation would give for that set of parameters. This makes it feasible to use methods like MCMC based purely on simulation results.

2.D.3 Emulator Requirements

For emulation in hydro simulations we want to use Gaussian processes to emulate scaling relations in the following form:

$$GP(y, x, \vec{\theta}). \tag{2.28}$$

We want to emulate scaling relations between a dependent variable y, as a function of the independent variable x and the model parameters $\vec{\theta}$. For each simulation many of these individual scaling relations can be calculated, for example the sizes of galaxies relative to their stellar mass, or the mass fraction of gas in galaxy clusters as a function of their mass. The individual object properties used in scaling relations can be measured from each individual simulation using a tool like VELOCIraptor (Elahi et al., 2019a).

Between simulations, the underlying parameters $\vec{\theta}$ can change, for instance the energy injected by each supernovae. Using an emulator, we want to be able to see how many scaling relations change as a function of these parameters like the supernova strength.

Emulators do not make a distinction between the independent x and the model parameters $\vec{\theta}$. An emulator will model y as a function of the combined

vector $\vec{\theta}' = (x, \vec{\theta})$. Getting the training data in the correct format can pose a significant challenge.

In order to save computational time, it is important to have an efficient sampling of the parameter space represented by $\vec{\theta}$. It may be more efficient to search the parameter space in a transformed coordinate space, like logarithmic space, if the expected viable range is over several orders of magnitude.

Once the emulator is working it can be challenging to perform standard tests to validate it. Things like cross-checks or parameter sweeps have to be implemented by hand, making proper use of emulators more difficult.

2.D.4 Why swift-emulator?

Many packages exist for Gausian process emulation, like george (Ambikasaran et al., 2015b, this provides the basis for swift-emulator), gpytorch (Gardner et al., 2018) and GPy (GPy, 2012). Additionally, a package like pyDOE (Baudin et al., 2012) can be used to set up efficient parameter samplings. However, most of these packages operate close to theory, and create a significant barrier for entry.

With swift-emulator we aim to provide a single 'python' package that interfaces with available tools at a high level. Additionally we aim to streamline the processes by providing i/o tools for the SWIFT simulation code [@Schaller2018; @Borrow2020]. This is done in a modular fashion, giving the users the freedom to change any steps along the way. swift-emulator provides many methods that work out of the box, removing the barrier to entry, and aim at making emulator methods easy to use. The more wide-spread use of emulators will boost the potential of future simulation projects.

swift-emulator combines these tools to streamline the complete emulation process. There are tools for experimental design, such as producing Latin hypercubes or uniform samplings of n-dimensional spaces. For simulations performed with SWIFT, parameter files can be created and simulation outputs can be loaded in through helper methods in the library. The results can then be used to train an emulator that can make predictions for the scaling relations in the simulation. There are also methods to perform cross-checks to find the accuracy of the emulator. In addition, for investigating the impact of individual parameters on a given scaling relation, there is a simple method to do a parameter sweep implemented. Finally, there are tools for comparing the emulated relations with other data, from a simple χ^2 method to complex model discrepancy structures.

swift-emulator is currently being used for two of the flagship simulation projects using the SWIFT simulation code, ranging across five orders of magnitude in mass resolution. The package is being used to allow modern simulations to reproduce key observations with high accuracy.

Finally swift-emulator has many options to optimise the methods for specific emulation problems. While the focus so far has been on integration with SWIFT,

the underlying API is structured in a simple enough way that using the emulator with a different simulation code is easy. swift-emulator is currently being used for simulation projects outside of the SWIFT project for the calibration of post-processing models.

2

2.D.5 Acknowledgements

We acknowledge support from the SWIFT collaboration whilst developing this project, with notable involvement from Richard Bower, Ian Vernon, Joop Schaye, and Matthieu Schaller. This work is partly funded by Vici grant 639.043.409 from the Dutch Research Council (NWO). This work used the DiRAC@Durham facility managed by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). The equipment was funded by BEIS capital funding via STFC capital grants ST/K00042X/1, ST/P002293/1, ST/R002371/1 and ST/S002502/1, Durham University and STFC operations grant ST/R000832/1. DiRAC is part of the National e-Infrastructure.

2

3 The FLAMINGO Project: A comparison of galaxy cluster samples selected on mass, X-ray luminosity, Compton-Y parameter, or galaxy richness

Authors: **Roi Kugel** et al. Joop Schaye, Matthieu Schaller, Ian G. McCarthy,- Joey Braspenning, John C. Helly, Victor J. Forouhar Moreno, Robert J. McGibbon
Accepted for publishing in MNRAS

Abstract

Galaxy clusters provide an avenue to expand our knowledge of cosmology and galaxy evolution. Because it is difficult to accurately measure the total mass of a large number of individual clusters, cluster samples are typically selected using an observable proxy for mass. Selection effects are therefore a key problem in understanding galaxy cluster statistics. We make use of the $(2.8\ \mathrm{Gpc})^3$ FLAMINGO hydrodynamical simulation to investigate how selection based on X-ray luminosity, thermal Sunyaev-Zeldovich effect or galaxy richness influences the halo mass distribution. We define our selection cuts based on the median value of the observable at a fixed mass and compare the resulting samples to a mass-selected sample. We find that all samples are skewed towards lower mass haloes. For X-ray luminosity and richness cuts below a critical value, scatter dominates over the trend with mass and the median mass becomes biased increasingly low with respect to a mass-selected sample. At $z \leq 0.5$, observable cuts corresponding to median halo masses between $M_{500c} = 10^{14}$ and $10^{15}\ \mathrm{M}_\odot$ give nearly unbiased median masses for all selection methods, but X-ray selection results in biased medians for higher masses. For cuts corresponding to median masses $< 10^{14}$ at $z \leq 0.5$ and for all masses at $z \geq 1$, only Compton-Y selection yields nearly unbiased median masses. Importantly, even when the median mass is unbiased, the scatter is not because for each selection the sample is skewed towards lower masses than a mass-selected sample. Each selection leads to a different bias in secondary quantities like cool-core fraction, temperature and gas fraction.

3.1 Introduction

Galaxy clusters are the largest virialized structures in the universe, and are found at the intersections of the filamentary network of the cosmic web. Following hierarchical structure formation, clusters are the last objects to form. Both the number density of clusters as a function of cluster mass, i.e. the halo mass function

(HMF), and the properties of individual clusters are sensitive to the underlying cosmological model (for a review see Allen et al., 2011).

The current standard model of cosmology involves a spatially flat universe dominated by dark energy and cold dark matter, and is denoted as ΛCDM. Recent weak lensing and distance ladder measurements have exposed tensions between the ΛCDM parameters recovered by measurements of the cosmic microwave background (e.g. Planck Collaboration et al., 2020) and observations of the late-time universe (e.g. Heymans et al., 2021; Abbott et al., 2022; Riess et al., 2022; Miyatake et al., 2023). Galaxy clusters are an independent probe that can help further investigate these tensions.

The cluster cosmology probe that is used the most is cluster counts, which are parameterised via the HMF. The HMF gives the abundance of clusters as a function of their total mass within some 3D aperture, which is generally not directly observable. Instead, measurements are limited to indirect probes of the total (3D) mass of the cluster, and selection effects have to be accounted for. Clusters are selected based on their Sunyaev-Zeldovich signal (SZ) (e.g. Planck Collaboration et al., 2016a; Bocquet et al., 2019; Bleem et al., 2024), X-ray luminosity (e.g. Pacaud et al., 2018; Chiu et al., 2023; Ghirardini et al., 2024), galaxy richness (e.g. Rykoff et al., 2014, 2016; Black & Evrard, 2022) combined with weak lensing signal (e.g. Costanzi et al., 2019). Future data releases of eRosita, and upcoming weak lensing missions like Euclid (Artis et al., 2022) will lead to an enormous increase in the number of detected clusters. With increased statistics the cosmology constraints will become much tighter.

Systematic differences between X-ray- and SZ-selected samples are well documented observationally. Lovisari et al. (2017) reports finding an excess of disturbed clusters in SZ selected samples with respect to X-ray-selected samples. Additionally, Andrade-Santos et al. (2017) and Rossetti et al. (2017) report a larger fraction of cool-core objects for X-ray selection compared to SZ selection. Chon & Böhringer (2017) argue that many of the differences in results for different samples are due to the difference between flux- and volume-limited selections rather than the specific selection observable used. There are quite a few comparisons of X-ray vs richness selected samples. In general, good agreement is found for the mass-luminosity relation, luminosity-richness relation, disturbed fraction and merger fraction when comparing X-ray and richness selected samples (e.g. Ramos-Ceja et al., 2022; Upsdell et al., 2023), but slight differences might exist in the X-ray luminosity-temperature relation (Giles et al., 2022). Additionally, Ota et al. (2023) find that clusters selected on having a high galaxy richness have a smaller fraction of relaxed clusters compared to X-ray selected samples. In general, galaxy richness selected samples contain a much larger number of clusters than X-ray selected samples. Grandis et al. (2021) find that this might originate from the fact that the contamination in richness selected samples increases towards lower values of richness. In a comparison between weak lensing

shear selected sources and X-ray selected sources by Willis et al. (2021), a large fraction of the sources is not matched between the catalogues. This is partially due to projection effects boosting the shear, but also because extended high flux sources were missed due to the morphological selection criteria and the XMM beam. Marini et al. (2024) show using mock observations that eROSITA is unable to find all group size objects, with a bias towards detecting objects with a high relative gas fraction. These differences show that every selection has a unique selection function.

Understanding the influence of selection effects on derived cluster properties is important beyond cluster cosmology. For example, scaling relations for clusters, in particular their baryon and gas content, provide constraints on how baryons impact the matter power spectrum (e.g. Chisari et al., 2019; Aricò et al., 2021; Giri & Schneider, 2021; Debackere et al., 2021; Salcido et al., 2023). Current measurements of the gas fraction in clusters (as collated by Kugel et al., 2023) indicate that selection effects start to dominate for haloes around the group mass of $M_{500c} \lesssim 10^{13.5}\,\mathrm{M}_{\odot}$.

The careful modelling of selection functions is one of the main ingredients of cosmological inference with cluster counts. As shown by Mantz (2019) a good grip on both the selection criteria and the mass-observable relation is necessary. In order to do unbiased cosmology inference, proper modelling of the observable relations and the effects of the selection procedures is key (Angulo et al., 2012). Power law relations with scatter are commonly assumed to relate observables to masses (e.g. Rozo et al., 2014; Evrard et al., 2014; Planck Collaboration et al., 2016a; Pacaud et al., 2018; Grandis et al., 2020; Chaubal et al., 2022). Especially for observations probing lower masses, these assumptions might break down, and lead to biased results. Recent cosmological inferences often combine X-ray or SZ selection with scaling relations based on lensing or richness (Bocquet et al., 2024; Clerc et al., 2024; Ghirardini et al., 2024), leading to additional complexity when modelling the selection function. A solution is to predict quantities that are directly observable. One candidate is the aperture lensing mass, as discussed by Debackere et al. (2022a,b). Similarly, Andreon et al. (2024) introduce the X-ray surface brightness within 300 kpc as a promising candidate that reduces observational biases when compared with X-ray, SZ or galaxy richness selected samples.

Cosmological constraints are typically inferred by comparing observed cluster counts to results based on (emulators of) the HMF of dark matter only simulations (e.g. Tinker et al., 2008; Bocquet et al., 2020). However, baryonic physics can lead to biases (e.g. Debackere et al., 2021). Additionally, dark matter only simulations cannot self-consistently model the gas that is needed to predict X-ray and SZ observables. As hydrodynamical simulations are computationally more expensive than dark matter only simulations, some of the state-of-the-art simulations like EAGLE (Schaye et al., 2015), Horizon-AGN (Kaviraj et al., 2017),

IllustrisTNG (Pillepich et al., 2018) and Simba (Davé et al., 2019) do not sample volumes sufficiently large to contain a representative sample of clusters. Simulations like BAHAMAS (McCarthy et al., 2017), MilleniumTNG (Pakmor et al., 2023) have volumes large enough to investigate typical clusters at low redshift, but for converged statistics for the halo mass distributions even larger volumes are needed. While the lowest-resolution simulations of the Magneticum suite (Hirschmann et al., 2014) have large volumes, so far only BAHAMAS uses subgrid models that have been calibrated to reproduce the gas fractions of clusters. Cosmological hydrodynamics simulations can be extended to the cluster range by making use of zoom-in simulations (e.g. Barnes et al., 2017; Bahé et al., 2017; Hahn et al., 2017; Cui et al., 2018; Pellissier et al., 2023; Nelson et al., 2023). While zooms enable simulating samples of massive clusters without the need to model very large volumes, they require selecting a sample from a large volume dark matter only simulation. Because a volume-limited sampled cannot be constructed from zooms, they cannot yield an unbiased study of selection effects.

For this work we make use of the FLAMINGO simulations (Schaye et al., 2023; Kugel et al., 2023). FLAMINGO is a suite of large-volume cosmological hydrodynamical simulations in box-sizes with side-lengths of 1.0 and 2.8 Gpc. At a resolution of $m_{\rm gas} = 1.07 \times 10^9$ M$_\odot$, using 5040^3 gas particles, the (2.8 Gpc)^3 FLAMINGO box is the largest cosmological hydrodynamics simulation evolved to $z = 0$. Additionally, FLAMINGO includes models that vary the resolution, cosmology, and feedback strength in boxes of (1.0 Gpc)^3. The cluster gas fractions and stellar mass function of the fiducial and feedback variations are calibrated to shifted observations.

The FLAMINGO simulations have been shown to be in good agreement with observations of hot gas in groups and clusters (Schaye et al., 2023; Braspenning et al., 2023). In particular, Braspenning et al. (2023) find that the X-ray luminosity, temperature and thermal SZ scaling relations are in good agreement with the data at all redshifts. The thermodynamic profiles also agree well with the observations, although the metallicities are too high in cluster cores. Braspenning et al. (2023) also find that the cool-core fractions are difficult to compare with observations, as they are very dependent on the measure used, and typically based on the properties of the gas at radii near or below our resolution limit, but that they are in agreement with other simulation projects. The cool-core fractions vary more strongly between the FLAMINGO feedback variations than is the case for the scaling relations and the outer thermodynamic profiles.

The simulation's very large volume, containing 461 (4100) clusters of mass $M_{\rm 500c}$[1] $> 10^{15}$ M$_\odot$(5×10^{14} M$_\odot$) at $z = 0$, the agreement with cluster observations, as well as the availability of convergence tests and model variations, make

[1] $M_{\rm 500c}$ is the mass enclosed by a sphere with radius $R_{\rm 500c}$, which is defined as the radius of a sphere centered on a halo within which the average density is 500 times the critical density.

FLAMINGO ideal for investigating the impact of selection effects on cluster counts.

We will compare selections based on X-ray luminosity, integrated thermal SZ effect and galaxy richness. We will contrast these selections with mass-selected samples for different redshifts. We will perform all these selections on theoretical quantities, without applying any other observational biases, projection effects, or noise. Our results are thus for a best-case scenario as selection effects are likely to become stronger when the sample selection is forward modelled using virtual observations. We choose to limit this work to theoretical quantities to increase the interpretability and because further steps towards forward modelling require choices that are survey specific, which will make the results less general. This is also why we leave an investigation of lensing masses to future work, as lensing only works in projection and requires the specification of a survey-specific redshift distribution of lenses galaxies. In future work we plan to model selection effects for specific observables and surveys by creating virtual observations based on FLAMINGO's lightcone output.

This paper is structured as follows: In Section 3.2 we discuss the FLAMINGO simulations, the quantities we select on and our definition of the sample mass bias, in Section 3.3 we present our results and we conclude and summarise our findings in Section 3.4.

3.2 Methods

In this section we describe the methods and data used. We discuss the FLAMINGO simulations and how we obtain halo catalogues in Section 3.2.1. The definitions used for the different quantities are described in Section 3.2.2 and the metrics with which we quantify the quality of the selections are described in Section 3.2.3.

3.2.1 FLAMINGO

This work makes use of the FLAMINGO simulations, described in detail by Schaye et al. (2023). FLAMINGO (Full-hydro large-scale structure simulations with all-sky mapping for the interpretation of next generation observations) is a suite of cosmological hydrodynamics simulations in large volumes with variations in baryonic feedback, cosmology, box size and resolution. In this work we make use of the simulations run at intermediate resolution ($m_{\rm gas} = 1.07\times10^9$ M$_\odot$) in a volume of $(2.8\,{\rm Gpc})^3$ which consist of 2×5040^3 gas and dark matter particles, and 2800^3 neutrino particles. The full output consists of 79 snapshots, of which we will use the snapshots at $z = [0, 0.3, 0.5, 1.0, 2.0]$.

The FLAMINGO simulations use the open source code Swift (Schaller et al., 2024). The simulations make use of the SPHENIX SPH scheme (Borrow et al.,

2022) with a Wendland (1995) C^2 kernel. Neutrinos are simulated using the δf method (Elbers et al., 2021). The ICs are generated using a modified version of Monofonic (Hahn et al., 2021; Elbers et al., 2022). The simulations use the '3x2pt + all external constraints' cosmology from the dark energy survey year 3 results of Abbott et al. (2022) ($\Omega_m = 0.306$, $\Omega_b = 0.0486$, $\sigma_8 = 0.807$, $H_0 = 68.1$, $n_s = 0.967$). Simulations with different cosmologies are available but not used in this work.

3

FLAMINGO includes subgrid models for element-by-element radiative cooling and heating (Ploeckinger & Schaye, 2020), star formation (Schaye & Dalla Vecchia, 2008), stellar mass loss (Wiersma et al., 2009; Schaye et al., 2015), feedback energy from supernova (Dalla Vecchia & Schaye, 2008; Chaikin et al., 2022b,a), seeding and growth of black holes, and feedback from active galactic nuclei (Springel et al., 2005; Booth & Schaye, 2009; Bahé et al., 2022). The fiducial models use a thermal model for AGN (Booth & Schaye, 2009), but we have two variations that use kinetic jets (Huško et al., 2022) (for a detailed description see Schaye et al., 2023). As for BAHAMAS, the important simulation parameters are set to match the observed $z = 0$ galaxy stellar mass function (Driver et al., 2022) and a compilation of data of gas fractions in clusters (Kugel et al., 2023). Unique to the FLAMINGO simulations is the method used to calibrate the subgrid physics. For FLAMINGO these parameters are fit to the observations by making use of emulators, as described by Kugel et al. (2023). This procedure is also used to constrain a set of feedback variations that skirt error bars on the calibration data. The variations are denoted by the change in the observations they are matched to. "fgas$\pm N\sigma$" denotes runs where the gas fraction is shifted up or down by $N\sigma$, "M*$-\sigma$" denotes runs where the stellar mass function is shifted to lower masses by 1σ and "Jet" denotes runs where AGN feedback is implemented in the form of kinetic jets instead of thermally-driven winds.

We identify cosmic structure using a recently updated version (see Forouhar Moreno et al. in prep) of the Hierarchical Bound Tracing algorithm (HBT+, Han et al., 2018), which leverages hierarchical structure formation to identify substructures more robustly than traditional halo finders. In short, it identifies structures as they form in isolation, by subjecting particles within spatial friends-of-friends (FOF) groups to an iterative unbinding procedure. The particles associated to these self-bound objects are tracked across outputs to provide a set of candidate substructures at later times. This allows the identification of satellites, as the particle memberships are retained once they have been accreted by the FOF of a more massive halo. Finally, each candidate substructure is subject to additional self-boundness and phase-space checks to decide whether it is still resolved, or if it has merged or disrupted. The HBT+ catalogue is further pro-

cessed by the Spherical Overdensity and Aperture Processor (SOAP[2]; McGibbon et al. in prep) , which computes a large selection of halo properties in a range of apertures. For this work we use properties inside R_{500c}, which is defined as the radius within which the enclosed density is 500 times the critical density. R_{500c} defines the mass M_{500c} which is defined as the mass within R_{500c}. Because observational studies of clusters focus on centrals, we consider only central galaxies, as identified by VR.

3.2.2 Observables used for selection

The X-ray luminosity within R_{500c} is defined as the intrinsic luminosity within the Rosat 0.5-2.0 keV broad band in the observer frame. This excludes star forming gas and gas at low temperatures ($T < 10^5$ K). We do not attempt to exclude satellites and sum over all particles within R_{500c}. The X-ray luminosity of each particle is computed by interpolating in redshift, density, temperature and individual element abundances, based on output from the photo-ionisation spectral synthesis code Cloudy (Ferland et al., 2017). A detailed description is given by Braspenning et al. (2023). Because the luminosities are measured in the observer frame, different parts of the rest-frame X-ray spectra will fall in the band at different redshifts.

We measure the thermal SZ Compton-Y in an aperture of $5 \times R_{500c}$ as done in Planck Collaboration et al. (2016b), but in Appendix 3.B we show some of the results also for an aperture of R_{500c}. Compton-Y is computed by summing over the Compton-Y contribution from each individual gas particle, y_i, which is stored in the snapshots. The contribution of the individual particles is computed at run-time following

$$y_i = \frac{\sigma_T}{m_e c^2} n_{e,i} k_B T_{e,i} \frac{m_i}{\rho_i}, \tag{3.1}$$

where σ_T is the Thomson cross section, m_e is the electron mass, c is the speed of light, k_B is the Boltzmann constant, $n_{e,i}$ is the electron number density, $T_{e,i}$ is the electron temperature m_i is the mass and ρ_i is the density of the particle with index i. The electron number density and temperature are obtained from the cooling tables. Selections based on the integrated Compton-Y are referred to as SZ-selections.

For both the X-ray luminosity and Compton-Y signal we exclude particles that in which AGN feedback energy has recently been directly deposited. This can affect the X-ray luminosity, particularly for outlier haloes with a high luminosity, but has a negligible effect on Compton-Y. AGN feedback in the fiducial FLAMINGO simulations is implemented thermally, heating a single particle to a high temperature. Particles that are heated tend to be close to the core of the halo and can have very high densities. This can lead to single particles having an

[2] https://github.com/SWIFTSIM/SOAP

unrealistically large contribution to the total X-ray luminosity and Compton-Y signal of the halo, potentially dominating over the rest of the halo, which would be unphysical. To avoid this we ignore the contribution to the X-ray luminosity and Compton-Y signal of particles that have been heated in the last 15 Myr and that have a temperature in the range

$$10^{-1}\Delta T_{\rm AGN} \leq T_i \leq 10^{0.3}\Delta T_{\rm AGN}, \tag{3.2}$$

where T_i is the temperature of the particle and $\Delta T_{\rm AGN}$ is the change in temperature when a particles is heated by a black hole, which has a value of $10^{7.78}$ K for the fiducial FLAMINGO model.

We define richness by counting the number of satellite galaxies above a mass threshold. Richness is defined as

$$\lambda = N_{\rm sats}(M_* > 10^{10.046}\ {\rm M}_\odot, r < R_{\rm 200c}) + 1, \tag{3.3}$$

where M_* is the stellar mass within a 50 proper kpc spherical aperture and r is the spherical radius from the centre of the cluster. These mass and radial limits were chosen to be similar to the cuts used for Redmapper (Rykoff et al., 2014). The mass limit is obtained from the fact that Redmapper uses a cut of $0.2L_*$, where L_* is the luminosity at the knee of a Schecter fit to the luminosity function. We convert this to $0.2M_*$ and use the mass at the knee from the stellar mass function of Driver et al. (2022), which FLAMINGO is calibrated to match. The Redmapper radial cut is a function of richness, and is optimised as part of the richness finding process. We instead opt for $R_{\rm 200c}$. This gives us the scaling of the radius with halo mass that is implicit in the Redmapper radial cut, but with a pre-defined radius for each halo. We pick $R_{\rm 200c}$ over $R_{\rm 500c}$ as the satellites in the interior of the clusters are more likely to be affected by resolution-dependent tidal disruption, and a larger radius leads to better convergence. For the values of richness that we recover $R_{\rm 200c}$ is usually a factor of a few larger than the scale cut used for Redmapper. As we do not fully forward model Redmapper, we choose to use a larger 3D volume instead of a cylinder as this leads to a more well defined sample. Qualitatively the differences between a 3D sphere and a 2D projection will be small without forward modelling. The FLAMINGO simulations are calibrated to reproduce the galaxy mass function down to a stellar mass of $10^{9.9}$ M$_\odot$. We wish to ensure that, on average, haloes down to $M_{\rm 500c} = 10^{13}$ M$_\odot$ still have more than one satellite above this mass, as a selection based on a richness of one returns all haloes. Note that Redmapper itself makes a probabilistic prediction for the number of satellites, and is hence not as affected by discreteness effects at low galaxy richness, though it will still be affected by small-number statistics for individual sources.

3.2.3 Sample Selection

A selection based on observable A is defined as the set of haloes that have $A > A_C$, where A_C is the selection limit. In order to compare selections based on different observables, we find the corresponding selection limits by taking the median of each observable at a fixed halo mass. In the case of an ideal scaling relation without scatter, such a selection would be equivalent to a mass selection. To compute this median value, we select haloes in a mass bin of 0.1 dex centred around the chosen mass limit. We then compute the median X-ray luminosity, thermal SZ signal or richness for these haloes. The cut, A_C, is defined as

$$A_C(M_C) = \text{median}\left[A\left(10^{-0.05}M_C < M_{500c} < 10^{0.05}M_C\right)\right], \tag{3.4}$$

where M_C is the target mass cut. By comparing sample selections $A > A_C(M_C)$ using the same target mass cut M_C, we can investigate how selections based on different observables deviate from the ideal case where A is exactly proportional to M_{500c} with no scatter.

Cluster count studies relate the counts in a sample to the HMF. To investigate how much the sample deviates from a mass-selected sample, we define the sample mass bias factor

$$b_{M_{500c}}(a, M_C) = \frac{\text{median}\,(M_{500c}|A > a)}{\text{median}\,(M_{500c}|M_{500c} > M_C)} - 1. \tag{3.5}$$

Hence, $b_{M_{500c}}(a, M_C)$ indicates the bias in the median M_{500c} of the sample $A > a$ compared to a sample for which $M_{500c} > M_C$. A bias of zero indicates an unbiased median mass. A negative (positive) bias indicates the median mass in the sample is lower (higher) than for the mass-selected sample. The bias factors for percentiles other than the median are defined analogously. Note that for the special case $a = A_C$ the bias is only a function of M_C. The bias has to be calculated separately for each redshift. By defining the sample mass bias in this way, we can quantify how much a cluster sample selected by a simple cut based on the value of an observable is influenced by lower-mass haloes that up-scatter into the sample. By investigating the bias in percentiles lower than the median, we will further quantify the level of contamination by lower mass haloes in the sample. We choose the 5th percentile as it strikes a good balance between probing the lower mass tail of each sample without being too influenced by small number statistics. The qualitative results are insensitive to the percentile picked, but in general a lower percentile that is further from the median leads to a larger value for the sample mass bias.

3.3 Results

In this section we compare the properties of cluster samples obtained with different selection cuts $A > A_C$, where A is mass M_{500c}, X-ray luminosity $L_{500c,0.5\text{-}2keV}$,

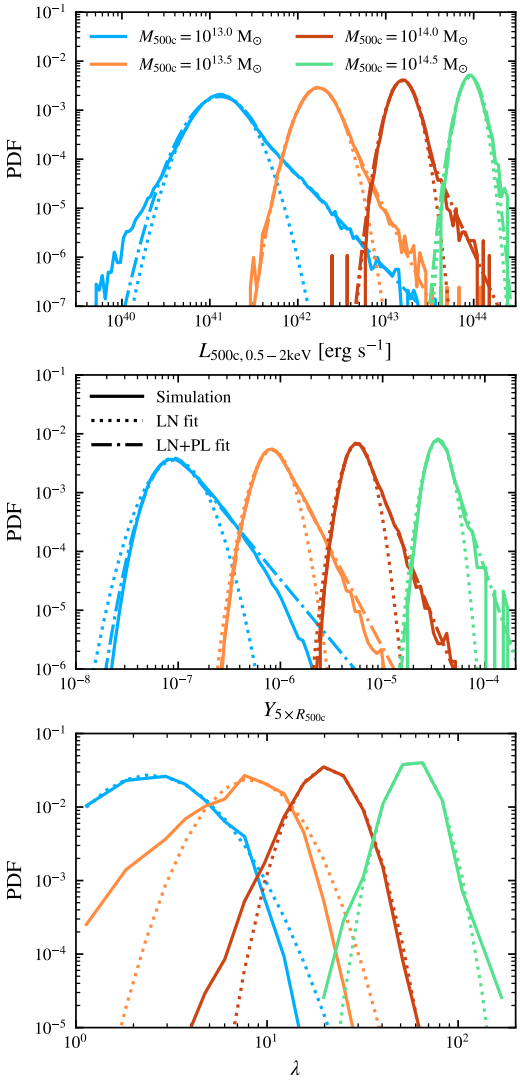


Figure 3.1: The distribution of X-ray luminosity (top), integrated thermal SZ Compton-Y (middle) and galaxy richness (bottom) at fixed mass at $z = 0.3$. The different colours indicate different mass bins of width ± 0.05 dex, around the central value. The dotted lines show the best-fitting lognormal function and, for X-ray and thermal SZ, the dot-dashed lines show the best-fitting lognormal plus power-law distributions (Eq. 3.6). For lower masses the lognormal distributions shift to smaller values and become narrower, while the power-law tails start at lower values and become shallower. Assuming lognormal distributions would underestimate the amount of upscatter from lower mass objects for a given cut on the value of the chosen mass proxy.

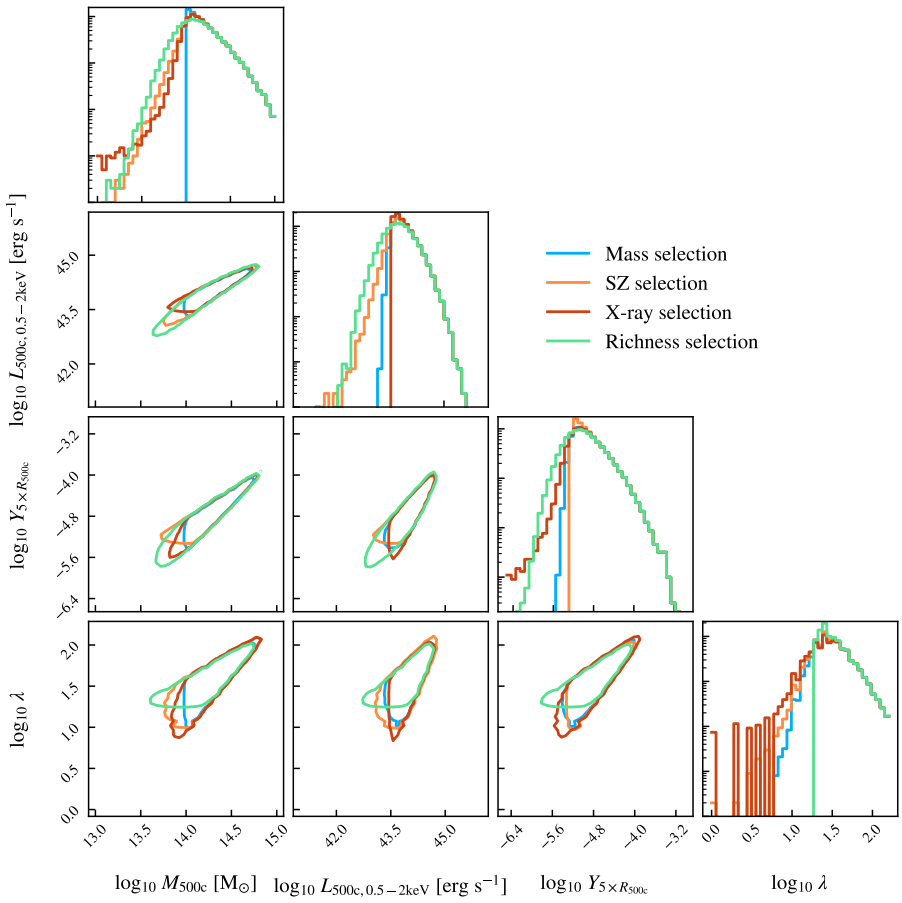


Figure 3.2: Corner plot showing the distribution of four different cluster properties A: M_{500c}, X-ray luminosity, Compton-Y and galaxy richness, for different selections of haloes in the 2.8 Gpc FLAMINGO fiducial volume at $z = 0.3$. Different colours correspond to samples selected based on different quantities A, as indicated in the legend. Each sample is defined to have $A > A_C$ where $A_C = \text{median}[A(M_{500c} = 10^{14}\ M_\odot)]$. The panels along the diagonal show histograms, while the off diagonal panels show two-dimensional distributions with each contour containing 95 per cent of the haloes in the sample. For each sample, the value of A_C corresponds to the sharp cutoff in the histogram shown in the top panel of the column with A plotted along the x-axis. The different samples converge for $A \gg A_C$ but there are differences for $A \lesssim A_C$.

Table 3.1: Values for the fits to Eq. 3.6 at $z = 0.3$. Top four rows are for X-ray luminosity, middle four for integrated Compton-Y and bottom for for galaxy richness. Note that for richness we only fit a regular lognormal, so we do not the include the power law parameters. Fits for other redshifts can be found in Appendix 3.A.

a	$M_{500c}[\mathrm{M}_\odot]$	A	μ	σ	$\log_{10} a_t$	α
X-ray	$10^{13.0}\,\mathrm{M}_\odot$	1.888×10^{-3}	41.1	0.35	41.6	1.90
X-ray	$10^{13.5}\,\mathrm{M}_\odot$	2.875×10^{-3}	42.2	0.23	42.6	3.67
X-ray	$10^{14.0}\,\mathrm{M}_\odot$	4.027×10^{-3}	43.2	0.17	43.5	3.96
X-ray	$10^{14.5}\,\mathrm{M}_\odot$	4.779×10^{-3}	44.0	0.14	44.9	6.50
SZ	$10^{13.0}\,\mathrm{M}_\odot$	3.882×10^{-3}	-7.05	0.23	-6.92	2.09
SZ	$10^{13.5}\,\mathrm{M}_\odot$	5.464×10^{-3}	-6.09	0.17	-5.98	3.27
SZ	$10^{14.0}\,\mathrm{M}_\odot$	7.010×10^{-3}	-5.25	0.13	-5.17	4.25
SZ	$10^{14.5}\,\mathrm{M}_\odot$	8.185×10^{-3}	-4.46	0.12	-4.31	5.04
λ	$10^{13.0}\,\mathrm{M}_\odot$	2.729×10^{-2}	0.39	0.34	-	-
λ	$10^{13.5}\,\mathrm{M}_\odot$	2.391×10^{-2}	0.91	0.24	-	-
λ	$10^{14.0}\,\mathrm{M}_\odot$	3.511×10^{-2}	1.31	0.17	-	-
λ	$10^{14.5}\,\mathrm{M}_\odot$	4.534×10^{-2}	1.77	0.13	-	-

thermal SZ signal $Y_{5\times R_{500c}}$, or galaxy richness λ. In Section 3.3.1 we show and fit the distributions of each of the selection observables at fixed mass. We describe the general correlations between the differently selected samples for a target mass cut of $M_\mathrm{C} = 10^{14}\mathrm{M}_\odot$ in Section 3.3.2. We show different percentiles of the mass distribution as a function of A_C in Section 3.3.3. We investigate the shift across redshift for selections based on cuts number density and observables in Sections 3.3.4 and 3.3.5, respectively. In Section 3.3.6 we investigate how the sample bias depends on mass and redshift. We finish by investigating how the different selections impact secondary cluster properties in Section 3.3.8.

3.3.1 Scatter at fixed mass

Before comparing samples defined by cuts in different observables, we will investigate the distribution of the observable mass proxies at fixed halo mass. Fig. 3.1 shows the scatter in X-ray luminosity (top panel), SZ Compton-Y (middle panel) and galaxy richness (bottom panel) in four different mass bins at $z = 0.3$. The mass bins are 0.1 dex wide, ±0.05 dex around the centre, and are centred on $\log_{10} M_{500c}/\mathrm{M}_\odot = 13.0$, 13.5, 14.0, and 14.5.

The distributions shift towards larger values for higher masses. Near their peaks, the distributions are well described by lognormal fits (dotted curves). However, the X-ray luminosity and Compton-Y distributions have tails towards higher values that deviate from the lognormal fits, skewing the distributions towards larger values. These distributions are well fit by lognormal plus power-law

functions (dot-dashed curves) parameterised as

$$N_{\rm haloes}(a) = \begin{cases} A\exp\left(-\frac{(\log_{10}(a)-\mu)^2}{\sigma^2}\right) & a \leq a_t, \\ Ba^{-\alpha} & a > a_t, \end{cases} \tag{3.6}$$

where,

$$B = \frac{A\exp\left(-\frac{(\log_{10} a_t-\mu)^2}{\sigma^2}\right)}{10^{-\alpha\log_{10} a_t}}. \tag{3.7}$$

The best-fitting values of the free parameters A, μ, σ, $\log_{10} a_t$ and α, which we obtained using least squares statistics, where each bin was weighted by $1/\sqrt{N}$, can be found in Table 3.1, and the result for other redshifts can be found in Appendix 3.A. The general trends described below also apply to the other redshifts.

For lower mass bins the lognormal parts become narrower, the power-law tails start closer to the peak and the slope become shallower. As a result, samples defined by a cut $A_{\rm C}$ will suffer from a slight increase in upscatter from low-mass bins and this upscatter will be underestimated if the distributions are assumed to be lognormal, which is the assumption conventionally adopted in the literature. X-ray is slightly more skewed, and Compton-Y is significantly more skewed than what was found for the stellar and gas mass by Farahi et al. (2018). However, for Compton-Y the deviations from a lognormal depend on the size of the aperture, which in this work we take to be $5R_{500c}$ for Compton-Y as appropriate for the Planck satellite. In Appendix 3.B we demonstrate that the deviations largely disappear when using a smaller aperture of R_{500c}, which suggests that the deviations visible in Fig. 3.1 are due to projection of/blending with nearby structures.

For richness we do not attempt to fit a lognormal plus power-law, since this shape is not clearly seen in the distributions. For the highest mass bins the shape is lognormal, and for the lower mass bins there is a tail extending towards lower values of richness.

For all three observables we find an increase in the lognormal scatter towards lower masses. We leave an investigation of the physical origin of the scatter in the different observables for future work.

3.3.2 Correlations between cluster properties

To better understand how different selections will relate to the different observables, we investigate the distributions of, and correlations between the observables we select on. In Fig. 3.2 we show a corner plot of the distribution of our selection quantities at $z = 0.3$. We pick an intermediate redshift, but note that the qualitative picture is similar at $z = 0$ and 0.5. The panels along the diagonal show histograms of the individual quantities. The off diagonal panels show the 95th percentiles for each combination of quantities. The light blue lines show the

mass selected distribution for a lower limit on M_{500c} of $M_C = 10^{14}\,M_\odot$. Each other colour shows the result for a different sample $A > A_C$, i.e. a selection based on the median value of the observable A indicated in the legend at $A(M_{500c} = M_C)$.

The light blue contours in the first column show that all observables are tightly correlated with halo mass for masses above the mass cut $M_{500c} = 10^{14}\,M_\odot$. For $M > M_C$ the differences between the different samples are small. Below this mass the distributions diverge. Richness shows the largest spread, and a cut of $M_{500c} = 10^{14}\,M_\odot$ is still high enough for it to not suffer from small number statistics.

The mass distributions for selections based on the other quantities can be seen in the topmost diagonal panel. At the selection limit the number of objects selected based on the quantity shown along the x-axis of the histogram drops to zero. The richness selection includes the largest number of haloes below the target mass M_C and starts to become incomplete, with respect to a mass selection, at masses below ≈ 0.2 dex above M_C. X-ray and Compton-Y selections are comparable to each other in terms of completeness at the target mass, and include less contamination from haloes with $M_{500c} < M_C$ than the sample selected on richness. At this redshift, X-ray selection yields the lowest number of haloes with mass smaller than the target mass.

3.3.3 Characteristic mass as a function of the cut in observable space

In addition to the complete distributions shown in Fig 3.2, it is interesting to look at each of the scaling relations between observable and mass that are used for the selection. The solid line in each panel of Fig. 3.3 shows the median M_{500c} for a sample defined by $A > A_C$ with A_C plotted along the x-axis at $z = 0.3$. Different panels show different choices for A. From top to bottom, the three panels show X-ray, SZ and richness selection. We also show the 5th and 95th percentiles of the sample, and for reference we indicate the median values at fixed mass, unlike the black lines which show the mass of the full sample with $A > A_C$, of the selection quantities at three fixed values of M_C using vertical dotted lines, with a circle to mark the mass the line corresponds to.

In all panels, the median lines cross each vertical line at a mass that is slightly higher than the mass that the vertical line is based on, indicated by the circle. Since the vertical lines and circles indicate the value of A for a sample with fixed mass M_C, while the black lines show the median based on the sample with $A > A(M_C)$, this is expected. The difference is not very large, due to the exponential nature of the high-mass end of the halo mass function, every sample will be dominated by its lowest mass haloes. There is a very slight trend where for richness the crossing point is closest to the fixed mass M_C compared with X-ray and SZ selections. As seen in Fig, 3.2, richness starts becoming incomplete at a

higher mass than the other selection methods, which will make the median mass in such a sample lower.

Except for very low X-ray luminosity cuts, for all panels and all values of A_C the median is closer to the 5th percentile than to the 95th percentile, indicating that the samples are skewed to lower halo masses. In Section 3.3.1 we showed that the intrinsic scatter in A at fixed mass is largely consistent with an un-skewed lognormal distribution, with only slight deviations at the high-end tail. The skew we see in Fig. 3.3 is due to the nature of the selection. Because there are more lower-mass haloes with relatively high values of A for their mass than there are higher-mass haloes with relatively low values of A for their mass, up scatter dominates over down scatter.

For most of the dynamic range shown in Fig. 3.3, all percentiles have a smooth, near power-law shape, with two exceptions. First, at low X-ray luminosities there is a sudden drop in the 5th percentile, indicating a large amount of scatter of the X-ray luminosity in haloes with masses $M_{500c} < 10^{13.5}\,M_\odot$. When we do not mask particles recently heated by AGN, the drop of the percentile moves to a higher X-ray luminosity. This suggests that for low halo masses increases in X-ray luminosity due to feedback are important. In Appendix 3.C we show that the drop in the 5th percentile does not disappear for a simulation with higher resolution, and is thus not a resolution effect. From Fig. 3.1 we know that for X-ray luminosity the importance of up-scatter increases for lower halo masses, as the distribution at fixed mass gains a tail towards higher X-ray luminosities. In particular, this deviation from lognormal is larger for X-ray than for SZ. Our findings in Fig. 3.3 indicate that the X-ray deviations from lognormal are strong enough to significantly skew the sample at masses $M_{500c} < 10^{13.5}\,M_\odot$. Second, for richness there is a clear deviation from the power-law shape for $\lambda < 10$. In addition, discreteness effects appear because a halo mass of about $10^{13}\,M_\odot$ is required for the richness to be larger than one. This behaviour is not affect by the resolution of the simulation, see Appendix 3.C, but does move to lower masses for higher resolutions.

3.3.4 Selection at fixed comoving number density

In the previous subsection we created samples by making a cut on a selection observable, and then compared the resulting sample with a mass-selected sample. Another approach of interest is to create an ordered list based on the values of a selection quantity and then selecting a sample based on a cut in the cumulative comoving number density of objects. We show the cumulative comoving number density as a function of mass, X-ray luminosity, SZ signal and galaxy richness in the different panels of Fig. 3.4. Different colours correspond to different redshifts. Comparing the different coloured solid lines, we see that a cut on comoving cumulative number density corresponds, as redshift increases, to a sample with

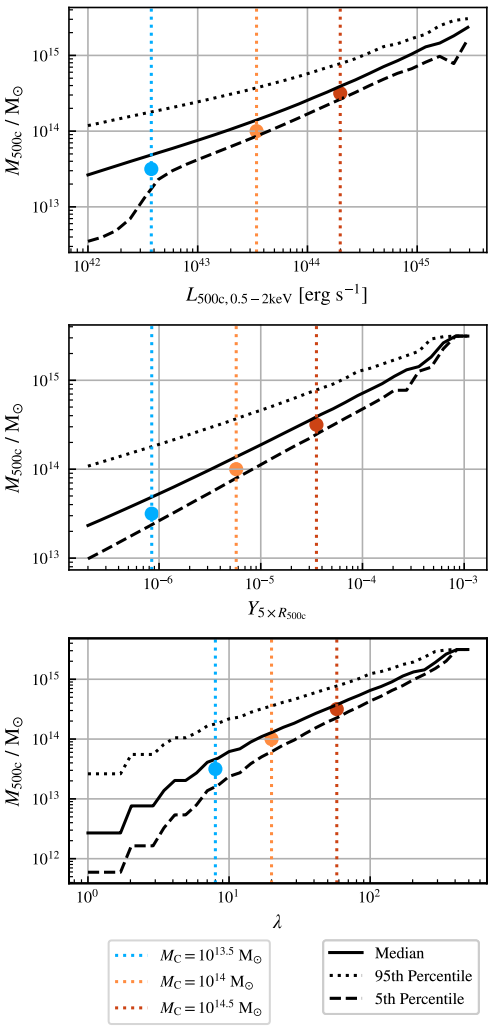


Figure 3.3: Solid lines show the median M_{500c} for a sample selected using a cut on the quantity plotted along the x-axis, i.e. median($M_{500c}|A > A_C$), for A_C given on the x-axis, at $z = 0.3$. From top to bottom, the different panels show the three different selection quantities A_C: X-ray, Compton-Y and galaxy richness. The dashed (dotted) line indicates the 5th (95th) percentile. The vertical dotted lines show the median values of each quantity at the masses indicated in the legend, with the corresponding mass indicated with a dot. Except for richness at $\lambda < 10$, the median relations are smooth and have shapes close to power-laws.

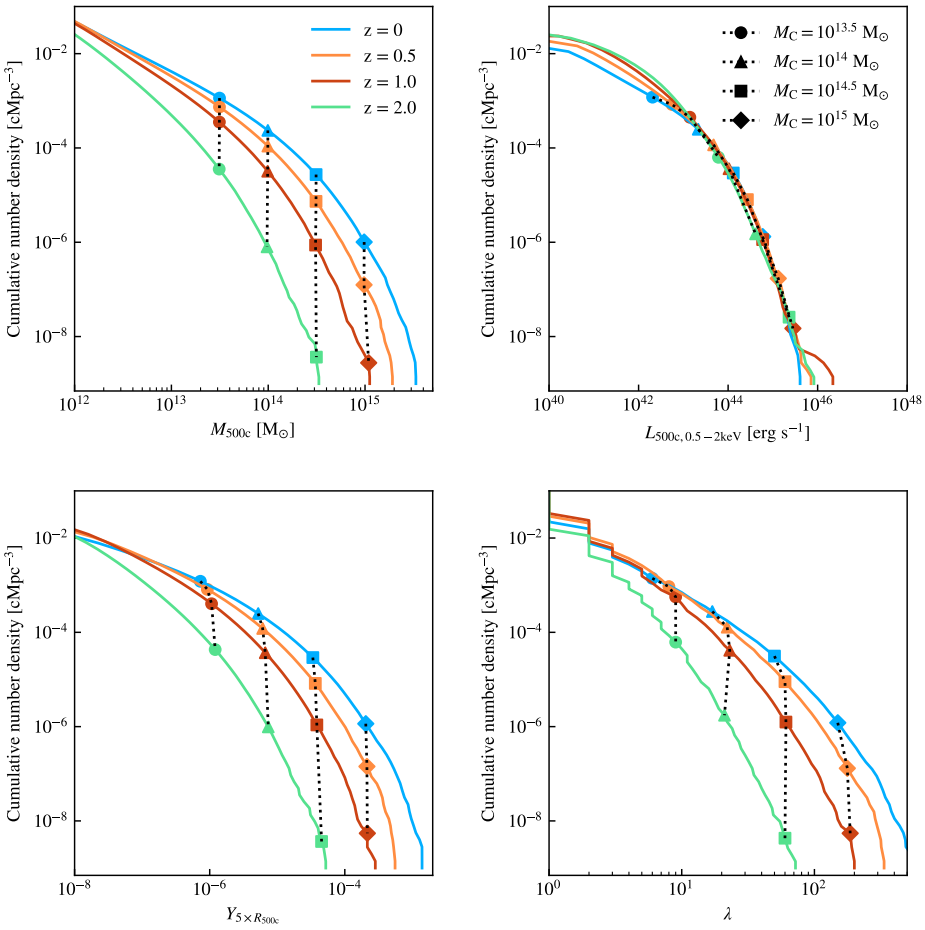


Figure 3.4: The cumulative comoving number density of all haloes as a function of M_{500c} (top left panel), X-ray luminosity (top right panel), SZ Compton-Y (bottom left panel) and galaxy richness (bottom right panel). The different colours correspond to different redshifts. The symbols indicate the number density and median value of each quantity at a fixed mass, with the different symbols corresponding to different masses. At higher redshifts, a cut at fixed comoving number density results in a sample with lower masses, Compton-Y and richness, but a distribution of X-ray luminosities that is almost independent of redshift. The dotted lines connect the markers for the same mass at different redshifts. Because at a fixed mass the X-ray luminosity increases with redshift, the number density above a fixed luminosity decreases less with redshift than for a mass-selected sample. Compton-Y and richness selected samples evolve similarly to a selection based on mass.

lower masses, Compton-Y values and richness values, but is close to an X-ray luminosity limited sample for number densities greater than $10^{-3}\,\mathrm{cMpc}^{-3}$.

Except for X-ray luminosity, the number density at a fixed value of the selection quantity decreases strongly with increasing redshift. For a mass-selected sample this is expected, because the halo mass function increases with time. For selection quantities for which the observable – mass relation does not evolve strongly, we expect the same qualitative trend, which is indeed seen for the SZ signal and, to a lesser extent, galaxy richness. Interestingly, for X-ray luminosity the different redshifts fall nearly on top of each other, except at the faint end. This implies that the evolution of the luminosity – mass relation nearly cancels the evolution of the mass function, with luminosity at fixed mass increasing with redshift. The very close agreement between the different redshifts must be a coincidence, because the number density – mass relation depends differently and more strongly on cosmology than the observable - mass relation. Note that, as a consequence, to create a mass-selected sample, we would need to select much higher X-ray luminosities, slightly higher Compton-Y values, and much smaller richness values at high redshift compared to $z = 0$.

It is helpful to consider the symbols connected with dotted lines, which inform us about the evolution of the selection quantity at fixed M_{500c}. For the SZ Compton-Y the dotted curves are nearly vertical, which implies that there is very little evolution in the mass – observable relation. For SZ the curves have negative slopes, bending slightly towards lower values at higher number densities. Because number density increases with time at fixed mass, this indicates a slight evolution towards smaller Compton-Y at fixed mass, as expected from the $E^{3/2}(z)$ scaling from self similarity (Kaiser, 1986, 1991). For X-ray luminosity the dotted lines bend strongly in the same direction, implying strong evolution towards lower luminosities at fixed mass, as expected from the $E^2(z)$ self-similar scaling. For galaxy richness the dotted curves behave similarly to Compton-Y, but slow slightly more evolution with redshift.

3.3.5 Sample mass bias as a function of the selection limit

The next step is to see how the sample mass bias changes with the selection limit $A > a$ and how it evolves with redshift. To indicate how different the sample is from a mass-selected sample with a mass cut M_C, which we will hold fixed at $10^{14}\,\mathrm{M}_\odot$, we compute the sample mass bias, as defined by Eq. 3.5, where we add one to the bias to allow for logarithmic plotting. The solid lines in Fig. 3.5 show the sample mass bias for the median, i.e. the factor by which the median mass of the sample with $A > a$, where a is plotted along the x-axis, differs from the median mass of the sample with $M_{500c} > M_C$. Similarly, the dashed lines show the sample mass bias for the 5th percentile. The different colours show different redshifts. The three panels show selections based on X-ray luminosity (top), SZ

signal (middle) and galaxy richness (bottom). The bias is defined with the respect to the sample with a mass cut of $M_C = 10^{14}\,M_\odot$.

For reference, the median values of observable A at the fixed mass M_C are indicated by the dotted vertical lines, one for each redshift. The vertical lines show strong redshift evolution of the value of the median X-ray luminosity at mass M_C, with the median luminosity increasing by over an order of magnitude from $z = 0$ to $z = 2$. For the SZ signal the effect is much milder, there is only a slight increase with redshift. Galaxy richness only exhibits mild evolution.

Observed clusters are distributed across a range of redshifts. If the observable-mass relation evolves, then applying a cut at a single value of the observable a can result in samples for which the mass distribution varies with redshift. This then leads to different sample mass biases for different redshifts. This effect is most pronounced for X-ray selection, as can be seen from the large differences between the different coloured solid lines. For example, while choosing a luminosity cut of 2×10^{43} erg s^{-1} yields a sample with a nearly unbiased median mass at $z = 0$, while at $z = 2$ the median mass is biased low by nearly an order of magnitude. Due to the strong evolution in the relation between X-ray luminosity and mass, any value selected for the X-ray luminosity cut will lead to a sample that becomes increasingly biased towards lower masses at higher redshifts. On the other hand, thanks to the mild redshift evolution for the SZ signal and galaxy richness, a cut on Compton-Y or λ will lead to a similar mass cut across different redshifts, thus allowing for the creation of a relatively unbiased sample. For a fixed cut in the observable, the value of the sample mass bias decreases with redshift for X-ray luminosity and SZ signal, but tends to increase with redshift for galaxy richness.

Examining the 5th percentiles, we see that they yield lower sample mass bias factors than for the medians (i.e. the dashed lines are below the solid lines of the same colour), indicating the sample is skewed towards lower masses. For cuts resulting in an unbiased median (i.e. samples with $A > a$ where the value a corresponds to the intersect of the vertical coloured dotted line and the horizontal black dotted line indicating $b_{M_{500c}} = 1$), the 5th percentile is biased low (i.e. the dashed line of the corresponding colour gives a bias value lower than unity). This means that the 5th percentile of the mass distribution of the sample with $A > a$, where a is chosen such that the median mass is the same as for a sample with $M > M_C$, is smaller than the 5th percentile of this mass-selected sample. This bias tends to increase with redshift and becomes particularly large for X-ray selection at $z = 2$.

Below a certain X-ray luminosity, the bias factor for the 5th percentile decreases rapidly to values $b_{M_{500c}} \ll 10^{-1}$. This suggests that for low halo masses, there is a large amount of scatter in the X-ray luminosity. This behaviour is similar to that for X-ray selection shown in Fig. 3.3. The sudden drop in the bias shifts to higher luminosities at higher redshifts. There are no similar drops in the bias

factor for SZ or richness selection.

3.3.6 Sample mass bias as a function of the target mass limit

Next we will investigate the bias in the median and 5th percentile M_{500c} for samples created with different observables as a function of the target mass M_C. To calculate the bias we use a cut based on the median of observable a at a fixed mass M_C (Eq. 3.4), which we denote as A_C. We then calculate the sample mass bias $b_{M_{500c}}$ for a range of M_C using Eq. 3.5. Since we use M_C to define the cut A_C, the sample mass bias becomes a function of only the target mass. This is shown for different observables in Fig. 3.6. From top to bottom, the mass cut is informed by a X-ray, SZ or richness selection limit. Each panel uses four distinct colours to represent various redshifts. The solid and dashed lines, respectively, depict the bias in the median and 5th percentile.

We first discuss some of the apparently odd features in each of the panels. Similar to what is shown in Figs. 3.3 and 3.5, there is a drop in the 5th percentile for the X-ray selection at low masses. The mass at which this happens increases with redshift and is 2×10^{13} $M_\odot$ at $z = 0$, increasing by a factor five at $z = 1$. Additionally, both biases exhibit a drop-off at the highest masses. This is caused by the fact that there are only very few halos for those mass bins. In that case even a few lower mass haloes that have a relatively high X-ray luminosity can quickly contaminate the sample and lead to a large bias.

At low masses, slightly above $M_{500c} = 10^{13}$ $M_\odot$, the richness selection demonstrates a sawtooth-like behaviour. This behaviour is directly linked to the discreteness issues inherent in our definition of richness. Every discrete value for the richness will have a range of halo masses for which it is the median at fixed mass. For this range of mass the richness selected sample will not change, and all the change is due to the mass selection. For each value of the richness there is a mass cut value that maximises the bias, and moving away from this value will always lead to an increasing bias. The decrease is turned around when the richness cut goes to the next discrete value, and then it will suddenly start to increase. This inherently leads to the lines going up and down with sudden changes in slope, which is seen in the figure as a saw tooth.

Now we will discuss the behaviour of the bias for each of the selections, starting with X-ray. At all redshifts the bias in the median mass for X-ray selection has a similar shape. Around $\sim 10^{14}$ $M_\odot$, the median bias is closest to zero, indicating a relatively unbiased selection, and it remains mostly flat around that mass range. Towards the highest masses, the bias has a sudden drop. The bias also slowly moves away from one towards lower masses. For $z = 1, 0.5$ and 0, the bias is close to -0.1 at the maximum, and only at $z = 2$ does the best possible bias decrease to just below -0.2. The 5th percentile exhibits more extreme evolution, with the plateau of roughly constant bias diminishing with increasing mass. At

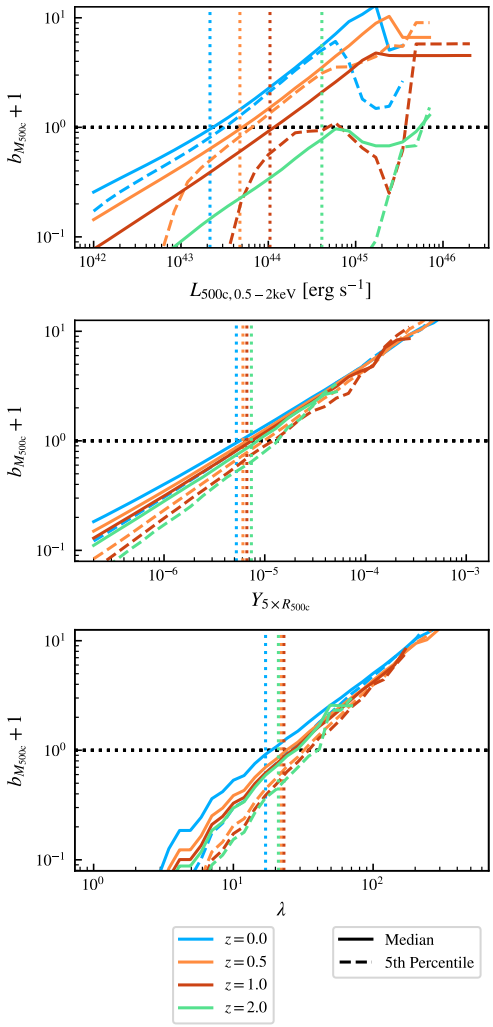


Figure 3.5: Each panel shows the bias factor for M_{500c} (Eq. 3.5) of a different sample relative to a mass-selected sample with a mass cut of $M_C = 10^{14}$ $M_\odot$. We add one to the sample mass bias to allow for logarithmic plotting. The panels show samples selected to have $A > a$, and thus contain all halos above the given threshold a, where A is X-ray luminosity (top panel), SZ Compton-Y (middle panel), or galaxy richness (bottom panel), and a is the value plotted along the x-axis. Solid and dashed lines show the bias for the median and the 5th percentile of the distribution, respectively. The different colours show the results for four different redshifts. The dotted vertical lines show the median value of each quantity at the mass M_C. For an unbiased sample both the solid and dashed curves would intersect the vertical dotted line at the y-axis value of unity.

$z = 2$, the 5th percentile is consistently biased by a factor of ten or more. In X-ray luminosity-based selections, optimal results are thus achieved by choosing a cut that maintains the median halo mass above $\sim 10^{14}$ $M_\odot$. This not only minimises bias but also prevents significant skewness in the distribution, especially at the 5th percentile.

The SZ selection consistently yields a median bias close to zero for the median across all masses and redshifts. The median bias increases slightly towards $\sim 10^{13} M_\odot$ but stays above -0.2 for all redshifts. This is in agreement with the results from Fig. 3.5. The SZ selection has little evolution with mass, and consistently provides relatively unbiased results ($-0.1 < b < 0$) for all redshifts. Significant evolution is observed for the 5th percentile, becoming more biased with increasing redshift. The bias in the 5th percentile shifts from -0.25 at $z = 0$ to approximately -0.5 at $z = 1$ and 2. The 5th percentiles become increasingly biased when the mass cut falls below $\sim 2 \times 10^{13}$ $M_\odot$. These results are for our fiducial SZ aperture of $5R_{500c}$. In Appendix 3.B we investigate the bias for a smaller aperture of R_{500c}. Using the smaller aperture the biases reduce further, leading to a nearly unbiased median over the entire mass range, and the bias in the 5th percentile reduces to only ~ 5 per cent. An aperture of R_{500c} thus leads to smaller biases for SZ selection.

With the exception of masses $M_{500c} \sim 10^{13}$ $M_\odot$ and at $z = 2$, richness selection leads to a median bias close to -0.1 that decreases slightly up to $z = 2$. At the lowest masses richness exhibits a slight sawtooth behaviour due to discreteness effects, but the bias does not drop significantly. At $z = 2$ the bias drops slightly more, reaching a value of slightly less than -0.2. The most interesting behaviour is found in the bias of the 5th percentile. Over the entire mass range, the bias in the 5h percentile increases with mass, going from -0.6 for $M_{500c} = 10^{13}$ $M_\odot$, to around -0.2 for $M_{500c} = 10^{15}$ $M_\odot$. The 5th percentile also becomes more biased at $z = 2$.

For masses below 10^{14} $M_\odot$ as well as at $z \geq 1$, using an SZ selection yields the least biased results for both the median and the 5th percentile. In those regimes, the X-ray selection exhibits a substantial influx of smaller haloes 'up-scattering' into the sample, resulting in a stronger bias. For richness the median sample mass bias is similar to the SZ selection, but there is a much larger skew in the 5th percentile. When we examine selections above $\sim 10^{14}$ $M_\odot$ at $z = 0$ and 0.5 the three selections exhibit closer bias values, and there is no longer a clear 'best' choice. Regarding the bias on the median, the only outlier occurs for masses close to and larger than 10^{15} $M_\odot$ in the case of an X-ray selection. In this scenario, both the median and the 5th percentile exhibit significant bias, and opting for either an SZ or richness selection yields better results.

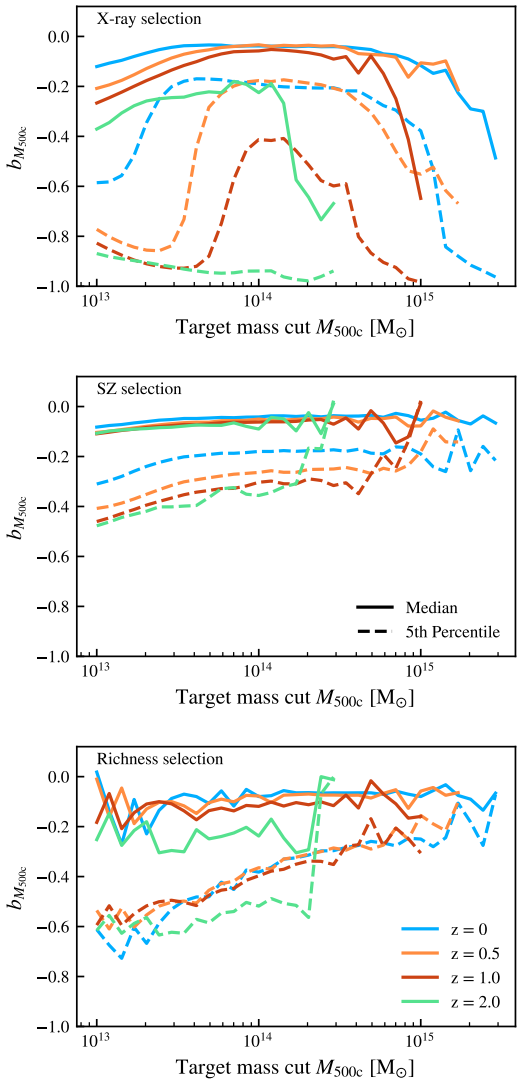


Figure 3.6: The sample M_{500c} bias (Eq. 3.5) as a function of the target mass cut. The cut used for X-ray luminosity (top panel), Compton-Y (middle panel) and galaxy richness (bottom panel) is the median value for the mass cut plotted along the x-axis (see Eq. 3.4). The solid and dashed lines shows the sample mass bias for the median and 5th percentile, respectively. The different line colours show the results at different redshifts. The bias in the median mass increases towards low target masses and, for X-ray selection also towards high masses. While there are target masses for which the median mass is only slightly biased low, the 5th percentiles of the mass distribution are always much lower than for a mass-selected sample. The biases tend to increase with redshift.

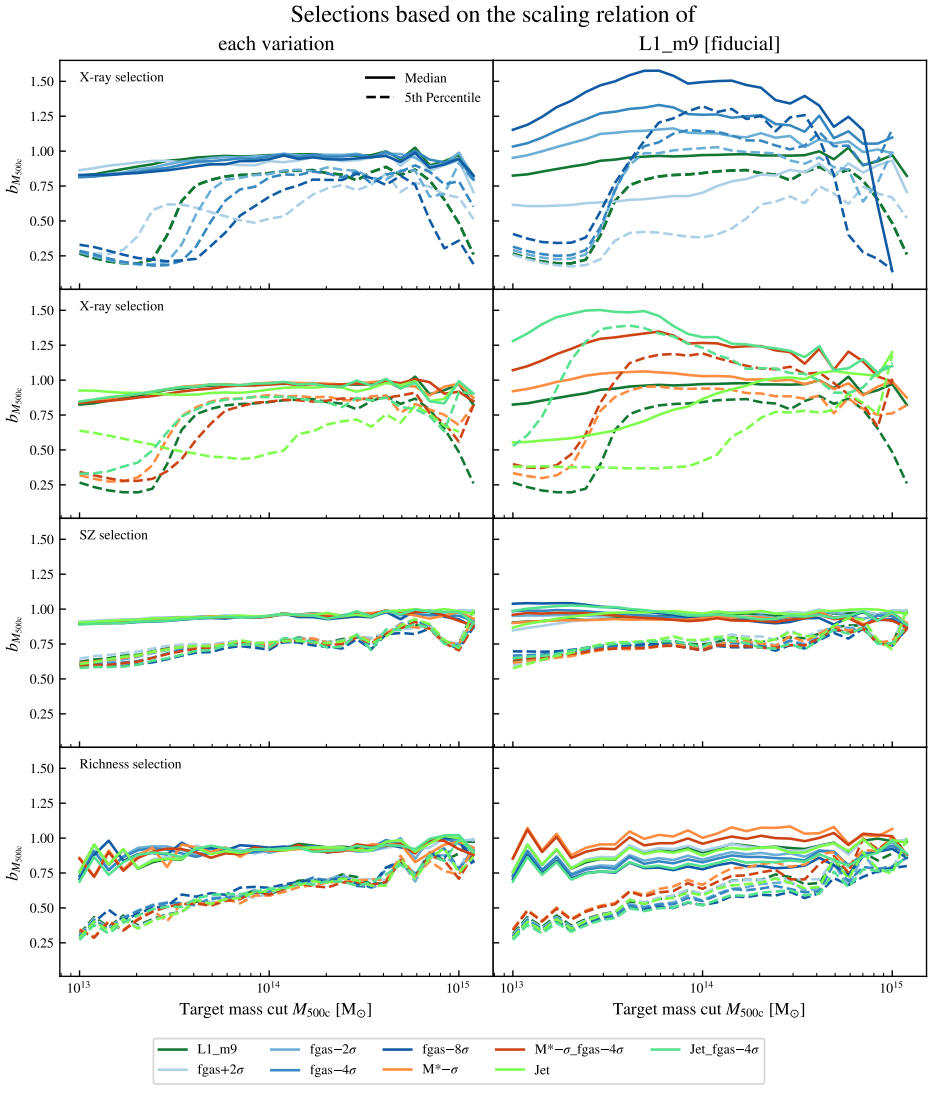

Selections based on the scaling relation of
each variation
L1_m9 [fiducial]
X-ray selection
Median
5th Percentile
X-ray selection
SZ selection
Richness selection
$b_{M_{500c}}$
Target mass cut M_{500c} [M$_\odot$]
L1_m9
fgas+2σ
fgas−2σ
fgas−4σ
fgas−8σ
M*−σ
M*−σ_fgas−4σ
Jet
Jet_fgas−4σ

Figure 3.7: Same as Fig. 3.6 for a single redshift, $z = 0.3$, but for different FLAMINGO models (different colours). In the left column the selection is based on the median of observable a in each simulation at the target mass cut plotted along the horizontal axis. Differences between models are thus due to the differences in the scatter in the observable-mass relations. For all models shown in the right column the selection is based on the median relation of observable a in the L1_m9 simulation. Differences are therefore due to both changes in the scatter and changes in the median of the observable-mass relations. The labels "fgas" indicate runs with a change in the gas fractions, where a lower number of sigma indicates a lower gas fraction and thus stronger feedback. The M* label indicates runs for which the stellar mass function is shifted to lower masses. The label "Jet" indicates that the AGN feedback model uses kinetic jets instead of the thermally driven winds used for the other runs. Note that the top two rows both show results for X-ray selection, but for different sets of simulations. The results for X-ray selection are distributed over two panels for visual clarity. Only X-ray selection and, to a lesser extent, richness selection are sensitive to changes in the gas fraction or the stellar mass function.

3.3.7 The effect of modelling uncertainty

One potential reason for concern is that our conclusions might be influenced by the properties of clusters realised in the simulation and that these properties may not be modelled with sufficient accuracy. To examine the effect of varying the cluster properties, Fig. 3.7 shows the sample mass bias as a function of the target mass cut at $z = 0.3$ for all the FLAMINGO feedback variations. As shown by (Schaye et al., 2023; Braspenning et al., 2023), the cosmology variation have no significant impact on the scaling relations, and are therefore not considered in this work. These variations consist of models that vary the hot gas content and/or the stellar mass function, by changing the strengths of stellar and AGN feedback, or that use jet-like instead of thermal AGN feedback. In the left column, the X-ray luminosity, Compton-Y, and richness cuts correspond to the median value of the observable as a function of the target mass cut. For each model variation the cut therefore corresponds to the same target mass cut. In the right column we instead fix the X-ray, SZ and richness cuts to those obtained for the fiducial L1_m9 simulation for the target mass cut. This translates to setting $A_{\rm C}(M_{\rm 500c}) = A_{\rm C,L1_m9}(M_{\rm 500c})$ for each variation, i.e. we assume a slightly wrong observable-mass scaling relation for the model variations. Therefore, the left column shows the effect of changes in the scatter in the observable-mass relation between the different models and the right columns shows the combined effect

of changing the scatter and ignoring the effect of the change in model on the median observable-mass relation.

Starting with the left column, which shows the effect of changing the scatter in the observable-mass relation, the results are similar for all model variations. Except for the 5th percentile of X-ray selected clusters for low target masses, the bias is generally insensitive to variations in the model. For X-ray there is a slight trend where a lower gas fraction (i.e. fgas$-N\sigma$) is associated with a slightly more biased median mass, but the effect is small. The shapes of the curves are different for the fgas+2σ and Jet models, particularly for the bias on the 5th percentile. For SZ- and richness-selected samples the bias factors are insensitive to the model.

In the right column, which shows the combined effect of the model variation on the scatter and the mass-observable relation, we find larger though still small model-dependence for the sample mass bias in SZ-selected samples. The variations change the bias by $\Delta b_{M_{500c}} \approx 0.05 - 0.1$ and the general shape of the dependence on the target mass does not change. For richness-selection, we find a slight trend with gas fraction, and a deviation of about 0.1 in bias for the models with a stellar mass function shifted to lower stellar masses. The dependence on stellar mass is expected, as we apply a stellar mass cut for our definition or richness. In contrast with SZ, the differences between feedback variations are larger for higher mass objects. For X-ray selection, changing the simulation without changing the selection limit to account for the change in the mass-observable relation has a large impact. The bias on the median mass changes from ≈ 0.5 to ≈ -0.5 going from the lowest to highest fgas variation. This implies that having complete knowledge of the true scaling relation is essential. Any deviations between the true scaling relation and the one that is assumed when modelling selections effects will lead to a biased sample.

The fact that X-ray selection is most affected by variations in the model is to be expected. From Braspenning et al. (2023) we know that the different variations have different electron densities in the cluster cores. The X-ray luminosity scales as ρ^2 and is therefore more sensitive to feedback processes affecting the core than Compton-Y, which scales as ρ. From Fig. 7 of Kugel et al. (2023) or Fig. 10 of Schaye et al. (2023) we can see that the gas fractions of all models start to converge for high masses, just as the sample mass bias start to converge for high masses in the top right panel of Figure 3.7, though substantial differences remain even at the highest masses. It is also clear that the behaviour is not fully determined by the gas fraction, as the bias for the jet and the M* variations do not agree with their corresponding fgas variations. This further emphasises the fact that direct knowledge of the observable-mass scaling relations is important, and that we cannot rely solely on indirect measurements.

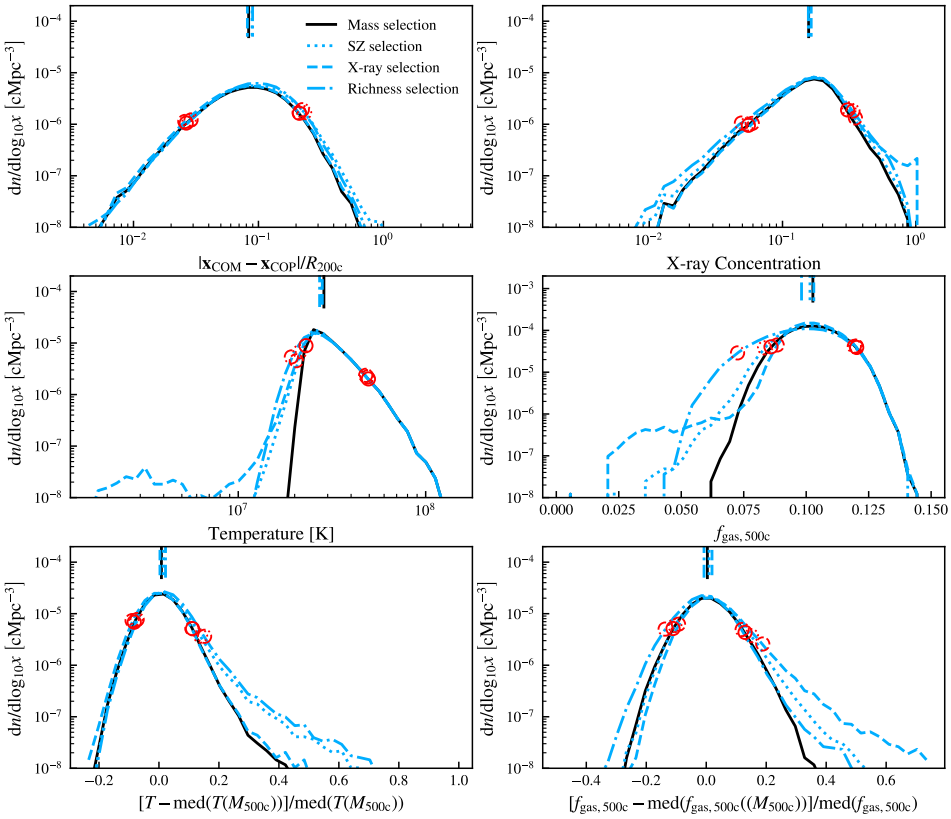


Figure 3.8: Distribution of relaxedness (top left; Eq. 3.8), X-ray concentration (top right; Eq. 3.9), mass-weighted mean temperature (middle left), gas fraction (middle right), relative deviation from the median mass-weighted mean temperature at fixed mass (bottom left) and relative deviation from the median $f_{\mathrm{gas,500c}}$ at fixed mass (bottom right) for selections with a target mass of $M_{\mathrm{C}} = 10^{14}\ \mathrm{M}_{\odot}$ at $z = 0.3$. The different line styles indicate selections based on different properties. The median of each sample is indicated with a vertical line at the top of each panel. The 5th and 95th percentiles are shown using red circles. Selection effects result in biased distributions of cluster properties relative to a mass-selected sample. This is mostly due to upscatter from lower masses, but the bottom row shows that there are biases even at fixed true mass.

3.3.8 Biases in properties other than mass

So far we have looked at how different selections bias the mass distribution of the cluster samples. When looking beyond the effects on cluster count cosmology, we want to inspect what the impact of different selections is on other properties of clusters. Even if the mass is measured independently, the lower mass objects that up-scatter into the selection could give a biased view of how scaling relations extrapolate towards lower masses.

There are a few cluster properties that are of particular interest. Lovisari et al. (2017), Rossetti et al. (2017) and Andrade-Santos et al. (2017) report differences in the disturbed fraction and the cool core fraction when comparing X-ray- and SZ-selected samples. Besides the disturbed fraction and cool core fraction, we also investigate biases in the median temperature and gas fraction.

To quantify the degree of disturbedness in FLAMINGO, we compute the relaxedness parameter, defined as

$$\text{Relaxedness} = \frac{|\mathbf{x}_{\rm COM} - \mathbf{x}_{\rm COP}|}{R_{\rm 200c}}, \tag{3.8}$$

where $\mathbf{x}_{\rm COM}$ is the position of the center of mass of the halo, defined by all the particles bound to the subhalo, $\mathbf{x}_{\rm COP}$ is the location of the most bound particle in the halo, and $R_{\rm 200c}$ is the radius within which the average density is equal to two hundred times the critical density. Note that a higher relaxedness value indicates a cluster that is more disturbed.

In order to trace whether a cluster is cool-core, we use the X-ray concentration, defined as

$$\text{X-ray concentration} = \frac{L_{X,r<0.15R_{\rm 500c}}}{L_{X,r<R_{\rm 500c}}}, \tag{3.9}$$

where $L_{X,r<0.15R_{\rm 500c}}$ is the X-ray luminosity in the core of the halo, defined by $0.15R_{\rm 500c}$ and $L_{X,r<R_{\rm 500}}$ is the total X-ray luminosity within $R_{\rm 500c}$. The higher the X-ray luminosity concentration, the more likely a cluster is to have a cool core.

We also measure the mass-weighted mean temperature, excluding gas below 10^5 K, and the gas mass fraction, each within $R_{\rm 500c}$. Additionally, since both the temperature and the gas fraction have a strong dependence on halo mass, we measure their deviations from the median at a fixed mass,

$$\Delta X = \frac{X - \text{median}(X(M_{\rm 500c}))}{\text{median}(X(M_{\rm 500c}))}. \tag{3.10}$$

This way we can investigate whether the lower mass haloes that up-scatter have different values for the temperature and gas fraction than a mass-selected sample.

To investigate how the different selections bias these quantities, we create a sample using a target mass $M_{\rm C} = 10^{14}$ $\rm M_\odot$ for each observable a, as well as a mass-selected sample. In Figure 3.8 we show the distributions of these quantities

at $z = 0.3$. On the y-axis we show the bin-size normalised number density. The different line styles indicate the different selection methods used. The mass selection (black solid curve) should be taken as the baseline to compare the other selections with. We show the median of each selection with a vertical line at the top of the plot, and the 5th and 95th percentiles using red circles.

The top left panel shows the distribution of relaxedness, the offset between the center of potential and center of mass. We do not find strong differences between the different selection methods, the medians and percentiles are similar, and close to those of the mass selected sample. For the most disturbed objects, with the highest value of the offset, there is a slight trend where an SZ selection yield more highly disturbed objects, but this trend is very slight.

In the top right panel we show the distribution of X-ray concentrations. Andrade-Santos et al. (2017) used a similar metric to divide clusters into cool-core and non-cool-core clusters. A higher value indicates a more centrally concentrated X-ray luminosity, implying that the cluster is more likely to be a cool-core cluster. Richness selection does not lead to a clear preference between more or less cool-core objects. For X-ray and SZ we find results qualitatively similar to those of Andrade-Santos et al. (2017). There is both an enhancement in the number of clusters with high X-ray concentration for the X-ray selection, and an enhancement of object with low X-ray concentration for SZ selection. However, as can been seen in the medians and percentiles, this difference is quite small.

The middle two panels of Figure 3.8 show the distributions of the temperature and gas fraction. For both these quantities, the differences relative to mass selection stem mainly from the fact that up-scattered haloes have lower halo masses, which implies that selections with more up-scattered haloes contain more objects with a low temperature and a low gas fraction. The most massive haloes, which have the highest temperatures and gas fractions, are included in each selection. This is reflected in the medians and 95th percentile, which do not change significantly, with the exception of the median gas fraction for richness selection. All selections are therefore complete for high temperatures and gas fractions. The samples selected on observables other than mass include more objects that have relatively low temperatures and gas fractions. For the temperature, the distribution of these objects is similar to what is found in Figure 3.2, indicating that the differences are largely mass-driven. These panels show that many of the lower-mass haloes that up-scatter into each selection have a significantly lower temperature and gas fraction than the haloes in a mass-selected sample. However, they do not tell us whether the haloes that are now included are different from other haloes of the same mass. This is investigated in the bottom two panels.

The bottom left panel of Figure 3.8 shows the relative deviation from the median temperature at the true halo mass for the different selections (see Eq. 3.10). By plotting this relative difference we can investigate how the temperatures are

biased with respect to the median at their given mass. In this case the X-ray selection does not bias the sample substantially, but for the SZ and richness selections there is a pronounced tail towards haloes with much higher temperatures. This slightly increases the 95th percentile, but does not change the median significantly.

The bottom right panel of Figure 3.8 shows the deviation from the median gas fraction within R_{500c} at fixed halo mass for the different selections. Richness selection increases the scatter, but there is no clear preference for higher or lower gas fractions as the median does not change. For both X-ray and SZ selections there is a preference for objects with a gas fraction that is high for their mass, though the median gas fractions are nearly the same. X-ray selection finds slightly fewer haloes with a relatively low gas fraction than mass selection. This implies that even for haloes of a fixed mass, X-ray selection will already lead to a slight bias towards higher gas fractions. For both SZ and X-ray selection the clusters in the sample tend to have gas fractions that are higher than the average population, even at a fixed mass. For the X-ray sample, the 95th percentile increases by about $\sim 20\%$ and higher percentiles are biased more strongly. This bias will be stronger closer to the survey selection limit, i.e. for lower masses. This is consistent with the findings by Kugel et al. (2023), who attributed the fact that the observed relation between X-ray gas fraction and mass flattens off below 7×10^{13} $M_\odot$ to selection effects.

For three of the four quantities investigated, i.e. X-ray concentration, temperature, and gas fraction, we find that selecting on an observable other than mass introduces slight biases compared to a mass-selected sample. For clusters with masses larger than the median of the sample these effects will be negligible. However, upscatter results in the addition of lower mass haloes with temperatures and gas fractions that are lower than for a mass-selected sample. For richness selection this upscatter results in significant negative biases for the median gas fraction, while the bias in the medians is negligible for other selections. Even for the 5th percentiles the biases are small, with the exception of richness selection. Comparing the temperatures and gas fractions to the median values for the true mass of each selected halo, we find again that the medians are nearly unbiased, but there is a tail towards higher temperatures and gas fractions. For richness selection the up-scattered haloes also have a tail towards lower gas fractions relative to that expected for their true mass. However, for the 5th and 95th percentiles the biases are still small. As all our selections are intrinsically volume-limited, we find results similar to Chon & Böhringer (2017), and we note that observed differences between differently selected samples may be more influenced by the difference between volume- and flux-limited surveys than the selection method.

While these results cannot explain the relatively large sample mass biases that we found in earlier sections, they do show that some of the biases in cluster properties other than mass are intrinsically correlated with the chosen selection

method.

3.4 Conclusions

Given their large volumes and good agreement with observation, as well as the availability of a large number of model variations, the FLAMINGO simulations (Schaye et al., 2023; Kugel et al., 2023) provide an opportunity to investigate how different galaxy cluster selection methods influence the resulting samples. This is crucial for cluster cosmology (e.g. Allen et al., 2011; Mantz, 2019), but also for understanding the role of selection biases in cluster scaling relations.

We used the FLAMINGO simulations to investigate how the samples obtained from cuts in X-ray luminosity, thermal SZ Compton-Y (integrated within an aperture of $5R_{500c}$), or galaxy richness (using satellite galaxies with stellar mass $> 10^{10.046}\,\mathrm{M}_\odot$) are biased in terms of the median and other percentiles of the mass distribution and certain secondary quantities. We summarise our findings as follows:

- The scatter in X-ray luminosity, Compton-Y and richness increases with decreasing halo mass (see Fig. 3.1). At fixed mass only the central parts of the distributions are lognormal. The distributions of X-ray luminosity and Compton-Y have power-law tails towards higher values, while for richness there can also be a tail towards lower values. The tails in the distributions cause the number of haloes that up-scatter into an X-ray or SZ selected sample to be underestimated when assuming lognormal scatter.

- In Fig. 3.2 we compared the distributions of halo mass, X-ray luminosity, Compton-Y and richness for a target mass cut of $M_{500c} = 10^{14}\,\mathrm{M}_\odot$ at $z = 0.3$ for samples selected by mass or by $A > A_C$ where $A_C = \mathrm{median}(A(M_{500c} = 10^{14}\,\mathrm{M}_\odot))$ and A is the observable. We found tight correlations between all quantities for $A > A_C$, but not for lower values. Selecting based on richness leads to the largest amount of contamination by low-mass haloes, while X-ray selection yields the least amount of contamination.

- As shown in Fig. 3.3, increasing the selection limit in terms of X-ray luminosity, Compton-Y or richness leads to a sample with a smoothly increasing median and 95th percentile mass. However, for an X-ray luminosity cut smaller than $10^{43}\,\mathrm{erg\,s^{-1}}$, the 5th percentile of the mass distribution dips to very low masses. This effect is converged with the numerical resolution (see Fig. 3.11) and is qualitatively robust to changes in the subgrid feedback modelling (see Fig. 3.7).

- The comoving number density above a fixed X-ray luminosity (richness) decreases less (more) with increasing redshift than for a mass-selected sample.

A Compton-Y or richness selected sample evolves similarly to a selection based on mass (see Fig. 3.4).

- For a fixed target mass cut of $M_{500c} = 10^{14}$ M$_\odot$, the corresponding X-ray luminosity cut increases by more than an order of magnitude from $z = 0$ to 2, while the richness cut decreases by about a factor of 3. For Compton-Y and richness the cut remains nearly constant with redshift (see Fig. 3.5).

- The bias in the median mass becomes stronger towards lower target masses and, for X-ray selection, also towards high masses. While there are target masses for which the median mass is only biased slightly low, the 5th percentiles of the mass distribution are always much lower than for a mass-selected sample. The samples tend to become more biased with increasing redshift. The target mass range for which the median sample mass bias is small is largest for SZ selection (see Fig. 3.6). When using a R_{500c} aperture instead of our fiducial $5R_{500c}$ aperture, the SZ selection becomes even less biased (Appendix 3.B).

- Except for the 5th percentile of X-ray selected samples, and provided the median observable-mass relation is known, the bias factors are nearly the same for models calibrated to yield different gas fractions or stellar masses, and also for models using a different implementation of AGN feedback (Fig. 3.7).

- The different selections lead to slight biases in cluster properties other than mass. In Figure 3.8 we demonstrated this for a target mass $M_{500c} = 10^{14}$ M$_\odot$. For X-ray selection, the lower mass objects that up-scatter into the sample have a very slight preference to have high X-ray concentrations, which is indicative of a cool core, while the opposite is true for selection based on richness. SZ selection includes slightly more clusters that are disturbed. Due to up-scatter of lower-mass haloes, all selections result in the inclusion of objects with temperatures and gas fractions that are much lower than are present in the mass-selected sample. However, compared with the median values for their true mass, the up-scattered objects tend to have high temperatures and gas fractions. Most of these effects are minor, leading to only small changes in the median and the 5th and 95th percentiles.

For each of the three selection methods, there are regimes in which the samples obtained have a small median sample mass bias. However, the 5th percentile of the mass distribution is nearly always biased significantly low and the biases tend to increase with redshift. Overall, SZ selection gives results that are closest to mass selection.

Overall, our results highlight how important it is that the scaling relations between mass and its observational proxies, including the scatter, are measured

and modelled accurately. Even slight biases in the mass distributions can lead to differences that are problematic for surveys aimed at measuring cosmological parameters using cluster counts. We aim to investigate the direct effect of these biases on clusters counts in future work.

We have shown that the objects with the lowest masses in each sample are more likely to be outliers with respect to the overall population when it comes to cluster properties other than the mass proxies. This can lead to biases when observationally determining scaling relations for quantities like the temperature and gas fractions.

In this work we have investigated selections based on observables in theory-space. We have ignored observational measurement errors, lightcone effects, projection effects, fore- and backgrounds, the effects of changing the cosmology, and other systematic effects, many of which will be survey specific. We have also implicitly assumed that observationally selection depends solely on the proxies investigated here, whereas in reality the signal-to-noise of a detection will depend on other properties. For example, X-ray selection likely depends not just on luminosity, but also on surface brightness (see e.g. Andreon et al., 2024). Galaxy richness does not rely on stellar mass selection, but depends on the luminosity and colour of the galaxies, as well as their distribution in phase space. It will be important to include such effects in future work, e.g. by forward modelling observational selection based on virtual observations created using the FLAMINGO lightcones.

Acknowledgements

This work is partly funded by research programme Athena 184.034.002 from the Dutch Research Council (NWO). This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 769130). VJFM acknowledges support by NWO through the Dark Universe Science Collaboration (OCENW.XL21.XL21.025). This work used the DiRAC@Durham facility managed by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). The equipment was funded by BEIS capital funding via STFC capital grants ST/K00042X/1, ST/P002293/1, ST/R002371/1 and ST/S002502/1, Durham University and STFC operations grant ST/R000832/1. DiRAC is part of the National e-Infrastructure.

Data Availability

All data presented in this paper will be shared upon reasonable request to the corresponding author.

References

Abbott T. M. C., et al., 2022, Phys. Rev. D, 105, 023520

Allen S. W., Evrard A. E., Mantz A. B., 2011, ARA&A, 49, 409

Andrade-Santos F., et al., 2017, ApJ, 843, 76

Andreon S., Trinchieri G., Moretti A., 2024, Observed abundance of X-ray low surface brightness clusters in optical, X-ray, and SZ selected samples (arXiv:2404.12435)

Angulo R. E., Springel V., White S. D. M., Jenkins A., Baugh C. M., Frenk C. S., 2012, MNRAS, 426, 2046

Aricò G., Angulo R. E., Contreras S., Ondaro-Mallea L., Pellejero-Ibañez M., Zennaro M., 2021, MNRAS, 506, 4070

Artis E., Melin J. B., Bartlett J. G., Murray C., Euclid Consortium 2022, in mm Universe @ NIKA2 - Observing the mm Universe with the NIKA2 Camera. p. 00004 (arXiv:2111.05432), doi:10.1051/epjconf/202225700004

Bahé Y. M., et al., 2017, MNRAS, 470, 4186

Bahé Y. M., et al., 2022, MNRAS, 516, 167

Barnes D. J., Kay S. T., Henson M. A., McCarthy I. G., Schaye J., Jenkins A., 2017, MNRAS, 465, 213

Black W. K., Evrard A., 2022, MNRAS, 516, 1170

Bleem L. E., et al., 2024, The Open Journal of Astrophysics, 7, 13

Bocquet S., et al., 2019, ApJ, 878, 55

Bocquet S., Heitmann K., Habib S., Lawrence E., Uram T., Frontiere N., Pope A., Finkel H., 2020, ApJ, 901, 5

Bocquet S., et al., 2024, arXiv e-prints, p. arXiv:2401.02075

Booth C. M., Schaye J., 2009, MNRAS, 398, 53

Borrow J., Schaller M., Bower R. G., Schaye J., 2022, MNRAS, 511, 2367

Braspenning J., et al., 2023, arXiv e-prints, p. arXiv:2312.08277

Chaikin E., Schaye J., Schaller M., Benítez-Llambay A., Nobels F. S. J., Ploeckinger S., 2022a, arXiv e-prints, p. arXiv:2211.04619

Chaikin E., Schaye J., Schaller M., Bahé Y. M., Nobels F. S. J., Ploeckinger S., 2022b, MNRAS, 514, 249

Chaubal P. S., et al., 2022, ApJ, 931, 139

Chisari N. E., et al., 2019, The Open Journal of Astrophysics, 2, 4

Chiu I. N., Klein M., Mohr J., Bocquet S., 2023, MNRAS,

Chon G., Böhringer H., 2017, A&A, 606, L4

Clerc N., et al., 2024, arXiv e-prints, p. arXiv:2402.08457

Costanzi M., et al., 2019, MNRAS, 488, 4779

Cui W., et al., 2018, MNRAS, 480, 2898

Dalla Vecchia C., Schaye J., 2008, MNRAS, 387, 1431

Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, MNRAS, 486, 2827

Debackere S. N. B., Schaye J., Hoekstra H., 2021, MNRAS, 505, 593

Debackere S. N. B., Hoekstra H., Schaye J., Heitmann K., Habib S., 2022a, MNRAS, 515, 3383

Debackere S. N. B., Hoekstra H., Schaye J., 2022b, MNRAS, 515, 6023

Driver S. P., et al., 2022, MNRAS, 513, 439

Elbers W., Frenk C. S., Jenkins A., Li B., Pascoli S., 2021, MNRAS, 507, 2614

Elbers W., Frenk C. S., Jenkins A., Li B., Pascoli S., 2022, MNRAS, 516, 3821
Evrard A. E., Arnault P., Huterer D., Farahi A., 2014, MNRAS, 441, 3562
Farahi A., Evrard A. E., McCarthy I., Barnes D. J., Kay S. T., 2018, MNRAS, 478, 2618
Ferland G. J., et al., 2017, Rev. Mex. Astron. Astrofis., 53, 385
Ghirardini V., et al., 2024, arXiv e-prints, p. arXiv:2402.08458
Giles P. A., et al., 2022, MNRAS, 516, 3878
Giri S. K., Schneider A., 2021, arXiv e-prints, p. arXiv:2108.08863
Grandis S., et al., 2020, MNRAS, 498, 771
Grandis S., et al., 2021, MNRAS, 504, 1253
Hahn O., Martizzi D., Wu H.-Y., Evrard A. E., Teyssier R., Wechsler R. H., 2017, MNRAS, 470, 166
Hahn O., Rampf C., Uhlemann C., 2021, MNRAS, 503, 426
Han J., Cole S., Frenk C. S., Benitez-Llambay A., Helly J., 2018, MNRAS, 474, 604
Heymans C., et al., 2021, A&A, 646, A140
Hirschmann M., Dolag K., Saro A., Bachmann L., Borgani S., Burkert A., 2014, MNRAS, 442, 2304
Huško F., Lacey C. G., Schaye J., Schaller M., Nobels F. S. J., 2022, MNRAS, 516, 3750
Kaiser N., 1986, MNRAS, 222, 323
Kaiser N., 1991, ApJ, 383, 104
Kaviraj S., et al., 2017, MNRAS, 467, 4739
Kugel R., et al., 2023, MNRAS, 526, 6103
Lovisari L., et al., 2017, ApJ, 846, 51
Mantz A. B., 2019, MNRAS, 485, 4863
Marini I., et al., 2024, Detecting Galaxy Groups and AGNs populating the local Universe in the eROSITA era (arXiv:2404.12719)
McCarthy I. G., Schaye J., Bird S., Le Brun A. M. C., 2017, MNRAS, 465, 2936
Miyatake H., et al., 2023, arXiv e-prints, p. arXiv:2304.00704
Nelson D., Pillepich A., Ayromlou M., Lee W., Lehle K., Rohr E., Truong N., 2023, arXiv e-prints, p. arXiv:2311.06338
Ota N., et al., 2023, A&A, 669, A110
Pacaud F., et al., 2018, A&A, 620, A10
Pakmor R., et al., 2023, MNRAS, 524, 2539
Pellissier A., Hahn O., Ferrari C., 2023, MNRAS, 522, 721
Pillepich A., et al., 2018, MNRAS, 473, 4077
Planck Collaboration et al., 2016a, A&A, 594, A24
Planck Collaboration et al., 2016b, A&A, 594, A27
Planck Collaboration et al., 2020, A&A, 641, A6
Ploeckinger S., Schaye J., 2020, MNRAS, 497, 4857
Ramos-Ceja M. E., et al., 2022, A&A, 661, A14
Riess A. G., et al., 2022, ApJ, 934, L7
Rossetti M., Gastaldello F., Eckert D., Della Torre M., Pantiri G., Cazzoletti P., Molendi S., 2017, MNRAS, 468, 1917
Rozo E., Bartlett J. G., Evrard A. E., Rykoff E. S., 2014, MNRAS, 438, 78
Rykoff E. S., et al., 2014, ApJ, 785, 104
Rykoff E. S., et al., 2016, ApJS, 224, 1
Salcido J., McCarthy I. G., Kwan J., Upadhye A., Font A. S., 2023, MNRAS, 523, 2247
Schaller M., et al., 2024, MNRAS, 530, 2378
Schaye J., Dalla Vecchia C., 2008, MNRAS, 383, 1210
Schaye J., et al., 2015, MNRAS, 446, 521
Schaye J., et al., 2023, MNRAS, 526, 4978
Springel V., Di Matteo T., Hernquist L., 2005, MNRAS, 361, 776
Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709
Upsdell E. W., et al., 2023, MNRAS, 522, 5267
Wendland H., 1995, Advances in Computational Mathematics, 4, 389

Wiersma R. P. C., Schaye J., Theuns T., Dalla Vecchia C., Tornatore L., 2009, MNRAS, 399, 574
Willis J. P., et al., 2021, MNRAS, 503, 5624

Appendix

3.A Fits at different redshifts

Tables 3.2 - 3.5 contain the fits similar to those in Section 3.3.1 at the four other redshifts considered in this work. For each redshift, the general trends are similar to those found at $z = 0.3$. We note that the fits for richness should be considered with care as they are not converged with the simulation resolution and, for most redshifts, the mass bins $M_{500c} > 10^{14.0}$ $M_\odot$ have a mean richness that is above ten. As the mean richness is very close to 10 for the $10^{13.5}$ $M_\odot$ mass bin, going to that mass or lower will likely lead to results that suffer from small-number statistics. We omit the highest mass bin at $z = 2$ as there are not enough high mass halos in the simulation volume to characterize the distribution.

Table 3.2: Values for the fits to Eq. 3.6 at $z = 0$. The top four rows are for selection based on X-ray luminosity, the middle four for integrated Compton-Y, and the bottom for for galaxy richness. Note that for richness we only fit a lognormal, so we do not the include the parameters for the power-law tail.

a	$M_{500c}[M_\odot]$	A	μ	σ	$\log_{10} a_t$	α
X-ray	$10^{13.0}$ $M_\odot$	1.810×10^{-3}	40.8	0.37	41.5	2.06
X-ray	$10^{13.5}$ $M_\odot$	2.675×10^{-3}	42.0	0.24	42.2	3.57
X-ray	$10^{14.0}$ $M_\odot$	3.487×10^{-3}	42.9	0.19	43.2	4.23
X-ray	$10^{14.5}$ $M_\odot$	4.481×10^{-3}	43.8	0.15	43.9	6.11
SZ	$10^{13.0}$ $M_\odot$	3.737×10^{-3}	-7.18	0.24	-7.03	2.14
SZ	$10^{13.5}$ $M_\odot$	5.608×10^{-3}	-6.15	0.17	-6.05	3.47
SZ	$10^{14.0}$ $M_\odot$	7.220×10^{-3}	-5.29	0.13	-5.21	4.45
SZ	$10^{14.5}$ $M_\odot$	8.317×10^{-3}	-4.48	0.11	-4.40	5.38
λ	$10^{13.0}$ $M_\odot$	3.213×10^{-2}	0.30	0.33	-	-
λ	$10^{13.5}$ $M_\odot$	2.050×10^{-2}	0.81	0.30	-	-
λ	$10^{14.0}$ $M_\odot$	3.347×10^{-2}	1.24	0.18	-	-
λ	$10^{14.5}$ $M_\odot$	4.317×10^{-2}	1.70	0.14	-	-

Table 3.3: As Table 3.2, but for $z = 0.5$.

a	$M_{500c}[M_\odot]$	A	μ	σ	$\log_{10} a_t$	α
X-ray	$10^{13.0}$ $M_\odot$	1.923×10^{-3}	41.3	-0.33	41.7	1.91
X-ray	$10^{13.5}$ $M_\odot$	2.977×10^{-3}	42.4	-0.22	42.8	3.11
X-ray	$10^{14.0}$ $M_\odot$	4.380×10^{-3}	43.3	0.15	43.6	2.45
X-ray	$10^{14.5}$ $M_\odot$	4.395×10^{-3}	44.1	-0.16	44.9	6.50
SZ	$10^{13.0}$ $M_\odot$	4.019×10^{-3}	-6.98	0.22	-6.86	2.12
SZ	$10^{13.5}$ $M_\odot$	5.475×10^{-3}	-6.05	0.17	-5.94	3.26
SZ	$10^{14.0}$ $M_\odot$	6.976×10^{-3}	-5.23	0.13	-5.14	4.20
SZ	$10^{14.5}$ $M_\odot$	7.975×10^{-3}	-4.44	0.13	-4.26	5.07
λ	$10^{13.0}$ $M_\odot$	2.506×10^{-2}	0.44	-0.34	-	-
λ	$10^{13.5}$ $M_\odot$	2.540×10^{-2}	0.95	0.23	-	-
λ	$10^{14.0}$ $M_\odot$	3.588×10^{-2}	1.34	0.17	-	-
λ	$10^{14.5}$ $M_\odot$	4.709×10^{-2}	1.79	0.13	-	-

Table 3.4: As Table 3.2, but for $z = 1$.

a	$M_{500c}[M_\odot]$	A	μ	σ	$\log_{10} a_t$	α
X-ray	$10^{13.0}$ $M_\odot$	2.036×10^{-3}	41.8	0.30	42.0	1.99
X-ray	$10^{13.5}$ $M_\odot$	3.218×10^{-3}	42.8	0.20	43.1	2.80
X-ray	$10^{14.0}$ $M_\odot$	4.727×10^{-3}	43.7	0.14	43.9	3.39
X-ray	$10^{14.5}$ $M_\odot$	4.923×10^{-3}	44.4	0.14	44.9	6.50
SZ	$10^{13.0}$ $M_\odot$	4.428×10^{-3}	-6.86	0.20	-6.75	2.38
SZ	$10^{13.5}$ $M_\odot$	5.701×10^{-3}	-5.99	0.16	-5.88	3.44
SZ	$10^{14.0}$ $M_\odot$	7.126×10^{-3}	-5.19	0.13	-5.10	4.49
SZ	$10^{14.5}$ $M_\odot$	8.161×10^{-3}	-4.41	0.12	-4.32	6.12
λ	$10^{13.0}$ $M_\odot$	2.324×10^{-2}	0.52	-0.33	-	-
λ	$10^{13.5}$ $M_\odot$	2.814×10^{-2}	0.99	0.21	-	-
λ	$10^{14.0}$ $M_\odot$	3.922×10^{-2}	1.36	0.15	-	-
λ	$10^{14.5}$ $M_\odot$	4.695×10^{-2}	1.80	0.13	-	-

Table 3.5: As Table 3.2, but for $z = 2$. Note that different from the other tables, we do not include the highest mass bin as there are insufficient halos to characterize the distributions at $z = 2$

a	$M_{500c}[M_\odot]$	A	μ	σ	$\log_{10} a_t$	α
X-ray	$10^{13.0}$ $M_\odot$	2.217×10^{-3}	42.5	-0.27	42.7	2.00
X-ray	$10^{13.5}$ $M_\odot$	3.330×10^{-3}	43.4	0.19	43.7	2.82
X-ray	$10^{14.0}$ $M_\odot$	4.190×10^{-3}	44.3	0.16	44.5	4.45
SZ	$10^{13.0}$ $M_\odot$	4.923×10^{-3}	-6.76	0.19	-6.62	3.11
SZ	$10^{13.5}$ $M_\odot$	6.225×10^{-3}	-5.93	0.16	-5.80	4.39
SZ	$10^{14.0}$ $M_\odot$	8.234×10^{-3}	-5.14	0.13	-5.04	6.94
λ	$10^{13.0}$ $M_\odot$	2.557×10^{-2}	0.54	-0.29	-	-
λ	$10^{13.5}$ $M_\odot$	3.074×10^{-2}	0.97	0.19	-	-
λ	$10^{14.0}$ $M_\odot$	4.414×10^{-2}	1.33	0.14	-	-

3.B Using Compton-Y within R_{500c}

For Compton-Y we use a fiducial aperture of $5R_{500c}$, which is motivated by Planck Collaboration et al. (2016b), but is much larger than the apertures of R_{500c} that we use for X-ray luminosity. In this section we investigate how the results for Compton-Y change if we use the same aperture as for the other observables. Fig. 3.9 shows the distribution of Compton-Y using a R_{500c} aperture, akin to Fig. 3.1. The different colours indicate different 0.1 dex wide mass bins. The dotted lines show the results of fitting a lognormal distribution to each mass bin. The values of the fitted distribution can be found in Table 3.6. Compared to the results for the $5R_{500c}$ aperture, the distributions now longer show a prominent power-law tail towards higher values of Y_{500c}. Furthermore, the distributions no longer overlap. Hence, the number of lower mass haloes that upscatter will be reduced.

In Fig. 3.10 we show the sample mass bias at four redshifts indicated with different colours for a sample selected on Compton-Y within R_{500c}, akin to Fig. 3.6. As expected from the previous figure, the smaller aperture leads to a a nearly mass-selected sample and the sample mass bias is close to zero for all target masses. Only the 5th percentile shows a slight (≈ 5 per cent) bias.

3.C Convergence with numerical resolution and simulation box size

In Fig. 3.11 we show the median (solid), 5th percentile (dashed) and 95th percentile (dotted) of M_{500c} obtained for samples based on different selection cuts for the three different FLAMINGO resolutions, in a $(1\ \mathrm{Gpc})^3$ box at $z = 0.3$. Addi-

Table 3.6: Values for the fitting the functional form $f(x) = A\exp\left[-\frac{(\log_{10}\mu - x)^2}{\sigma^2}\right]$ to the distribution of Compton-Y within R_{500c} for objects in a 0.1 dex width bin around the mass given in the first column. The second column gives the amplitude of the distribution, the third column the mean and the fourth column the scatter.

Mass M_{500c} [$M_\odot$]	Amplitude	$\log_{10}\mu$	σ
10^{13}	4.79×10^{-3}	-7.79	2.12×10^{-1}
$10^{13.5}$	6.86×10^{-3}	-6.59	1.50×10^{-1}
10^{14}	9.95×10^{-3}	-5.60	1.05×10^{-1}
$10^{14.5}$	1.10×10^{-2}	-4.73	9.52×10^{-2}

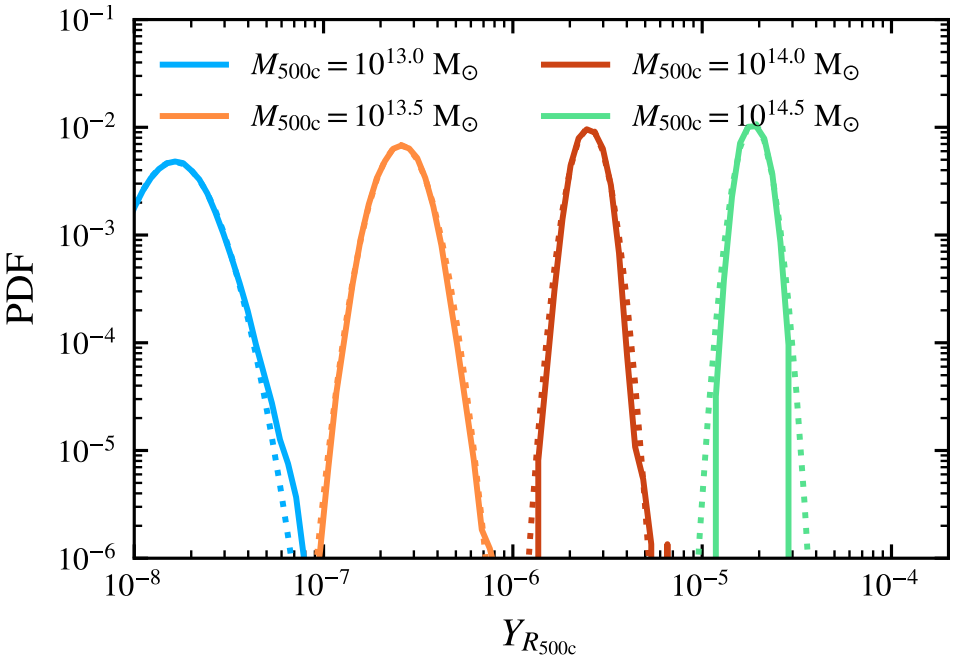


Figure 3.9: As the middle panel of Fig. 3.1, but for an aperture of R_{500c} instead of $5R_{500c}$. The shape of the distribution is very close to lognormal for every mass bin.

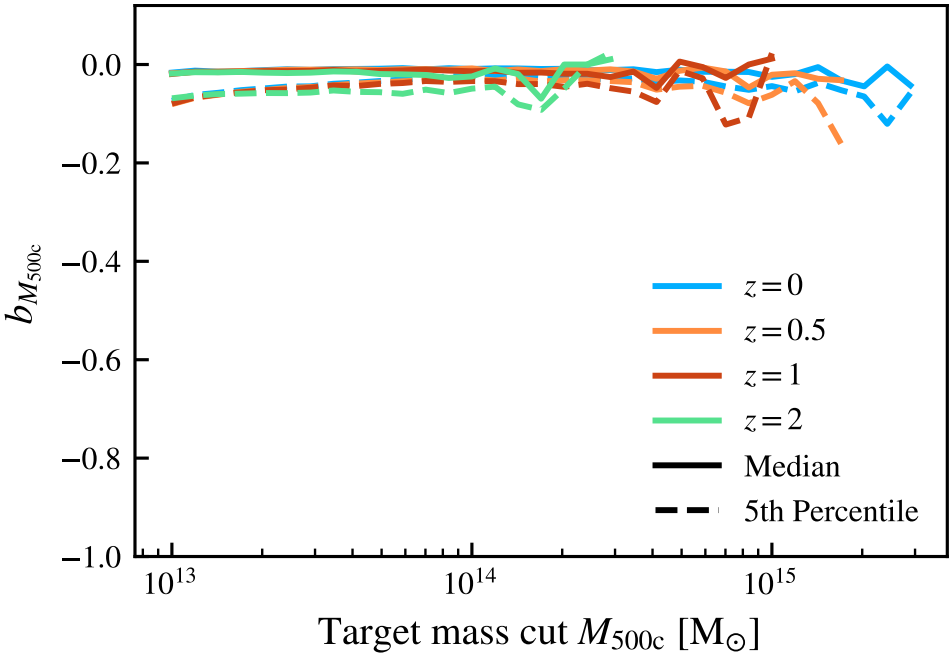


Figure 3.10: As the middle panel of Fig. 3.6, but for an aperture of R_{500c} instead of $5R_{500c}$. The sample mass bias is close to zero across all masses and redshifts.

tionally, we show the results for the $(2.8\ \mathrm{Gpc})^3$ box at intermediate resolution, see Schaye et al. (2023) for the naming convention. Comparing L1_m9 and L2p8_m9 we find converged results for all but the largest halo masses, for which the sampling is much better in L2p8_m9. The only box size effect is due to the improved statistics in a larger volume. For the SZ selection (middle panel) all percentiles are converged for all resolutions.

For the X-ray luminosity selection (top panel) the median and 95th percentile are very close to being converged, with only the lowest resolution (m10) run decreasing slightly at the lowest masses. The largest difference is found for the 5th percentile. While the dip remains at roughly the same mass across the three resolutions, the 5th percentile drops more towards with higher resolution. This implies that the existence of the dip is not directly due to resolution effects, and could be caused due to an increase in scatter for haloes with this luminosity. The fact that the dip gets deeper with increasing resolution implies that at m9 resolution we do not yet resolve the full range of haloes that can up-scatter in our selection for the lowest luminosities. For our fiducial resolution (m9) the median is converged over the full target mass range and the 5th percentile is converged for target mass cuts of $M_{500c} \gtrsim 10^{14}\,\mathrm{M}_\odot$.

For the richness selection we make use of both a cut in stellar mass and radius.

As the stellar mass - halo mass relation is not converged at the high-mass end, see Fig. 9 from Schaye et al. (2023), it is not surprising that richness is not converged either. If we make a cut using the bound subhalo mass of the satellites instead of stellar mass, and pick a subhalo mass limit of 2×10^{11} $M_\odot$, which selects haloes close to the stellar mass limit of $10^{10.046}$ $M_\odot$, then the m9 and m8 simulations do agree, as shown in Fig. 3.12. The subhalo mass cut is too low for the low resolution (m10) simulation to be converged, but the other two resolution simulations are converged if we select on subhalo richness. As the subhaloes are converged between m9 and m8, the differences we see for galaxy richness in the bottom panel of Figure 3.11 are caused by the differences in the stellar mass - halo mass relation. Both simulations match the stellar mass function up to $M_* = 10^{11.5}\,M_\odot$, so the fact that the richness selection is not fully converged is due to a combination of differences in the satellite fractions and imperfect calibration of the galaxy stellar mass function.

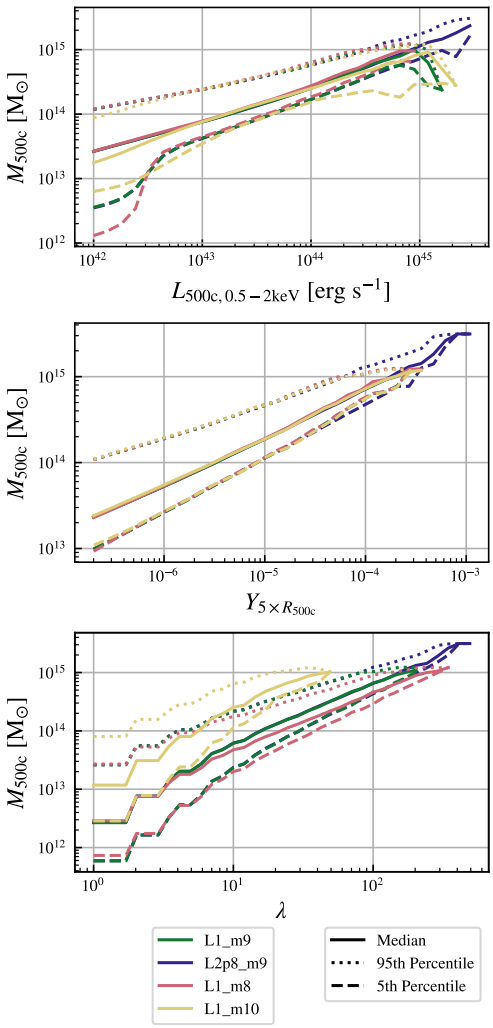


Figure 3.11: The solid lines show the median M_{500c} for a sample selected using the cut on the quantity plotted along the x-axis. The different panels show the three different selection quantities, X-ray luminosity (top), thermal SZ Compton-Y (middle) and galaxy richness (bottom). The different colours show the results for the three different FLAMINGO resolutions and for the 1 and 2.8 $(\mathrm{Gpc})^3$ m9 boxes. The dashed (dotted) line indicates the 5th (95th) percentile mass for the sample after the cut. The mass distributions are converged with resolution for X-ray and SZ selection, but not for selection based on richness.

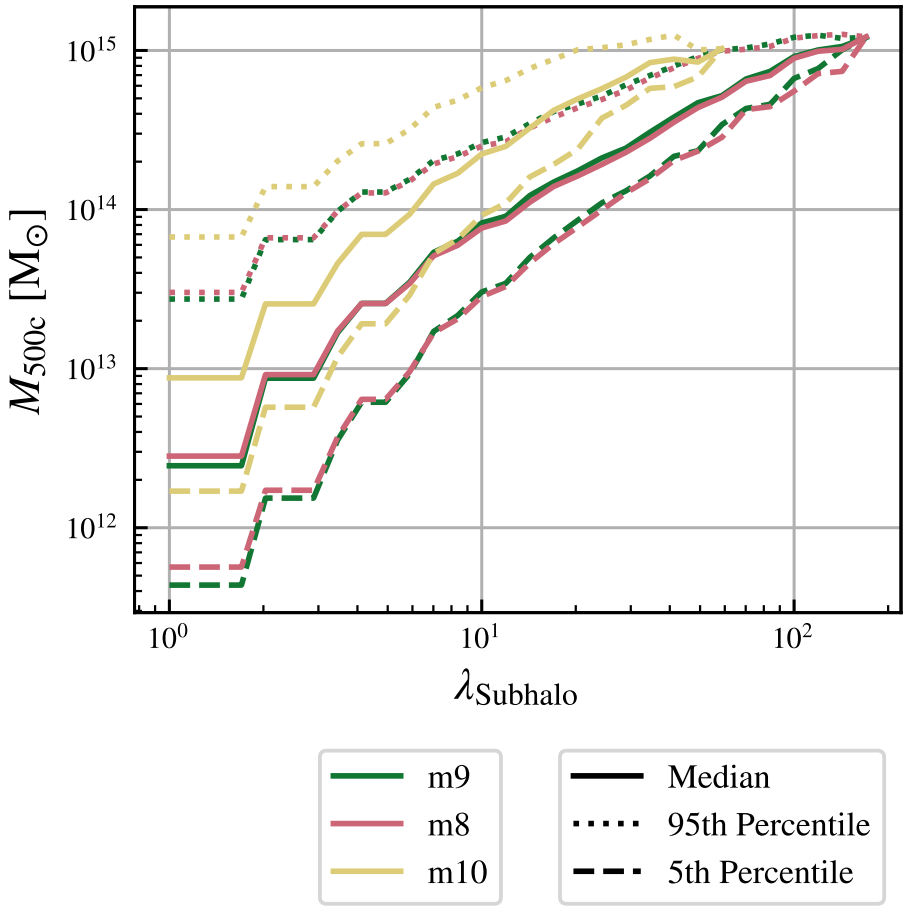


Figure 3.12: The solid lines show the median M_{500c} for a sample selected using a cut on subhalo richness, defined using satellite subhaloes with a bound mass above 2×10^{11} $M_\odot$, and plotted along the horizontal axis. The different colours show the results for the three different FLAMINGO resolutions. The dashed (dotted) line indicates the 5th (95th) percentile mass for the sample after the cut. The halo mass distributions for samples selected by cuts on subhalo richness are converged for m8 and m9, but not for m10.

4 The FLAMINGO Project: An assessment of the systematic errors in the predictions of models for galaxy cluster counts used to infer cosmological parameters

Authors: **Roi Kugel**, Joop Schaye, Matthieu Schaller, Victor J. Forouhar Moreno, Robert J. McGibbon
Submitted to MNRAS

Abstract

Galaxy cluster counts have historically been important for the measurement of cosmological parameters and upcoming surveys will greatly reduce the statistical errors. To exploit the potential of current and future cluster surveys, theoretical uncertainties on the predicted abundance must be smaller than the statistical errors. Models used to predict cluster counts typically combine a model for the dark matter only (DMO) halo mass function (HMF) with an observable -mass relation that is assumed to be a power-law with lognormal scatter. We use the FLAMINGO suite of cosmological hydrodynamical simulations to quantify the biases in the cluster counts and cosmological parameters resulting from the different ingredients of conventional models. For the observable mass proxy we focus on the Compton-Y parameter quantifying the thermal Sunyaev-Zel'dovich effect, which is expected to result in cluster samples that are relatively close to mass-selected samples. We construct three mock surveys based on existing (Planck and SPT) and upcoming (Simons Observatory) surveys. We ignore measurement uncertainties and compare the biases in the counts and inferred cosmological parameters to each survey's Poisson errors. We find that widely used models for the DMO HMF differ significantly from each other and from the DMO version of FLAMINGO, leading to significant biases for all three surveys. For upcoming surveys, dramatic improvements are needed for all additional model ingredients, i.e. the functional forms of the fits to the observable-mass scaling relation and the associated scatter, the priors on the scaling relation and the prior on baryonic effects associated with feedback processes on the HMF.

4.1 Introduction

The standard model of cosmology, ΛCDM, consisting of a universe filled mostly by dark energy and cold dark matter, has so far been successful at describing

observations, but cracks may have started to appear. When comparing the values of cosmological parameters inferred from the cosmic microwave background (Planck Collaboration et al., 2020a) with the same values obtained from local measurements (e.g. Riess et al., 2022; Heymans et al., 2021; Abbott et al., 2022) tensions have started to pop up. These tensions exist in particular for the Hubble constant, H_0, and for the clustering parameter σ_8. With the advent of surveys like Euclid (Euclid Collaboration et al., 2024) and LSST (Ivezić et al., 2019), we are close to getting a statistically robust measurement of σ_8. It it thus imperative for models assuming ΛCDM to have high accuracy and precision, while keeping the door open to potential extensions.

The tension in σ_8 is currently at the $2-3\sigma$ level (Heymans et al., 2021; Abbott et al., 2022; Miyatake et al., 2023; McCarthy et al., 2023). Surveys like Euclid and LSST will provide us with much tighter constraints based on weak lensing and galaxy clustering. Another promising avenue is to approach the tension from many different directions. In additions to lensing, we can, for example, use the Sunyaev-Zel'dovich (SZ) effect power spectrum (e.g. Sunyaev & Zeldovich, 1972; Planck Collaboration et al., 2014b; Bleem et al., 2022), cluster counts (e.g. Planck Collaboration et al., 2016a; Bocquet et al., 2024; Ghirardini et al., 2024), CMB-lensing (e.g. Planck Collaboration et al., 2020b) and all their cross-correlations to constrain cosmology. These independent probes are not only a way to increase our constraints on σ_8, but, if the tensions persist, they can also provide valuable insight into how to extend the ΛCDM model.

Here we will investigate cosmological constraints via cluster count cosmology (for reviews see Allen et al., 2011; Pratt et al., 2019). The number of halos as a function of halo mass is described by the halo mass function (HMF), which is sensitive to the cosmological parameters. In particular, the high mass end, where clusters are found, is very sensitive to changes in σ_8. Because of this sensitivity, even a statistic as simple as counting the number of observed clusters per unit volume can provide valuable constraints on σ_8 and, to a lesser extent, Ω_m, the matter content of the Universe.

In order to count clusters above some mass limit, we need an observable proxy for halo mass with which we can select clusters. As clusters are very massive objects, they can be observed across almost the entire electromagnetic spectrum. The three most favoured methods use the X-ray emission from the hot intra-cluster medium (e.g. Pacaud et al., 2018; Liu et al., 2022, 2024), the number of member galaxies in the optical (e.g. Black & Evrard, 2022; Artis et al., 2022), and the distortion of the CMB spectrum due to their high electron pressures via the SZ effect (e.g. Planck Collaboration et al., 2016b; Hilton et al., 2018; Bleem et al., 2024). Additionally, methods using the cluster lensing signal are starting to be used (Hamana et al., 2020; Chen et al., 2024). Kugel et al. (2024) use the FLAMINGO cosmological simulations (Schaye et al., 2023; Kugel et al., 2023) to compare the performance of the different observables and found that SZ-

selection is generally less prone to selection effects than X-ray or galaxy richness selection (using a window of size $5R_{500c}$[1], as appropriate for the Planck satellite).

In order to measure cosmological parameters from cluster counts, we need accurate predictions for the expected number of clusters. The standard approach combines a theoretical HMF with a scaling relation between halo mass and an observable mass proxy with an assumed level of scatter around it. The scaling relation is further assumed to be a power-law with lognormal scatter with a fixed σ across the mass range (e.g. Rozo et al., 2014; Evrard et al., 2014; Planck Collaboration et al., 2016a; Bocquet et al., 2019; Costanzi et al., 2019; Chiu et al., 2023; Ghirardini et al., 2024). HMF models typically do not account for baryonic effects or do so assuming a particular galaxy formation model. These assumptions might induce biases in the analysis because it is unlikely that scaling relations are perfect power-laws with mass-independent and lognormal scatter (Kugel et al., 2024) and because poorly constrained astrophysical processes associated with galaxy formation are expected to modify the halo masses (e.g. Velliscig et al., 2014; Cui et al., 2014; Cusworth et al., 2014; Bocquet et al., 2016; Schaye et al., 2023) and to bias the total masses measured using weak gravitational lensing (Debackere et al., 2021). Getting a good grip on these assumptions and the associated systematic uncertainties is key in making sure we can derive unbiased cosmology results (Angulo et al., 2012; Mantz, 2019).

Compared with cosmic shear and galaxy clustering measurements, cluster counts are less constraining for cosmological parameters. However, thanks to ongoing and upcoming X-ray (e.g. Liu et al., 2024), optical (e.g. Artis et al., 2022) and SZ surveys (e.g. Hilton et al., 2018; Ade et al., 2019; Klein et al., 2024; Bleem et al., 2024) it is likely that the constraints will tighten in the near future. The main goal of this work is to investigate if and how standard models and assumptions used for cluster cosmology affect the cosmological inference from cluster counts.

The models of the HMF are usually based on cosmological dark matter only (DMO) simulations. One of the difficulties here is that to accurately predict galaxy clusters, simulations with a very large box size are required. Clusters are very rare, and a large enough sample is needed to make statistically robust predictions. In order to model the HMF different methods are used. A common method is the one described by Jenkins et al. (2001), where an empirical formula is fitted to DMO simulations and is then used to predict the HMF directly from the matter power spectrum. This is the method employed by for example Tinker et al. (2008, 2010), Bocquet et al. (2016) and Euclid Collaboration et al. (2023). Additionally, emulators have also started to be used for the HMF (e.g. Bocquet et al., 2020). In this case the HMF is obtained by interpolating between a set of DMO training simulations of different cosmologies.

[1] R_{500c} is the radius around a halo within which the enclosed density is 500 times the critical density. The mass within this radius is defined as M_{500c}.

These standard models are mostly based on DMO simulations. However, it is known that the HMF is modified by baryonic effects. The effects are strongest for clusters with mass $M_{500c} \lesssim 10^{14} M_\odot$ (e.g. Velliscig et al., 2014; Bocquet et al., 2016; Schaye et al., 2023). In addition to a DMO HMF, Bocquet et al. (2016) use the HMF obtained from the Magneticum hydrodynamical simulation (Hirschmann et al., 2014) to constrain the fitting parameters directly to the HMF predicted by the hydro simulation. If baryonic effects are not taken into account correctly, the analysis is inconsistent, as the scaling relations are based on the true halo masses measured using X-rays or weak lensing rather than from DMO simulations (e.g. Planck Collaboration et al., 2014a; Lovisari et al., 2015; Akino et al., 2022).

Baryons need to be considered to predict observable mass proxies other than lensing. Feedback processes associated with galaxy formation change the distribution of the gas, which will change the halo mass function and modify the functional form of, and the scatter in the mass -observable relation. Hydrodynamic simulations make direct predictions for both the number of clusters and the observable mass proxies. However, most state-of-the-art simulations, such as EAGLE (Schaye et al., 2015) and IllustrisTNG (Pillepich et al., 2018), do not have volumes large enough to provide a statistical sample of clusters. For cluster counts, doing zooms of individual objects (e.g. Barnes et al., 2017) is not feasible as we need volume-limited statistics. Although projects like The Three Hundred Project (Cui et al., 2018) and TNG-cluster (Nelson et al., 2023) can use zooms to simulate a hundreds of clusters from a single volume, they are unable to provide volume-limited samples of clusters. For cluster counts it is necessary to use large volumes even if this means lowering the resolution relative to simulations focusing on galaxy evolution. This is the approach taken in projects like CosmoOWLS (Le Brun et al., 2014), BAHAMAS (McCarthy et al., 2017) and MilleniumTNG (Pakmor et al., 2023). In addition, it is necessary to vary the uncertain strength of feedback processes associated with galactic winds driven by star formation and particularly by Active Galactic Nuclei (AGN) and to do so in a manner constrained by observations relevant for cluster cosmology (Le Brun et al., 2014; McCarthy et al., 2017).

For this work, we will make use of the cosmological hydrodynamic simulations of the FLAMINGO project (Schaye et al., 2023; Kugel et al., 2023). FLAMINGO is a suite of very large volume simulations designed specifically to investigate the interplay between effects due to baryonic processes, massive neutrinos, and cosmology. The simulation suite includes a hydrodynamic simulation in a $(2.8\ \mathrm{Gpc})^3$ volume using $(5040)^3$ gas particles and many model variations in $(1.0\ \mathrm{Gpc})^3$ volumes. The largest volume, contains 461 (4100) clusters of mass $M_{500c} > 10^{15}\ M_\odot$ ($5 \times 10^{14}\ M_\odot$) at $z = 0$. The FLAMINGO variations include variations in feedback, cosmology and resolution. Due to its large volumes, FLAMINGO provides statistically significant samples up to very high halo

masses. The subgrid stellar and AGN feedback in the simulation have been calibrated to match the observed low-redshift galaxy stellar mass function and cluster gas fractions. The simulation have been shown to provide provide a good match to X-ray observations of clusters at the profile level (Braspenning et al., 2023). FLAMINGO was calibrated using machine learning assisted Gaussian process emulation (Kugel et al., 2023), and the feedback variations can each be related to systematic shifts in the cluster gas fractions and/or the galaxy mass function. In addition, the suite includes simulations that use an alternative AGN feedback prescription that uses kinetic jets instead of thermally driven winds, which is calibrated to match the same observations as the thermal models.

For this work we assume that the predictions from the FLAMINGO simulation are the ground truth that we compare our models against. In Kugel et al. (2024) we found that, in FLAMINGO, the scatter around the mass-observable scaling relation for X-ray luminosity within R_{500c} and for Compton-Y within $5R_{500c}$ has power-law tails. However, we also showed that for Compton-Y within R_{500c}, which is the aperture we use here, the scatter is close to lognormal, but the scatter still changes with mass and redshift. Additionally, the scaling relations deviate from single power-laws and change between different feedback variations (Schaye et al., 2023; Braspenning et al., 2023). We will investigate how deviations from assumed scaling relations and scatter influence cluster counts and the cosmology inferred from them. We choose to focus on SZ-selected samples since those are intrinsically the least biased compared to a mass-selected sample (Kugel et al., 2024). The systematic errors that we report here are thus likely smaller than for other selection methods. In this work we do not attempt to forward model the selection effects using virtual observations. In the observations, SZ-selection is accomplished by applying a matched filter to CMB maps at different frequencies (see e.g. Melin et al., 2006, 2012; Hilton et al., 2018). At the higher resolutions of current and upcoming CMB surveys, this will introduce additional selection effects via, for example, source confusion, foregrounds (see e.g. Melin et al., 2018; Zubeldia et al., 2023), and the effect of beam smearing. In this work we neglect these effects and leave them for a future study, noting that the data products needed for such a work are available within the FLAMINGO suite of simulations.

The paper is structured as follows. In Section 4.2 we describe the FLAMINGO simulations, the models and functional forms used to predict cluster counts, and how we use those predictions to constrain cosmological parameters. In Section 4.3 we show how different assumptions and changes due to the variations in the FLAMINGO simulation suite affect the predictions for cluster counts and the inferred cosmology. We summarise the results in Section 4.4.

4.2 Methods

In this section we describe the methods we use to predict cluster counts using the FLAMINGO simulations. We briefly introduce the FLAMINGO simulation and how we identify haloes in §4.2.1. We describe our predictive model for cluster counts in §4.2.2. The three sample definitions used throughout this paper are introduced in §4.2.3. We finish this section by describing the likelihood and how it is sampled in §4.2.4.

4.2.1 FLAMINGO

To construct virtual cluster catalogues we make use of the HMF, observable-mass scaling relations and their scatter obtained from the FLAMINGO simulations (Schaye et al., 2023; Kugel et al., 2023). We will make use of the simulations at FLAMINGO's intermediate resolution ($m_{\rm gas} = 1.09 \times 10^9$ M$_\odot$) in box sizes of (1 Gpc)^3. These simulations use 2×1800^3 gas and dark matter particles, 1000^3 neutrino particles, and all assume the Dark Energy Survey year 3 (Abbott et al., 2022) cosmology ($\Omega_{\rm m} = 0.306$, $\Omega_{\rm b} = 0.0486$, $\sigma_8 = 0.807$, $\rm H_0 = 68.1$ km/s/Mpc, $\rm n_s = 0.967$). The simulations are run with the cosmological smooth particle hydrodynamics and gravity code SWIFT (Schaller et al., 2024) using the SPHENIX SPH scheme (Borrow et al., 2022). The initial conditions are obtained from a modified version of MONOFONIC (Hahn et al., 2021; Elbers et al., 2022), and neutrinos are implemented with the δf method (Elbers et al., 2021).

The FLAMINGO subgrid physics is an evolution the models developed for OWLS (Schaye et al., 2010) and used in BAHAMAS (McCarthy et al., 2017). It includes radiative cooling (Ploeckinger & Schaye, 2020), star formation (Schaye & Dalla Vecchia, 2008), stellar mass loss (Wiersma et al., 2009), kinetic supernova feedback (Dalla Vecchia & Schaye, 2008; Chaikin et al., 2022b,a), super massive black holes (Springel et al., 2005; Bahé et al., 2022) and AGN feedback in both thermal (Booth & Schaye, 2009) and kinetic form (Huško et al., 2022). One of the new features is that the subgrid physics was calibrated by fitting the simulations to match the $z = 0$ galaxy stellar mass function (in the mass range 10^{10} M$_\odot \lesssim$ M$_* \lesssim 10^{11.5}$ M$_\odot$) and the gas fractions in low-z groups and clusters (up to a mass of $M_{\rm 500c} = 10^{14.3}$ M$_\odot$) using machine learning (Kugel et al., 2023). The same method was used to design a set of variations which are directly based on the observed error bars. Of particular interest for this work are the variations in the cluster gas fraction and AGN model, which are denoted as fgas_$\pm N\sigma$ and Jet_fgas_$\pm N\sigma$. For these models the $N\sigma$ denotes by how many observed standard deviations the gas fractions have been shifted up or down with respect to the fiducial model. The Jet models make use of kinetic jets for the AGN feedback instead of the thermal model used for all other runs. These Jet models are calibrated to match the same data as the corresponding thermal AGN feedback models. More details can be

found in Schaye et al. (2023).

To obtain catalogues of (sub)haloes from the simulation, we make use of a modified version of Hierarchical Bound Tracing+ (HBT+) (Han et al., 2012, 2017, Forouhar Moreno et al. in prep). The modifications include required extensions required for the application to hydrodynamic simulations. HBT+ tracks haloes trough "history space". After a halo has been identified as a friends-of-friends group, HBT+ tracks the particles of the halo trough cosmic time, and converts it to a satellite when it becomes part of another friends-of-friends group. For each central halo in the HBT+ catalogue we use the Spherical Overdensity and Aperture Processor (SOAP[2], McGibbon et al. in prep) to calculate M_{500c} and the integrated Compton-Y. The Compton-Y contribution of each particle is stored in the snapshots and is calculated via

$$y_i = \frac{\sigma_T}{m_e c^2} n_{e,i} k_B T_{e,i} \frac{m_i}{\rho_i}, \tag{4.1}$$

where σ_T is the Thomson cross section, m_e is the electron mass, c is the speed of light, k_B is the Boltzmann constant, $n_{e,i}$ is the electron number density, $T_{e,i}$ is the electron temperature, m_i is the mass and ρ_i is the density of the particle with index i. The contribution for each halo can be found by summing the contribution of all particles within an aperture. In this work we use the integrated Compton-Y within R_{500c}.

4.2.2 The cluster count model

There are three key ingredients to predict the number of clusters that are observed in a cluster survey:

1. The number of halos per unit mass and volume via the HMF.

2. The probability for a halo with a given mass to have a certain observable property via a median scaling relation and the associated scatter.

3. The probability for a halo with a certain observable property to be observed in the survey via the selection function.

Our primary goal is to investigate the effect of the different assumptions that are made for the first two of these items. Because we focus on the deviations on the theory side, we omit from our analysis the conversion between the observable and the signal measured at the telescope. Using these assumptions, we can predict the

[2] https://github.com/SWIFTSIM/SOAP

number of clusters using the following integral

$$N(X_{\rm C}) = \int_0^{z_{\rm max}} \int_0^{A_{\rm sky}} \int_{M_{\rm 500c,min}}^{M_{\rm 500c,max}} \phi(M_{\rm 500c}, z, \boldsymbol{\theta}, a_{\rm Astro}) \times \chi(M_{\rm 500c}, z, X_{\rm C}) \frac{{\rm d}V}{{\rm d}\Omega {\rm d}z}(\boldsymbol{\theta}) {\rm d}M_{\rm 500c} {\rm d}\Omega {\rm d}z, \tag{4.2}$$

where $\phi(M_{\rm 500c}, \boldsymbol{\theta}, a_{\rm Astro}, z)$ is the HMF, which explicitly depends on the astrophysical parameters $a_{\rm Astro}$, like the gas fraction on clusters and the assumed AGN model, $\chi(M_{\rm 500c}, z, X_{\rm C})$ gives the probability of observing a cluster with mass $M_{\rm 500c}$ at redshift z given a cut (i.e. selection limit) $X_{\rm C}$ on observable X, $A_{\rm sky}$ is the survey angular area on the sky, and $\frac{{\rm d}V}{{\rm d}\Omega {\rm d}z}(\boldsymbol{\theta})$ is the differential comoving volume, which, like the HMF, depends on the cosmological parameters $\boldsymbol{\theta}$. We note here that in the integral $M_{\rm 500c}$ refers to the true total mass, including baryonic effects. We integrate from redshift zero to $z_{\rm max} = 2$, noting that going to higher redshifts makes a negligible difference for the cuts we are interested in. For $M_{\rm 500c}$ we integrate from 10^{12} $M_\odot$ to 10^{16} $M_\odot$, noting that we run out of halos before we reach the maximum mass. The minimum mass of 10^{12} $M_\odot$ ensures we do not miss any halos for the observational selection cuts we apply. We are specifically interested in selecting clusters via their Compton-Y contribution within $R_{\rm 500c}$. In this case X becomes $Y_{\rm 500c}$, which we will denote simply as Y, with the cut being defined as $Y_{\rm C}$.

For the selection function we use a simple step function

$$P(Y|Y_{\rm C}) = \begin{cases} 1 & \text{if } Y > Y_{\rm C}, \\ 0 & \text{if } Y < Y_{\rm C}. \end{cases} \tag{4.3}$$

Since we are interested in the effect of theoretical uncertainties, this selection function gives us results that are easily interpretable. To obtain $\chi(M_{\rm 500c}, z, X_{\rm C})$ we use the mass observable scaling relation $P(Y|\hat{Y}(M_{\rm 500c}, z))$. In this case $\hat{Y}$ is the median value of Y at a given $M_{\rm 500c}$ and z as given by the scaling relation.

We consider four cases designed to separate the effects of two of the different model ingredients: the power-law observable-mass scaling relation and the associated scatter:

1. Taking both the scaling relation and scatter from the simulation.
2. Using a power-law scaling relation with lognormal scatter, both fit to the simulation results, referred to as "PL+LN".
3. Using the scaling relation from the simulation with lognormal scatter fit to the simulation results, referred to as "LN".
4. Using a power-law scaling relation fit to the simulation results with scatter from the simulation, referred to as "PL".

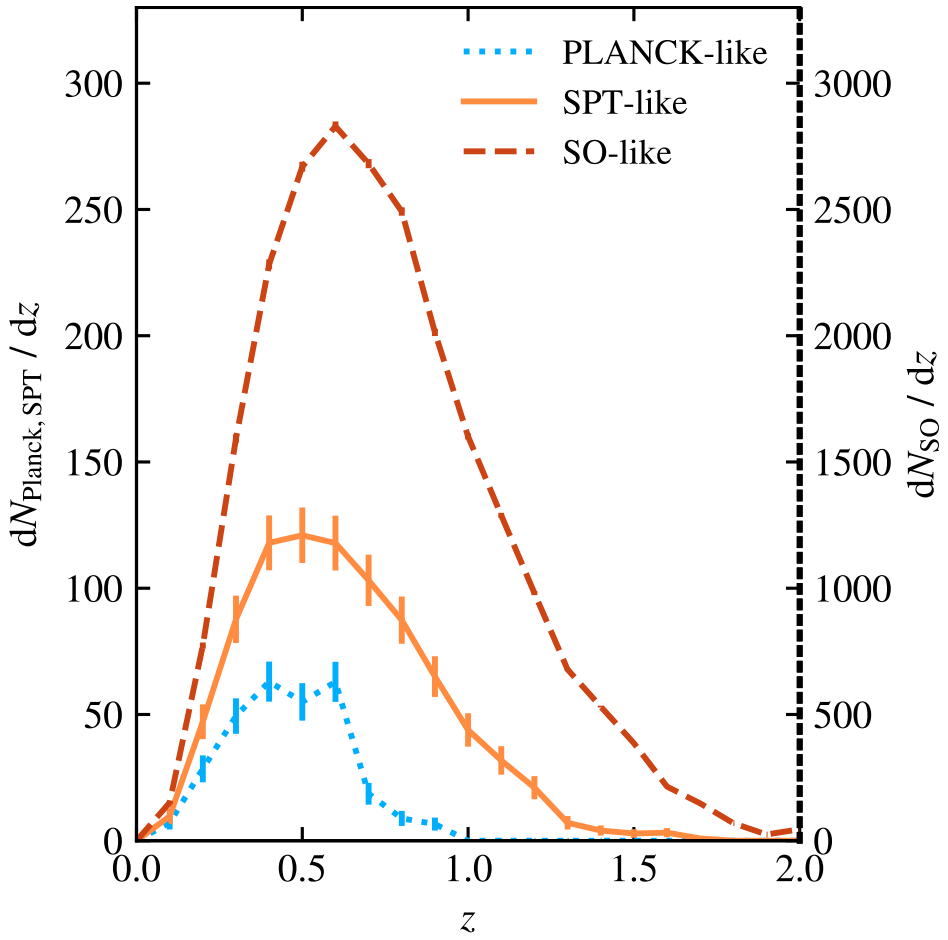


Figure 4.1: The redshift distribution of cluster number counts for our fiducial (L5p6DMO HMF and L1_m9 $Y - M_{500c}$ scaling relation and scatter) model. The three different lines show the predictions for the three mock survey sizes we investigate in this paper. The error bars show the Poisson errors. Note that the PLANCK- and SPT-like surveys use the left y-axis, while the SO-like survey uses the right y-axis. Note that the SO-like errors are smaller than the line width in this figure.

By comparing the four cases we can assess potential systematics due to conventional assumptions. When describing the changes between these four cases, it is important to remember that the quantity we need is $\chi(M_{500c}, z, X_C)$, the probability of observing an object with a given mass and redshift given a selection cut.

In the case where we use both the scaling relation and the scatter from the

simulation, we use only the binned outputs of the simulation. To obtain the scaling relation from the simulations we set up a grid in M_{500c} and z. We sample the mass in bins of width 0.1 dex between $M_{500c} = 10^{12}$ $M_\odot$ and 10^{16} $M_\odot$, and we sample the redshift in bins of 0.05 from $z = 0$ to $z = 3$. In each bin we calculate the number density of objects and the median value for Y. Additionally, at each mass and redshift we use 800 bins between a Y of 10^{-10} and 10^{-2} to sample the distribution of Y for each mass and redshift bin. To obtain the probability of observing a halo in this mass-redshift bin, we want to find the fraction of halos that have $Y > Y_C$. For our selection function, this is equivalent to

$$\begin{aligned}\chi(M_{500c}, z, X_C) &= \int_{-\infty}^{\infty} P(Y|\hat{Y}(M_{500c}, z))P(Y|Y_C), \\ &= 1 - C_{M_{500c},z}(Y_C),\end{aligned} \tag{4.4}$$

where $C_{M_{500c},z}(Y_C)$ is the normalised cumulative distribution function (CDF) at fixed mass and redshift. For each mass-redshift bin we create the CDF by interpolating the 800 Y bins, giving us a smooth function. As the simulation binning is already a combination of the scatter and scaling relation, we do not need to take any additional steps when the goal is using both the scaling relation and scatter from the simulations.

In the case of using a power-law scaling relation with lognormal scatter, we instead assume a functional form to obtain the required probabilities. The functional form we use was used by Planck Collaboration et al. (2014a, 2016a) and Salvati et al. (2022) and is given by

$$E^{-\beta}(z)\left[\frac{D_A^2(z)\hat{Y}}{10^{-4}\,\mathrm{Mpc}^2}\right] = Y_*\left[\frac{h}{0.7}\right]^{-2+\alpha}\left[\frac{(1-b)M_{500c}}{6\times10^{14}\,M_\odot}\right]^{\alpha}, \tag{4.5}$$

where the fitting parameters Y_*, the power-law amplitude, α, the power-law slope, β, the slope of the redshift scaling, and b, the hydrostatic bias, are fit to the results from the simulation. Any change is thus purely due to slight deviations from a power-law in the simulation. The values and priors can be found in Table 4.1.

For lognormal scatter we can compute the CDF via

$$1 - C_{M_{500c},z}(Y_C) = \frac{1}{2}\left[1 - \mathrm{erf}\left(\frac{\log_{10} Y_C - \log_{10}\hat{Y}(M_{500c}, z)}{\sqrt{2}\sigma}\right)\right], \tag{4.6}$$

where σ is the standard deviation parameterising the lognormal scatter in Y in dex and $\hat{Y}(M_{500c}, z)$ is the median. The value of σ is taken from Planck Collaboration et al. (2014a, 2016a). Note that the implementation of lognormal scatter has the scaling relation as a direct input. So we can use both the power-law from Eq. 4.5 or directly use the simulation prediction, covering cases two and three.

Our final case is using a power-law for the scaling relation, with scatter from the simulation. In this case we again make use of the binned interpolators that

Table 4.1: Priors on the cosmological and power-law fit parameters (Eq. 4.5) taken from the MiraTitanEmulator and Planck Collaboration et al. (2016a) respectively and the fiducial FLAMINGO values. Here $\mathcal{U}(x,y)$ stands for a uniform distribution between x and y, $\mathcal{N}(\mu,\sigma)$ stands for a normal distribution with mean μ and standard deviation σ. The last two parameters, β and b, are fixed, though we do give the uncertainty on the parameter β from Planck Collaboration et al. (2016a) for context.

Parameter	Prior	FLAMINGO
Ω_m	$\mathcal{U}(0.259, 0.334)$	0.306
σ_8	$\mathcal{U}(0.7, 0.9)$	0.807
$\log_{10} Y_*$	$\mathcal{N}(-0.19, 0.02)$	−0.098
α	$\mathcal{N}(1.79, 0.08)$	1.66
$\sigma_{\log_{10}}$	$\mathcal{N}(0.075, 0.01)$	0.081
β	0.66 (±0.50)	0.89
b	0.743	0.743

are used for the simulation only case. To conserve the scatter, we calculate for each mass and redshift bin what the difference is between the power-law scaling relation fit to the simulation prediction and the actual simulation scaling relation. We then recenter the distribution in each mass-redshift bin on the power-law predictions. This conserves the shape of the scatter in the simulation, but the median now follows the power-law scaling relation.

To obtain the HMF from the simulation, we use the same binning approach the get the number density of objects at each mass and redshift. However, in order to be able to vary and fit the cosmology, we need to know how the HMF changes with cosmology. We investigate the accuracy of the available models for the HMF by Tinker et al. (2010), Bocquet et al. (2016) and Bocquet et al. (2020, which we will refer to as the MiraTitanEmulator). They can be divided into two categories. The models by Tinker et al. (2010) and Bocquet et al. (2016) are based on the fitting formula introduced by Jenkins et al. (2001) and they derive the HMF directly from the linear theory matter power spectrum. The MiraTitanEmulator is an emulator that was trained on the MiraTitan set of DMO simulations (Heitmann et al., 2016). Tinker et al. (2010) and Bocquet et al. (2016) directly provide the HMF based on M_{500c}. For the MiraTitanEmulator, we convert from M_{200c} to M_{500c} assuming an NFW density profile and using the mass-concentration relation by Diemer & Joyce (2019).

All these HMFs model the DMO HMF. To investigate the effects of baryons on the HMF, we allow for the following correction

$$\phi(M_{500c,\mathrm{Hydro}}, z) = \phi_\mathrm{DMO}(M_{500c,\mathrm{DMO}}, z) F_\mathrm{cor,bar}(M_{500c,\mathrm{DMO}}, z), \tag{4.7}$$

where $F_\mathrm{cor,bar}(M_{500c}, z)$ gives the correction factor to convert from a DMO HMF

to the HMF affected by baryons. We measure the correction factor from the FLAMINGO simulations using the same binning for each HMF as described earlier. If either the DMO or hydro HMF has no halos in a bin, we set the ratio to one. We note here that not using this correction causes an inconsistency in the calculation as the scaling relations are always based on the hydrodynamic masses.

As we have continuous analytic alternatives for each of the ingredients of the cluster count model, we have verified that our choice of binning (0.1 dex in M_{500c} and steps of 0.05 in z) does not lead to significant numerical errors. We evaluate our continuous assumptions on the same grid of inputs as the simulations. This allows for fair comparisons between the simulation and analytic predictions and allows for a significant speedup of the code.

To ensure that the number counts vary continuously as a function of redshift, which is necessary as we want to be able to deviate from the simulation snapshot redshifts when defining our redshift bins, we first perform the following integral

$$\frac{\mathrm{d}N(Y > Y_C)}{\mathrm{d}z}(z) = \int_0^{A_{\rm sky}} \int_{M_{500c,\rm min}}^{M_{500c,\rm max}} \phi(M_{500c}, z, \boldsymbol{\theta}, a_{\rm Astro}) \times \chi(M_{500c}, z, Y_C) \mathrm{d}M_{500c} \mathrm{d}\Omega, \tag{4.8}$$

resulting in a redshift density distribution. We use linear interpolation to obtain a continuous version of the distribution. We then numerically integrate this continuous function for each bin using a quadratic integrator. We have verified that these choices do not lead to significant numerical errors.

4.2.3 Definition of samples

We will use three fiducial sample definitions, based very loosely on available SZ cluster samples. We include a shallow full-sky sample, using $Y_C = 10^{-4}$ Mpc^{-2}. This has a total of about 360 clusters for our fiducial simulation, which is similar to the number of objects in the Planck survey (Planck Collaboration et al., 2016a). We also include a 5000 deg^2 survey with a cut of $Y_C = 3 \times 10^{-5}$ Mpc^{-2}. This leads to about 1000 objects, making it similar to the current SPT survey (Bocquet et al., 2024). As a reference for future observations, we also include a Simons Observatory (SO)-like survey. This survey covers 40% of the sky with a cut of $Y_C = 10^{-5}$ Mpc^{-2}. This results in about 27000 objects, in line with SO forecasts (Ade et al., 2019). To create model data we sample the distribution at every 0.1 in redshift (i.e. a coarser redshift sampling than used for the simulations) between $z = 0$ and $z = 2$, resulting in 19 bins. For each of these samples we will assume the observable mass scaling relation from our fiducial (1 Gpc)^3 volume, L1_m9, with the HMF from our biggest (5.6 Gpc)^3 DMO volume. This ensures we use our calibrated model for the hydro, while using a large volume to reduce cosmic variance errors on the HMF. This does imply that our fiducial model has the common inconsistency that the scaling relation is based on hydro

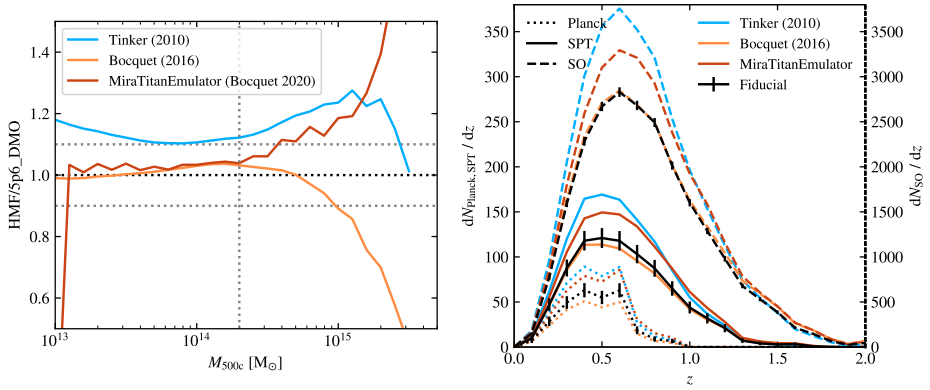


Figure 4.2: The left panel shows the ratio between the three HMF models specified in the legend and the FLAMINGO L5p6_m10_DMO for the FLAMINGO cosmology at $z = 0$. The black dotted horizontal line shows the one-to-one line and the grey dotted horizontal line shows the ten per cent region. The grey dotted vertical line roughly coincides with the lowest cluster masses that will be in the deepest (SO) sample. The right panel shows the distribution of clusters with redshift for the three different surveys, given by different line-styles. The different coloured lines show the cluster counts when assuming different HMF models. The black line with error bars shows the cluster counts from the fiducial model (L1_m9 observable-mass scaling relation and L5p6_m10_DMO HMF). The error bars indicate the Poisson errors in each bin. Note that the left y-axis is for the dotted (Planck) and solid (SPT) lines, while the right y-axis is for the dashed lines (SO).

masses, while the HMF uses DMO masses, without correcting for it. We choose instead to investigate this potential effect separately. The distribution of sources with redshift is shown in Fig. 4.1. From the figure it is clear that the error bars become much smaller for the larger samples. There is a also a slight shift towards sampling higher redshift objects with deeper surveys. Due to the limited redshift range of the MiraTitanEmulator, we are unable to use its predictions for number counts at $z > 2$. However, from this figure it is clear that this leads to missing a negligible fraction of all sources, even for the deepest survey.

4.2.4 The likelihood

To fit for cosmological parameters, we make use of a Poisson likelihood. We have verified that this is a valid assumption by bootstrap resampling different parts of the sky in the halo lightcone, finding that the error on the cluster counts follows a Poisson distribution. Taking the model counts λ_i and the observed counts x_i for each redshift bin i, the likelihood of binned data $\mathbf{x}$ being the result of model data

$\boldsymbol{\lambda}$ can be expressed as

$$P(\mathbf{x}|\boldsymbol{\lambda}) = \Pi_i^N \frac{e^{-\lambda_i}\lambda_i^{x_i}}{x_i!}, \tag{4.9}$$

$$\ln P = \sum_i^N (x_i \ln \lambda_i - \lambda_i - \ln(x_i!)), \tag{4.10}$$

where we will use $\ln P$ as the log likelihood. For our cosmological fits we vary only Ω_m and σ_8, keeping all the other values fixed to the values from Abbott et al. (2022), the cosmology used for FLAMINGO. For the cosmological parameters we use flat priors that are based on the available parameter space for the MiraTitanEmulator, which can be found in Table 4.1. We use the publicly available package emcee (Foreman-Mackey et al., 2013) using the ensemble sampler. We run the chains using 40 walkers for 2500 steps, where the first 500 are discarded.

To quantify the absolute and statistical biases of our posteriors with respect to the truth, we employ two different measures of the bias. The first gives us the fractional error of the median compared with the truth

$$\text{Fractional error} = \frac{x-\mu}{\mu}, \tag{4.11}$$

where x is the median and μ is the true value. We calculate the fractional error separately for σ_8 and Ω_m. Additionally, we calculate how many sigmas the truth is away from the posterior median values using the geometric distance assuming a covariant Gaussian

$$N\sigma = (x-\mu)^T \Sigma^{-1} (x-\mu), \tag{4.12}$$

where Σ is the covariance matrix. We provide these two different bias estimates to give an indication of the level of precision at which certain biases start to become dominant and to show how statistically significant the biases are for a purely cosmic-variance limited cosmological inference. Hence, the number of σ should be interpreted as a worst case scenario for a given fractional error, as observational uncertainties will likely significantly reduce this number.

4.3 Results

In this section we will investigate how the following assumptions and uncertainties affect cluster counts:

- The DMO HMF model (§4.3.1)
- The effect of baryons on the HMF (§4.3.2)
- The functional forms for the analytic observable-mass scaling relation and scatter (§4.3.3)

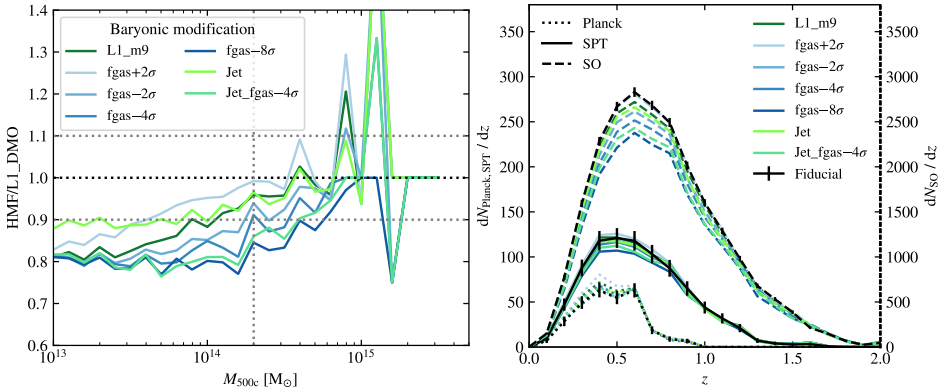


Figure 4.3: The left panel shows the ratio between the HMF of the different FLAMINGO feedback variations, given by the legend, and the FLAMINGO L1_m9_DMO at $z = 0$. The black dotted horizontal line shows the one-to-one line and the grey dotted horizontal lines show ten per cent deviations. The grey dotted vertical line roughly coincides with the lowest mass clusters that will be in the deepest (SO) sample. The functions in the left panel are used to modify the HMF of the fiducial DMO model and to create the right panel which shows the distribution of clusters with redshift for the three different surveys, indicated by different line-styles. The different coloured lines show the effect that baryons have on the HMF for each FLAMINGO simulation. The black lines with error bars shows the cluster counts from the fiducial model (L1_m9 observable-mass scaling relation and L5p6_m10_DMO HMF). The error bars are the Poisson errors in each bin. Note that the left y-axis is for the dotted (Planck) and solid (SPT) lines, while the right y-axis is for the dashed lines (SO).

- The effect of changes in the assumed scaling relation (§4.3.4)

For the effect of baryons on the HMF and the effect of changes in the assumed scaling relation we will make use of the variations in gas fraction and AGN feedback model in the FLAMINGO suite of simulations. After investigating the effect on cluster counts for each type of assumption separately, we will compare the biases induced by different methods (§4.3.5) and see quantitatively how they impact the inference of σ_8 and Ω_m (§4.3.6).

4.3.1 DMO Halo mass function models

First we investigate the effect of different models for the DMO HMF. Fig. 19 of Schaye et al. (2023) shows that the FLAMINGO DMO HMF agrees well with models taken from the literature. However, those plots were made using the Velociraptor (Elahi et al., 2019) halo finder for $M_{200\mathrm{m}}$, while we use the HBT+ (Han

et al., 2017) halo finder and M_{500c}. Furthermore, there were some systematic deviations that could lead to biases. We focus on comparing three widely used DMO HMF models at fixed cosmology with FLAMINGO: Tinker et al. (2010), Bocquet et al. (2016) and the MiraTitanEmulator (Bocquet et al., 2020). We compare their predictions for the HMF and the resulting cluster counts for the FLAMINGO cosmology and compare with our fiducial setup, for which the HMF is based on the L5p6_m10_DMO simulation. We investigate the accuracy of the models in Fig 4.2. In the left panel we show the ratio between the different HMF models and the L5p6_m10_DMO FLAMINGO simulation at $z = 0$. In the right panel we show the distribution of clusters with redshift for the three different mock surveys indicated with different line styles. In both panels the colours indicate the different HMF models.

The Tinker et al. (2010) model and the MiraTitanEmulator fall outside of the error bars for all three mock surveys. As can be seen in the left panel, the Tinker et al. (2010) HMF over-predicts the number counts at all masses. From the right panel it is clear that the over-prediction is worst at intermediate redshifts, but that the agreement becomes better for $z \geq 1$. However, this is beyond the redshift range that contain most objects for our three mock surveys.

The MiraTitanEmulator is relatively accurate at low mass, overshooting by only a few per cent up to a mass of $M_{500c} \approx 2 \times 10^{14}$ $M_\odot$. Above this mass there is a large upturn and the model over-predicts the number of massive clusters. In the right panel it can be seen that this leads to a systematic over-prediction of the number of objects for all three mock surveys for the entire redshift range. To compute the M_{500c} mass function using the MiraTitanEmulator, which predicts the M_{200c} HMF, we assumed a NFW density profile and the mass-concentration relation from Diemer & Joyce (2019). At the steep end of the HMF, slight deviations from the truth and scatter in the mass-concentration relation has a big effect on the results. Our results highlight the fact that it is important to use models that emulate the quantity of interest directly, as conversion factors will introduce additional biases.

The HMF model that best recovers the FLAMINGO cluster counts is the model from Bocquet et al. (2016). From the left panel of Fig. 4.2 it is clear that this is due to the fact that the model agrees with the simulation at the few per cent level up to $M_{500c} \approx 7 \times 10^{14}$ $M_\odot$. Above this mass the model underpredicts the number of clusters. This explains the behaviour of this model in the right panel. For the Planck-like survey, the model undershoots the results, as the Planck-like survey is mostly limited to very massive clusters. For the deeper surveys, that probe objects down to a lower mass than where the model starts to deviate, there is reasonable agreement. For the SPT-like survey it still undershoots, but stays largely within the error bars. For much of the redshift range, there is good agreement with the SO-like survey, only deviating at high redshifts, though it is hard to see in this plot whether it stays within the error

bars. Because of this relatively good agreement, we will assume the model by Bocquet et al. (2016) for later comparisons and fits where we need to vary the cosmology.

From this comparison of HMF models it is clear that the choice of model will have a direct influence on the inferred cosmology, and that there are big differences between the models. For $M_{500c} \lesssim 2 \times 10^{14}$ $M_\odot$ the models by Bocquet et al. (2016) and Bocquet et al. (2020) are in good agreement with the FLAMINGO DMO simulation. However, most objects, especially in SZ-selected surveys, will have higher masses than this. While the sensitivity of the high mass end of the HMF to cosmology makes it an interesting target, the model predictions are very sensitive to any of the choices made. This includes, for example, mass conversions, simulation box size effects, halo definitions and the halo finder used. Therefore it might be more beneficial to focus on slightly lower mass objects that suffer less from these systematic errors, something that will happen naturally with upcoming surveys.

4.3.2 Baryonic effects on the halo mass function

In addition to biases introduced by the choice of DMO HMF model, all of the models neglect the effect of baryons on the halo masses. If the observable-mass scaling relation is obtained from hydrostatic masses (Planck Collaboration et al., 2016a) or weak lensing masses (Bocquet et al., 2024), then it is the total mass, that includes both baryons and dark matter, that is probed. If the halo baryon fraction deviates from the universal fraction, then this leads to an inconsistency between the scaling relation and the HMF used for modelling. For the largest objects the effect is small, but the effect is expected to increase towards lower masses (see e.g. Velliscig et al., 2014; Cui et al., 2014; Schaye et al., 2023). To quantify the uncertainty in the baryonic effect on the HMF, we will make use of the FLAMINGO feedback variations, which span the uncertainties in the observed gas fraction data (Kugel et al., 2023).

The impact of the baryonic effects on the HMF is shown in Fig. 4.3. Similar to Fig. 4.2, the left panel shows the ratio between the HMF of the variations and the L1_m9_DMO simulations at $z = 0$. We choose to compare with the L1_m9_DMO instead of the L5p6_m10_DMO as these simulations use the same initial conditions and therefore provide a fairer comparison when investigating just the effect of baryons. Baryonic effects suppress the mass function more strongly at lower masses. For the fiducial model the suppression increases from ≈ 20 per cent at $M_{500c} \sim 10^{13}$ $M_\odot$ to ≈ 10 per cent at 10^{14} $M_\odot$. As expected, models with lower gas fractions lead to stronger suppression. A $(1\ \mathrm{Gpc})^3$ volume only contains a limited number of objects with $M_{500c} > 10^{15}$ $M_\odot$. Therefore, the ratio becomes quite noisy for the highest masses. As noted in Section 4.2.2, we set the ratio to one when either of the bins (in the DMO or the Hydro simulation) is empty, which

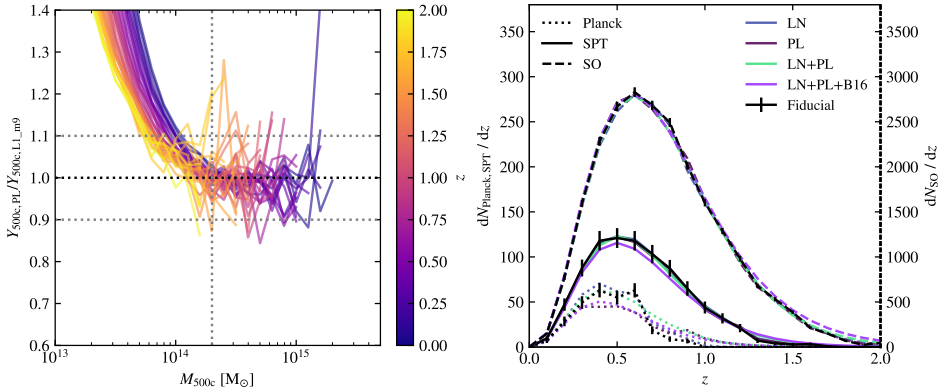


Figure 4.4: The left panel shows the ratio between the best-fitting power-law $Y_{500c}(M_{500c})$ to the scaling relation of the L1_m9 simulation and the actual scaling relation at different redshifts, indicated by the different colours. The black dotted horizontal line shows the one-to-one line and the grey dotted horizontal line shows ten per cent deviations. The grey dotted vertical line indicates the lower mass limit used for fitting which roughly coincides with the lowest mass clusters that will be in the deepest (SO) survey. The right panel shows the distribution of clusters with redshift for the three different surveys, indicated by different line-styles. The different coloured lines show the effect of assuming the scaling relation to be a power-law (PL), the scatter to be lognormal (LN), the HMF to be given by the MiraTitanEmulator (MTE) and combinations of these assumptions. The black line with error bars shows the cluster counts from the fiducial model (L1_m9 observable-mass scaling relation and L5p6_DMO HMF). The error bars are the Poisson errors in each bin. Note that the left y-axis is for the dotted (Planck) and solid (SPT) lines, while the right y-axis is for the dashed lines (SO).

can be seen for the largest masses.

The right panel shows how the deviations from DMO, shown in the left panel, propagate into changes in the cluster counts. Because the ratio is noisy for the highest masses, we can see in the right panel that some of the models fall outside the error bars for the Planck-like survey. The fact that we find disagreement for the sample that covers the highest mass clusters indicates that we are dominated by small number statistics due to the limited simulation volume. Therefore, we choose not to make any definite statements on whether the effect of baryons on the HMF would lead to a large bias for a Planck-like sample.

For the SPT-like survey most variations fall within the Poisson error bars, only the most extreme models, fgas-8σ and Jet_fgas-4σ, fall outside the error bars. These models start to deviate by over 10% close to the mass cut. For the SO-like survey, nearly every baryonic modification leads to disagreements beyond the Poisson error bars over most of the redshift range, especially at $z \approx 0.5$. Only the

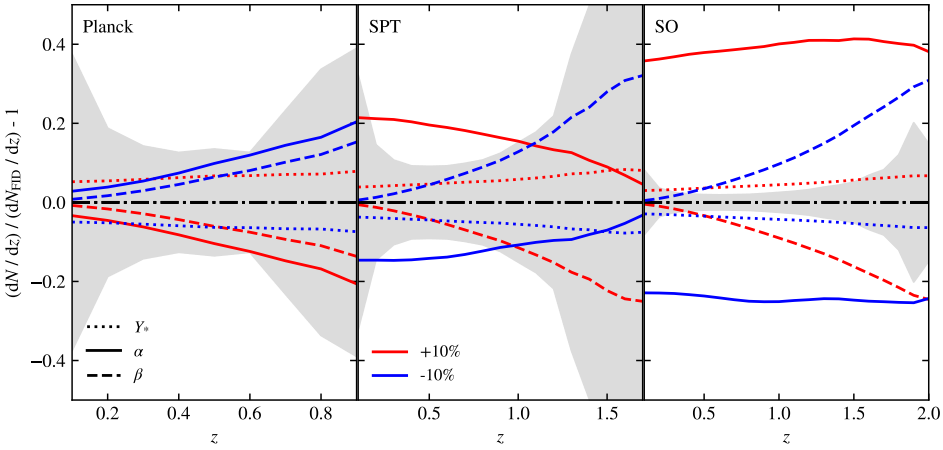


Figure 4.5: The relative impact of changes to the power-law parameters to the cluster counts. From left to right, the three panels show the impact for the Planck-, SPT- and SO-like surveys respectively. The different line-styles indicate the three different power-law parameters (Eq. 4.5): the amplitude Y_*, the scaling with mass α and the redshift scaling β. The red (blue) lines increase indicate an increase (decrease) by 10% relative to the fiducial value of each parameter. The black dot-dashed line indicates the one-to-one line. The grey band indicates the Poisson errors on each sample. Accuracy higher than 10% is needed for every parameters for SPT and SO. The sensitivity to the mass scaling increases for deeper surveys.

weakest feedback model has a small enough impact to still be consistent with the fiducial model. This can also be seen in the left panel, where the fgas+2σ model is consistent within a few per cent for the masses probed by the SO mock survey, indicated with a vertical dotted line. We conclude that future observations cannot neglect the effect that baryons have on the HMF.

4.3.3 Fit to the observable-mass scaling relation

The next two subsections explore the effect of varying the assumptions for the mass-observable scaling relation, i.e. the relation between M_{500c} and Y_{500c}, and its scatter. In this section, we investigate the effect of assuming a power-law with lognormal scatter to model the scaling relation. In subsection §4.3.4 we explore the effect of getting the scaling relation slightly wrong by making use of the FLAMINGO feedback variations.

The power-law parameterization we use is described in Section 4.2.2 and is taken from Planck Collaboration et al. (2016a) (Eq.4.5). Instead of using their best fitting values, we refit the amplitude Y_*, the slope α, the redshift scaling β

and the lognormal scatter σ. We re-fit to isolate the effects of assuming a power-law functional form. We only fit these parameters for masses $M_{500c} > 2{\times}10^{14}\,\mathrm{M}_{\odot}$. The result of this can be seen in the left panel of Fig. 4.4. The re-fit ensures that the relation leads to a match for the masses probed by our three surveys, across all redshifts, though extrapolation to lower cluster masses would result in large errors. The results of the fit are listed in Table 4.1.

If we compare the values we find with those given in literature, we find that the amplitude Y_* and the slope α both have to change by more than 1σ. The redshift scaling β is within the 1σ range. However, we note that changing β within the 1σ range can lead to large changes in the predicted cluster counts. The value $\beta = 0.89$ that we find for FLAMINGO is also significantly different from the self-similar scaling, $\beta = 2/3$. The hydrostatic mass bias b is degenerate with the amplitude Y_*. The parameters β and b are both kept fixed in the analysis by Planck Collaboration et al. (2016a). To get rid of the degeneracy due to the hydrostatic bias we assume a fixed value given by the FLAMINGO calibration by Kugel et al. (2023). However, we note that using the value of the hydrostatic bias used by Planck Collaboration et al. (2016a) ($b = 0.8$) has a big impact on the resulting value of Y_*. The log-normal scatter we find for FLAMINGO, $\sigma = 0.081$, is consistent with the Planck prior, $\sigma = 0.075 \pm 0.01$. As the scatter for Compton-Y within R_{500c} is close to lognormal (Kugel et al., 2024), any deviations found here are more likely to be an effect of the scatter not being constant with mass or redshift.

While we obtain a good match to the Compton-Y values for the masses we fit to, we find that the scaling relation quickly starts to deviate at lower masses, indicating that a double power-law might better describe the scaling of the integrated SZ signal with mass. However, with current facilities we cannot probe these masses efficiently observationally, so we do not attempt to constrain a double power-law. Above the mass cut we match the relation up to redshift two.

The impact of the assumptions for the scaling relation on the cluster counts can be seen in the right panel of Fig. 4.4. The different abbreviations stand for lognormal scatter (LN), power-law (PL) and using the model by Bocquet et al. (2016) (B16) for the HMF. For the Planck-like survey, the power-law leads to a substantial decrease in the number of objects across most redshifts. At $z > 0.7$, all assumptions lead to large overestimates. Just like for the HMF modification, it is likely that this behaviour is in large part due to the limited statistics for very massive clusters in the L1_m9 simulation. For the SPT-like survey we see that the agreement is generally very good. At nearly all redshifts, the lognormal scatter and power-law assumptions have a negligible effect. Only when combined with the HMF model from Bocquet et al. (2016) (LN+PL+B16), which we know from Fig. 4.2 leads to deviations, do we see a noticeable impact. As shown in a previous section, this is driven by errors in the HMF.

For the SO-like survey, the relative differences seen in the plot are very small,

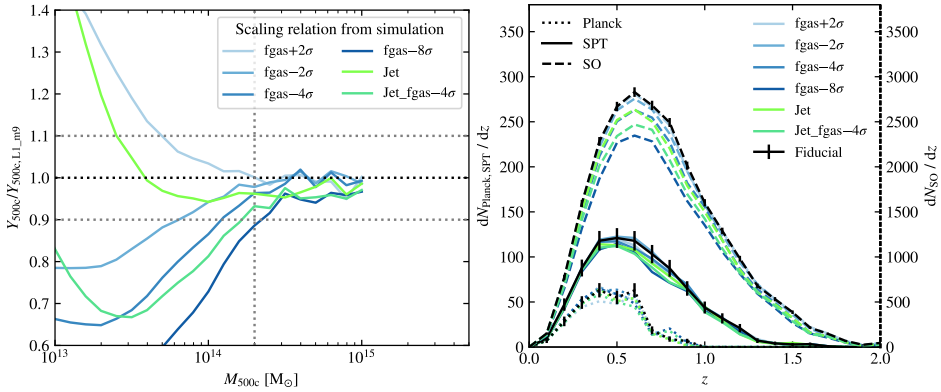


Figure 4.6: The left panel shows the ratio of the $Y_{500c}(M_{500c})$ scaling relation of the different FLAMINGO feedback variations, listed in the legend, and that of the fiducial L1_m9 simulation at $z = 0$. The black dotted horizontal line shows the one-to-one line and the dotted grey horizontal line show ten per cent deviations. The grey dotted vertical line roughly coincides with the lowest masses that will be in the deepest (SO) sample. The right panel shows the distribution of clusters with redshift for the three different surveys, indicated by different line-styles. The different coloured lines show the effect changing the scaling relation to one of the variations has for each FLAMINGO simulation. The black line with error bars shows the cluster counts from the fiducial model (L1_m9 observable-mass scaling relation and L5p6_m10_DMO HMF). The error bars are the Poisson errors in each bin. Note that the left y-axis is for the dotted (Planck) and solid (SPT) lines, while the right y-axis is for the dashed lines (SO).

and there are almost no systematic offsets. However, as the sample is much larger, the requirements for a hypothetical cosmological parameter inference to be accurate down to cosmic variance levels are also much stricter. The largest deviation is caused by the introduction of the HMF model by Bocquet et al. (2016), as expected from section 4.3.1. This is seen for $z > 1$ where there is an over-prediction of the number of clusters. Other deviations can be seen around the peak of the distributions, where all the assumptions lead to an underestimate that is outside the Poisson error bars either before the peak, for LN, or after the peak, for PL and LN+PL. As discussed in Section 4.3.5, these minute differences still lead to larger errors than those due to cosmic variance.

Provided the observed scaling relations can be constrained in an unbiased way, these results show that the commonly assumed power-law functional form is unlikely to lead to a large bias, at least not for surveys that are only sensitive to masses $M_{500c} > 2 \times 10^{14}\,\odot$. However, the values of the parameters of the scaling relation will need to be fit with high accuracy. To investigate what accuracy is needed, we look at the effect of changing each parameter in Fig. 4.5. In contrast

with previous figures, we show the relative change with respect to a model that uses the fiducial values for the fitting parameters. In other words, the difference between the one-to-one line and the coloured lines is only due to the change in a single parameter. The different line-styles indicate the three different parameters we vary. The colour indicates whether the parameter is increased (red) or decreased (blue) by 10%. The grey coloured regions indicate the Poisson errors on each survey. For reference, for the constraints used by Planck Collaboration et al. (2016a) the relative uncertainty is about 10% for Y_*, 5% for α and 75% for β. This is improved upon by Bocquet et al. (2024). While they use a different definition for the scaling relations, their respective parameters for Y_*, α and β have an accuracy of 8.7%, 2.3% and 14.8% respectively in their final analysis.

Looking at the different line-styles in Fig. 4.5, we can see that for a large part of the redshift range, the parameters α and β have the largest effect at a fixed relative uncertainty of 10 per cent. The uncertainty in β leads to larger errors with increasing z. The uncertainty in α leads to bigger errors as the survey gets deeper. This is as is to be expected. The functional form for the power-law has a pivot mass at $M_{500c} = 6\times10^{14}\,\mathrm{M}_\odot$. An error in α leads to an increasing bias in the scaling relation when the masses differ more from the pivot mass. The amplitude Y_* leads to a similar level of uncertainty of about 5 to 10 per cent for each survey, slightly decreasing for the deeper surveys.

Going panel-by-panel, it is clear that a 10% uncertainty in the fitting parameter only leads to an uncertainty in the cluster counts that is within the Poisson error bars for the Planck-like survey. For the SPT and SO mocks, a higher accuracy is needed for each of the three parameters. The constraints by Planck Collaboration et al. (2016a) on the parameter β have a much larger uncertainty (75%) than the 10% shown in this figure. Furthermore, a 10% deviation from the best fitting value found for FLAMINGO is still inconsistent with self-similar evolution. This implies that assuming the evolution is self-similar might lead to significant biases. As the deeper surveys also probe higher z, it is important that the redshift evolution is captured correctly. Additionally, for deeper surveys the pivot mass should be moved to lower masses to reduce the error due to the uncertainty in α. At the current pivot, an error in α of less than a per cent is required to keep the results within the SO Poisson error bars. For the SPT mock, we see that the constraints used by Bocquet et al. (2024) are not tight enough to push the uncertainties below the level of the Poisson errors. It is important to note that these parameters and their uncertainties are marginalised over when inferring cosmology and that the uncertainties on these parameters will factor into the uncertainty in the cosmology. Our results highlight the importance of fully understanding the selection effects when constraining the scaling relations, as small errors can lead to large biases in the inferred cosmology.

Table 4.2: Compilation of the level of agreement between the cluster number counts predicted using the various assumptions tested in this work and those predicted by the fiducial model, which uses the L1_m9 observable mass scaling relation and the L5p6_m10_DMO HMF. Note that in each case the cosmology is fixed to the true one. The first column indicates the four different types of assumptions tested, the second column specifies which model was used for that type of assumption. The final six columns give the χ^2 and the p-value for the hypothesis that the model is consistent with the mock survey, split by the three mock surveys. Every time an assumption leads to a discrepancy that has a p-value lower than 0.2, it is shown in boldface. Entries with a higher value lead to systematics that are consistent within the Poisson errors.

Assumption	Model	Planck χ^2	Planck p	SPT χ^2	SPT p	SO χ^2	SO p
DMO HMF	Tinker (2010)	50.7	$\mathbf{< 10^{-6}}$	107.1	$\mathbf{< 10^{-6}}$	1790.7	$\mathbf{< 10^{-6}}$
	Bocquet (2016)	10.7	$\mathbf{1.54 \times 10^{-1}}$	2.9	1.00×10^{0}	68.9	$\mathbf{< 10^{-6}}$
	MiraTitanEmulator	36.0	$\mathbf{7.38 \times 10^{-06}}$	64.4	$\mathbf{< 10^{-6}}$	712.5	$\mathbf{< 10^{-6}}$
Baryon effect on the HMF	L1_m9	4.2	7.56×10^{-1}	2.1	1.00×10^{0}	22.4	2.14×10^{-1}
	fgas+2σ	11.5	$\mathbf{1.18 \times 10^{-1}}$	3.3	1.00×10^{0}	5.9	9.97×10^{-1}
	fgas−2σ	1.7	9.75×10^{-1}	2.4	1.00×10^{0}	96.7	$\mathbf{< 10^{-6}}$
	fgas−4σ	1.1	9.93×10^{-1}	3.4	1.00×10^{0}	226.9	$\mathbf{< 10^{-6}}$
	fgas−8σ	0.4	1.00×10^{0}	9.2	8.68×10^{-1}	563.1	$\mathbf{< 10^{-6}}$
	Jet	1.3	9.88×10^{-1}	2.6	1.00×10^{0}	80.0	$\mathbf{< 10^{-6}}$
	Jet_fgas−4σ	0.4	1.00×10^{0}	6.2	9.77×10^{-1}	404.0	$\mathbf{< 10^{-6}}$
Scaling relation fit	LN	14.5	$\mathbf{4.29 \times 10^{-02}}$	2.5	1.00×10^{0}	35.6	$\mathbf{8.03 \times 10^{-03}}$
	PL	22.3	$\mathbf{2.22 \times 10^{-03}}$	1.9	1.00×10^{0}	20.5	3.04×10^{-1}
	LN+PL	49.1	$\mathbf{< 10^{-6}}$	13.5	5.66×10^{-1}	127.6	$\mathbf{< 10^{-6}}$
	LN+PL+B16	33.2	$\mathbf{2.44 \times 10^{-05}}$	23.2	$\mathbf{7.97 \times 10^{-02}}$	412.9	$\mathbf{< 10^{-6}}$
Scaling relation variations	fgas+2σ	4.3	7.48×10^{-1}	2.6	1.00×10^{0}	4.0	1.00×10^{0}
	fgas−2σ	2.6	9.18×10^{-1}	0.9	1.00×10^{0}	31.3	$\mathbf{2.69 \times 10^{-02}}$
	fgas−4σ	2.9	8.90×10^{-1}	2.1	1.00×10^{0}	156.3	$\mathbf{< 10^{-6}}$
	fgas−8σ	20.8	$\mathbf{4.00 \times 10^{-03}}$	11.7	6.99×10^{-1}	762.6	$\mathbf{< 10^{-6}}$
	Jet	2.3	9.41×10^{-1}	7.8	9.32×10^{-1}	191.4	$\mathbf{< 10^{-6}}$
	Jet_fgas−4σ	16.9	$\mathbf{1.81 \times 10^{-02}}$	12.3	6.59×10^{-1}	462.3	$\mathbf{< 10^{-6}}$

4.3.4 Systematic uncertainty in the observable-mass scaling relation

In the previous section we have investigated the impact of assuming a power-law with lognormal scatter on the cluster counts and quantified how tightly the parameters of the power-law need to be constrained to get unbiased results. In this section we investigate the effect of changes to the scaling relation itself. To do this, we make use of the FLAMINGO variations. Similar to previous figures, we show both the change in the varied relation and the resulting effect on the cluster counts in Fig. 4.6.

In the left panel of Fig. 4.6, we show the ratio between the $Y_{500c}-M_{500c}$ scaling relation of each of the simulation variations, and the scaling relation from the L1_m9 simulation used for our fiducial model at $z = 0$. The right panels shows the impact of these deviations from the fiducial model on the cluster counts for our three mock surveys, indicated by the different line-styles. The black lines show the fiducial model, including Poisson errors.

At $z = 0$, and for masses included in all samples (the mass limit is indicated by the vertical grey line in the left panel), the changes in the scaling relations mostly fall within the 10% region, but the deviations become much larger at lower masses. All models are similar for the largest objects. This is as expected, since AGN feedback is unable to offset the inflow of gas in the most massive haloes. Towards lower masses the different variations diverge, with models with higher gas fractions having a higher integrated Compton-Y at fixed mass. There is an offset between the thermal and 'Jet' models at fixed gas fractions, indicating that the feedback model changes the resulting scaling relation's dependence on gas fraction.

Looking at the right panel of Fig. 4.6, the quantitative difference due to changing the scaling relation is similar to what was found for the baryon effect on the HMF. For Planck most models are quite close to the Poisson error bars, and at high z they are likely affected by noise as they do not deviate systematically. For SPT, A large fraction of the models falls within the Poisson error bars. Only the models with the lowest gas fractions deviate significantly, especially for $z \approx 0.75$. It is only for the SO-like mock survey that almost every variation leads to a significant difference. Only the fgas+2σ model agrees with the fiducial model within the Poisson error bars.

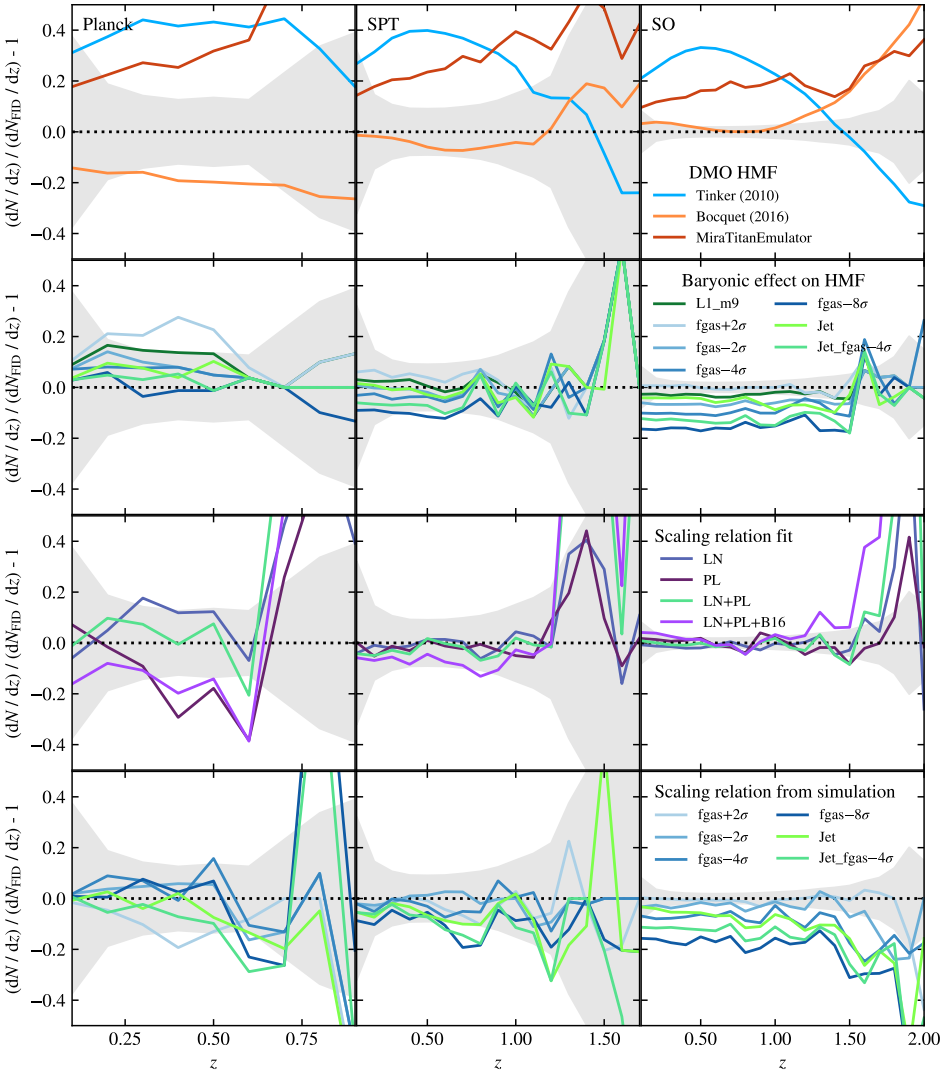

Planck
SPT
SO
DMO HMF
Tinker (2010)
Bocquet (2016)
MiraTitanEmulator
Baryonic effect on HMF
L1_m9
fgas+2σ
fgas−2σ
fgas−4σ
fgas−8σ
Jet
Jet_fgas−4σ
Scaling relation fit
LN
PL
LN+PL
LN+PL+B16
Scaling relation from simulation
fgas+2σ
fgas−2σ
fgas−4σ
fgas−8σ
Jet
Jet_fgas−4σ
(dN / dz) / (dN_FID / dz) - 1
z

Figure 4.7: A compilation of all the cluster count comparisons in this work. Each panel shows the ratio between the cluster number counts predicted by a model making the indicated assumption and those predicted by and our fiducial model, which uses the L1_m9 mass-observable scaling relation and scatter, and the L5p6_m10_DMO HMF. The different model assumptions are indicated by the different coloured lines. The Poisson uncertainty is given by the grey shaded region. The three columns show the results for the three mock surveys considered in this work, from left to right, Planck, SPT and SO. The first row varies the model for the HMF, the second row varies the baryonic modification of the DMO HMF, the third row investigates different fits to the observable-mas relation predicted by the fiducial model and the fourth row varies the simulation from which the mass-observable scaling relation and scatter is taken. In the third row the abbreviations are lognormal (LN), power-law (LN) and using the HMF model by Bocquet et al. (2016) (B16). The biggest systematic errors for all surveys are caused by the HMF models, however, for the SO-like survey, nearly all assumptions lead to a significant errors.

Assuming that deviations from the models used for cluster cosmology are as small as for our mildest feedback variations, it seems that for previous and current generation surveys the difference is still small enough for the effect to be within the Poisson error bars. However, for future surveys, the necessity of constraining the scaling relation more tightly is clear.

4.3.5 Quantifying the biases in the number counts

In this section we will investigate and compare the effects on the cluster number counts of the different assumptions that we investigated for the three different mock surveys. This way we can get a good overview for all the assumptions that might lead to biases in the inferred cosmology.

In Table 4.2 we compile the effect of each assumption in terms of χ^2 and the p-value with respect to our fiducial model, which consists of the observable-mass scaling relation from the L1_m9 simulation with the L5p6_m10_DMO HMF. To highlight the problematic cases, all deviations that have a p-value below 0.2 are highlighted using boldface. Like in the previous sections, the different assumptions are divided into four categories: the DMO HMF model, baryonic effects on the HMF, assumptions made when fitting the observable-mass scaling relation and variations in the scaling relation itself. Additionally, we summarise all the findings so far in Fig. 4.7. This figure combines the right panels of Figs 4.2, 4.3, 4.4 and 4.6, but instead shows the ratio of the cluster number counts with respect

to our fiducial model. The three mock surveys are shown in the three columns with their Poisson errors shown as the grey shaded region.

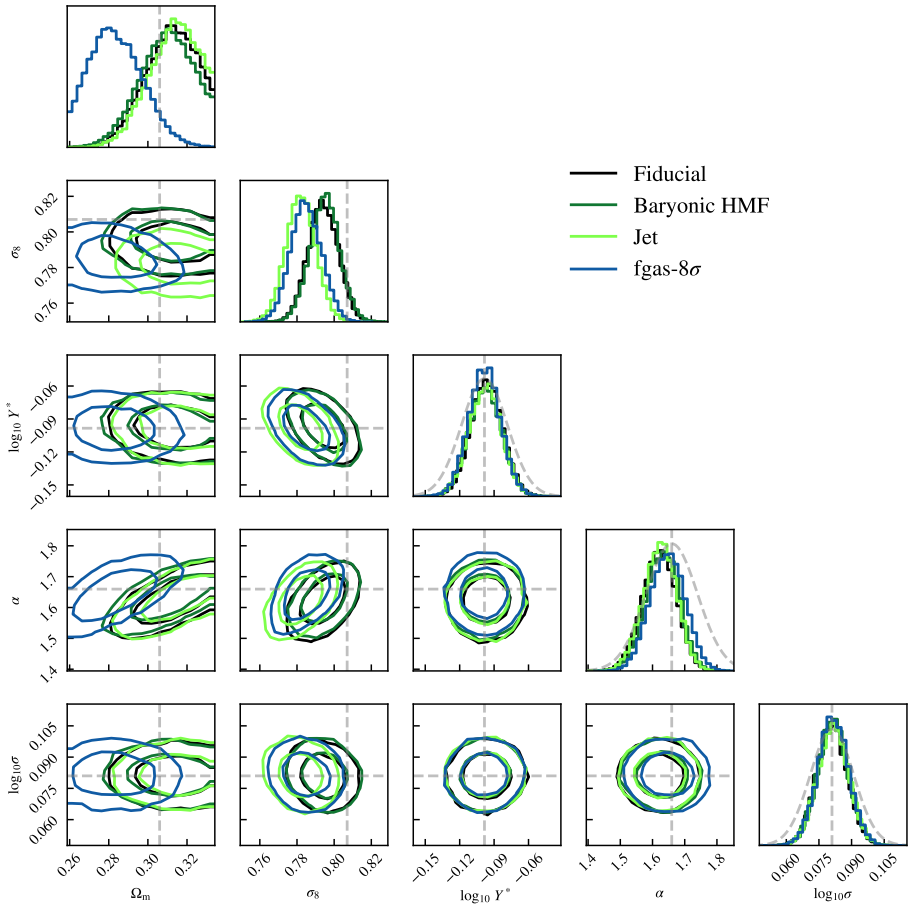


Figure 4.8: Posteriors from fitting the "Full" fitting model for cluster counts, which assumes a power-law mass-observable scaling relation with lognormal scatter and the DMO HMF model from Bocquet et al. (2016), to the cluster count predictions for four different mocks setups at the Simons observatory detection limit: Our fiducial mock survey ("Fiducial"; L1_m9 observable-mass scaling relation and scatter and a L5p6_m10_DMO HMF), a mock survey that includes the baryonic modification of the HMF ("Baryonic HMF"; as the Fiducial model but with the baryonic modification of the HMF predicted by L1_m9), a setup that uses an alternative model for AGN feedback that employs kinetic jets instead of thermally-driven winds ("Jet"; the mass-observable scaling relation, scatter and baryonic modification of the HMF are taken from the Jet simulation) and a setup that uses the simulation with the strongest reduction in cluster gas fractions ("fgas-8σ"; the mass-observable scaling relation, scatter and baryonic modification of the HMF are taken from the fgas-8σ simulation). We fit for the cosmological parameters Ω_m and σ_8, the two power-law observable-mass scaling relation parameters Y^* and α (see Eq. 4.5), and the lognormal scatter around this relation σ. The horizontal and vertical grey dashed lines show the true values. In the diagonal panels for the parameters Y^*, α and σ, we also show the priors using grey dashed lines. The two contour levels show the 68th and 95th percentiles. There are significant biases in the cosmological parameters for all mocks.

By focusing on the boldfaced entries in Table 4.2, we get a clear overview of the findings that we reported earlier in this work. There are a few groups of problematic assumptions to highlight. One of these groups is found for some of the results for the Planck-like survey. As explained earlier, these deviations are mostly due to the limited (1 Gpc)^3 volume used for this work, which leads to additional noise for the cluster masses found by the Planck-like survey.

The second group consists of the DMO HMF models, which all lead to significant biases for all surveys, with the one exception being the model by Bocquet et al. (2016) for the SPT-like survey. This can also be seen in Fig. 4.7. In all other cases the models fall outside the error bars for most of the redshift range. For the SO-like survey the requirements are particularly stringent. The model by Bocquet et al. (2016) agrees very well for a large part of the redshift range, but outside of that range the deviations quickly become so large that it leads to a p-value that indicates a high level of disagreement.

Finally, nearly the entire SO column is boldfaced. None of the DMO HMF models are accurate enough. All FLAMINGO variations but the ones with the weakest feedback have enough of an effect on the HMF and scaling relation to lead to significant differences. The assumption of lognormal scatter on the mass-observable scaling relation, especially in combination with the assumption that

the scaling relation is a power-law and the previously mentioned inaccuracies of the DMO HMF model assumptions, lead to large deviations. From Fig. 4.7 it is clear that for the baryonic variations, the effects on the HMF and on the observable-mass scaling relation are insensitive to redshift, leading to systematic offsets across the entire redshift range. For the lognormal scatter and power-law assumptions, the discrepancies are greatest at $z > 1$. These findings highlight the point that for future surveys our current models use assumptions that are too restrictive.

Our findings are in line with those of Bocquet et al. (2016), who also find that post-SPT like surveys need to take baryonic corrections into account, and that we need to be careful in constructing the HMF models as different current models in the literature will lead to the inference of significantly different cosmologies.

4.3.6 Biases in cosmological parameters

As a final test we want to quantify the impact of the assumptions that are commonly made in models predicting cluster counts on inferred cosmological parameters for SPT- and SO-like surveys. We exclude the Planck-like survey from this section as we found previously that the systematics are limited by our finite simulation volume.

To fit mock survey data, we make use of the "full" model, LN+PL+B16, which assumes the DMO HMF model from Bocquet et al. (2016), the power-law observable-mass scaling relation given by Eq. 4.5 with lognormal scatter. We fit for five parameters: the cosmological parameters Ω_m and σ_8, the power-law observable-mass scaling relation parameters Y_* and α, and the lognormal scatter in the scaling relation σ. For the cosmological parameters we use the flat priors given by the limits of the MiraTitanEmulator [$\Omega_\mathrm{m} = \mathcal{U}(0.259, 0.334), \sigma_8 = \mathcal{U}(0.7, 0.9)$]. For the power-law and scatter parameters we use Gaussian priors taken from Planck Collaboration et al. (2016a) centered on the values that best fit our fiducial simulation L1_m9 [$\log_{10} Y_* = \mathcal{N}(-0.098, 0.02), \alpha = \mathcal{N}(1.66, 0.08), \sigma = \mathcal{N}(0.081, 0.01)$]. We keep the redshift scaling β and the hydrostatic bias b fixed [$\beta = 89, b = 0.0743$]. See also Table 4.1.

Besides our full fitting model, we use another fitting model, referred to as "HMF", that also assumes the DMO HMF from Bocquet et al. (2016), but uses the actual L1_m9 observable-mass scaling relation and scatter, rather than the power-law fit with lognormal scatter used for the full model. This model will aid in separating the errors induced by the use of the DMO HMF model from Bocquet et al. (2016) from the errors induced by the assumption of a power-law scaling relation with lognormal scatter. In the HMF case we only fit for the cosmological parameters Ω_m and σ_8.

For each type of survey (SPT- or SO-like), we use four different mock survey setups to quantify the effect of biases in the fitting models. The first setup uses the L1_m9 scaling relation and scatter (rather than a power-law fit and lognormal scatter) and the L5p6_m10_DMO HMF. This mock is fit using both the HMF and full models. As this mock setup and the fitting models both use a DMO HMF, these two fits help us quantify errors due to the choice of DMO HMF model, and errors due to the assumption of a power-law observable-mass scaling relation with lognormal scatter.

The second mock setup, referred to as "baryonic HMF", again uses the L1_m9 scaling relation and scatter and the L5p6_m10_DMO HMF, but we modify the DMO HMF by the baryonic effect predicted by the L1_m9 simulation, which we fit using the full fitting model (LN+PL+B16). This way we can quantify the level of bias introduced by baryonic effects on the HMF not being taken into account by the fitting model.

The final two mock setups, which we fit using the full fitting model, use the L5p6_m10_DMO HMF, but the observable-mass scaling relation, scatter and baryonic modification of the HMF are taken from either the Jet or fgas-8σ simulations. These last two mocks can inform us about the impact of inconsistencies between the priors in the observable-mass scaling relation, which are centered on the predictions from L1_m9, and reality (assumed to be either the Jet or fgas-8σ predictions). In the case of the Jet model these inconsistencies are due to a change in the prescription for AGN feedback at fixed cluster gas fractions, while for fgas-8σ they are due to a large reduction in the gas fractions. This yields ten fits in total, five each for the SPT- and SO-like surveys.

The results are shown in Table 4.3. The full model can only reach the level of accuracy necessary to improve on the Planck Collaboration et al. (2020a) errors (i.e. about 2% in $\Omega_{\rm m}$ and 1% in σ_8) for the Fiducial mock setup for the SPT-like survey. The agreement is worst for the Jet and fgas-8σ mock setups. For the SO mocks the fractional systematic errors are generally reduced with respect to the SPT mocks, but, because of the increased precision reached by the SO-like survey, the biases are more more significant, i.e. the errors correspond to a larger number of σ.

Interestingly, for the SO survey the biases are less significant for the baryonic HMF mock setup than for the Fiducial setup. As the full fitting model does not take baryonic effects on the DMO HMF into account, this implies that unaccounted for systematic errors are compensating each other. For all fits using the full model the number of σ is close to or exceeds unity, signaling that all models tested suffer from significant systematics. Comparing the biases for the full and HMF fitting models, it is clear that the DMO HMF model drives part of the errors for both surveys. For the SPT survey about half of the fractional error in $\Omega_{\rm m}$ and a quarter of the fractional error in σ_8 are caused by the DMO HMF model. For SO it accounts for most of the $\Omega_{\rm m}$ error and about half of the σ_8 error.

The posteriors for the SO-like model fits are shown in Fig. 4.8. Focusing first on the power-law and scatter parameters, it is clear that they are all mostly prior driven. When using broad flat priors, the posteriors for the power-law and scatter parameters posteriors extend far beyond realistic ranges. The posteriors are somewhat tighter than the priors, indicated by the dashed grey lines in the diagonal panels, but the differences are small. There are some degeneracies between the power-law and cosmological parameters. Such degeneracies can lead to biases. This can be seen for α, which is biased slightly low due to its degeneracy with σ_8. This illustrates the importance of priors, i.e. independent constraints on the scaling relation parameters, in order to reduce the biases that can be incurred due to movement along the lines of degeneracy. It is also clear from the fits to the Jet and fgas-8σ mocks that changes in the observable-mass scaling relation are not picked up by the fitting. The prior is so dominant that any differences between the prior scaling relation and the truth will manifest itself in a change in cosmological parameters. Relaxing the priors might aid in relative consistency, but this would also degrade the cosmological constraining power.

Clearly, improvements are needed to the modelling of both baryonic effects and the DMO HMF. With current models, the fractional systematic errors exceed the current Planck constraints. Without improvements on the modelling side, we will not be able to use cluster counts to make any definite statements about for example the σ_8 tension.

4.4 Conclusions

Galaxy cluster counts have the potential to provide an alternative avenue to explore cosmological tensions between the cosmological parameters measured by different types of observables. For cluster surveys, selection based on the amplitude of the thermal Sunyaev-Zel'dovich (SZ) effect, i.e. the Compton-Y parameter, is predicted to result in smaller biases relatively to a mass-selected sample compared with e.g. X-ray selection (e.g. Kugel et al., 2024). The number of objects detected in SZ is rapidly increasing (Planck Collaboration et al., 2016a; Hilton et al., 2018; Bocquet et al., 2019, 2024) and future observatories are predicted to dramatically increase our sample size (e.g. Ade et al., 2019). To exploit the statistical power of these surveys, any biases in the predictions of the models must be small compared with the statistical errors.

In the light of these advances we investigated some of the assumptions that go into the models used to measure cosmological parameters using galaxy cluster samples, particularly those from SZ surveys. These models need to assume a model for the halo mass function (HMF), which is typically based on a framework built on dark matter only (DMO) simulations, which is conventionally combined with a power-law mass-observable scaling relation with lognormal scatter. We aim to investigate whether these model ingredients, in particular the choice of

Table 4.3: The fractional and number of σ (relative to Poisson) systematic error (see Section 4.2.4) in the cosmological parameters Ω_m and σ_8 when fitting different models to the cluster counts predicted by the different mock setups for SPT- and SO-like mock surveys. The first column lists the survey from which the selection limit and sky area are taken, either SPT or SO. The second column lists the model that is fit to the mock data. "HMF" uses the true observable-mass scaling relation and scatter from the fiducial simulation (L1_m9) and the DMO HMF from Bocquet et al. (2016). For this fitting model the only free parameters are Ω_m and σ_8. "Full" also uses the B16 DMO HMF, but assumes a power-law (2 free parameters) plus lognormal scatter (1 free parameter) scaling relation. The third column lists the setup of the mock surveys that the models are fit to. "Fiducial" stands for the true L1_m9 scaling relation and scatter combined with the L5p6_m10_DMO HMF. "Baryonic HMF" uses the Fiducial setup but modifies the DMO HMF with the baryon response predicted by L1_m9. The Jet and fgas-8σ setups use the scaling relation, scatter and baryonic modification of the HMF from their respective simulation, with the L5p6_m10_DMO HMF. The fourth and fifth columns list the per cent error in the median of the cosmological parameters obtained from the fits. The final column shows the number of sigma the Ω_m-σ_8 posterior medians are away from the truth, taking into account degeneracies between the parameters.

Survey	Fit model	Mock setup	Ω_m [%]	σ_8 [%]	$N\sigma$
SPT	HMF	Fiducial	1.8	0.7	0.40
SPT	Full	Fiducial	3.8	3.2	1.38
SPT	Full	Baryonic HMF	7.6	5.6	1.36
SPT	Full	Jet	9.3	1.7	0.96
SPT	Full	fgas-8σ	17.3	10.6	1.12
SO	HMF	Fiducial	2.3	0.7	1.51
SO	Full	Fiducial	2.1	1.6	1.74
SO	Full	Baryonic HMF	1.3	1.5	1.43
SO	Full	Jet	2.6	3.1	3.17
SO	Full	fgas-8σ	7.5	2.8	3.54

DMO HMF model, the assumptions of a power-law scaling relation with lognormal scatter, and the uncertainties in baryonic effects on the HMF and on the scaling relation, lead to a biased cosmological inference.

We make use of the FLAMINGO suite of hydrodynamical simulations (Schaye et al., 2023; Kugel et al., 2023). With its large box sizes (at least 1 Gpc on a side) and many variations spanning the uncertainties in the strength of feedback processes constrained by observations of the galaxy mass function and cluster gas fractions, it provides an ideal laboratory to compare cluster count models with. The FLAMINGO simulations self-consistently predict the observables needed for cluster counts, in our case the Compton-Y parameter. We take the predictions of a FLAMINGO simulation as the baseline truth, and test the validity and implications of the different assumptions that go into the theoretical models for cluster counts that are used to measure cosmological parameters from observed cluster counts. We construct three mock surveys based on the number of objects in the Planck (Planck Collaboration et al., 2016a), SPT (Bocquet et al., 2024) and future Simons observatory (Ade et al., 2019) SZ surveys. Since we are focused on uncertainties on the modelling side, we to not create virtual observations and assume pure Poisson errors based on the depth and area of the survey. To identify potential biases, we compare the systematic deviations to the Poisson errors of each survey. We summarise our main results in Fig. 4.7 and Table 4.2. Our main results are:

- There are large and highly significant deviations between the predictions of widely used DMO HMF models. The models differ both from each other and from the $(5.6\ \mathrm{Gpc})^3$ DMO FLAMINGO simulation (see Fig. 4.2). For all three mock surveys, the results from Tinker et al. (2010) and the MiraTitanEmulator (Bocquet et al., 2024) fall outside the Poisson error bars. The model from Bocquet et al. (2016) performs least badly, but it still diverges from FLAMINGO at the high mass end, leading to disagreement for the Planck-like sample, and it is not accurate enough for the SO-like sample.

- The effect of baryons on the HMF will need to be accounted for. This is particularly important for an SO-like survey (see Fig. 4.3), for which even some of our weaker feedback variations lead to deviations that exceed the Poisson errors.

- The assumption of a power-law with lognormal scatter does not lead to large systematic errors, at least not for surveys that are only sensitive to masses $M_{500c} > 2 \times 10^{14}\ \odot$ (see Fig. 4.4). For deeper surveys a single power-law is likely inadequate. For the Planck- and SPT-like samples, the systematic errors due to these assumptions are smaller than Poisson. For the SO-like sample, the assumption of a constant lognormal scatter leads to signif-

icant deviations, which become worse when combined with the power-law assumption.

- Ten per cent uncertainties on the power-law parameters of the observable-mass scaling relation are only good enough for the Planck-like survey (see Fig. 4.5). For SPT and particularly for SO, (much) tighter constraints are required to push the systematic deviations below the Poisson errors. For these surveys current constraints on the scaling relation are insufficiently tight to reduce the systematic deviation below the Poisson errors.

- The effect of changing the scaling relation is similar in magnitude to that of the baryonic modification of the HMF (see Fig. 4.6). For Planck- and SPT-like surveys, only the most extreme models fall outside the Poisson errors. However, for the SO-like sample the differences are large with respect to the Poisson errors. As the changes in the scaling relation due to residual uncertainties in the observationally-constrained galaxy formation physics are at the 10 per cent level, this further reinforces the fact that improving constraints on the observable-mass scaling relation is necessary to ensure the cosmology inference remains unbiased.

- Comparing all the results (see Fig. 4.7 and Table 4.2), we find that the largest source of systematic error on the number counts is the DMO HMF modelling, which leads to a significant bias for each survey. Additionally, for future surveys, improvements will be needed across all the modelling ingredients, for example by introducing a mass- and redshift-dependent scatter and by obtaining better external constraints on the mass-observable relations and the effects of baryons on the HMF.

- When fitting cluster counts with a model assuming a power-law observable-mass scaling relation with lognormal scatter, with Gaussian priors taken from Planck Collaboration et al. (2016a) but centered on the values that best fit our fiducial simulation L1_m9, and the DMO HMF from Bocquet et al. (2016), we find significant errors in the cosmological parameters for all mock surveys (see Fig. 4.8 and Table 4.3). Because cluster counts on their own do not constrain the power-law parameters, any systematic deviations in the prior on the scaling relation lead directly to biases in the cosmological parameters. At the current level of model accuracy, the systematic errors are so large that cluster counts cannot be used to shed light on the σ_8 tension.

We have used the FLAMINGO simulations to highlight some of the shortcomings of current cluster count cosmology models. In particular, the large deviations caused by the assumed theoretical DMO HMF models will need to be addressed as they could lead to large biases even for current surveys. Although

we found that the Bocquet et al. (2016) HMF gives better agreement than popular alternatives, this may be largely coincidental given that we use a different halo finder and that the overestimate of the abundance for masses near the selection limit compensates the strong underestimate for high masses (see Fig. 4.2). The divergence between different models for the HMF at high mass suggests that the abundance of these objects might simply be too sensitive to systematic effects, including simulation volume, halo finder and the mass-concentration relation, to give reliable cosmological constraints. Instead it might be favourable to use objects with a lower mass, something that will happen naturally with new surveys. Improvements to the constraints on the observable-mass scaling relation and the effect of baryons on the HMF will also be necessary for next generation surveys.

Two major steps that we have excluded from this work are the measurement of the observable mass proxy and the selection of clusters based on such measurements. Forward modelling using virtual observations would allow including these steps. This would enable application of full multi-frequency SZ selection methods (Melin et al., 2006, 2012; Hilton et al., 2018) to the FLAMINGO lightcones, including foregrounds and noise specific to each survey. Systematic errors in the calibration of the observable-mass relation using e.g. gravitational lensing could also be modelled using virtual observations.

Another avenue worth exploring, is to abandon the classical HMF based on the spherical overdensity definition of dark matter haloes. Instead, we can predict the abundance of clusters as a function of the observable (e.g. Debackere et al., 2022a). This approach can reduce baryonic uncertainties (e.g. Debackere et al., 2022b) and is well-suited to emulation. Indeed, given that emulators have already replaced (semi-)analytic models for the DMO HMF, this would be a natural next step. Emulators based on hydrodynamic simulations can directly predict cluster counts as a function of observables such as the limiting Compton-Y parameter and the cosmological parameters. Moreover, they can be used to marginalize over the uncertainties in the baryon physics and to constrain such variations using external data.

Acknowledgements

We are grateful to all members of the FLAMINGO collaboration for their help in creating the simulation suite. In particular, we thank Ian McCarthy for his helpful discussions. For this work we made use of the python packages colussus (Diemer, 2018), hmf (Murray et al., 2013), corner (Foreman-Mackey, 2016) and emcee (Foreman-Mackey et al., 2013). This work is partly funded by research programme Athena 184.034.002 from the Dutch Research Council (NWO). VJFM acknowledges support by NWO through the Dark Universe Science Collaboration (OCENW.XL21.XL21.025). This work used the DiRAC@Durham facility managed by the Institute for Computational Cosmology on behalf of the

STFC DiRAC HPC Facility (www.dirac.ac.uk). The equipment was funded by BEIS capital funding via STFC capital grants ST/K00042X/1, ST/P002293/1, ST/R002371/1 and ST/S002502/1, Durham University and STFC operations grant ST/R000832/1. DiRAC is part of the National e-Infrastructure.

Data Availability

The code and data used for this work can be found on GitHub[3]. The HMF and M_{500c}-Compton-Y scaling relation data from FLAMINGO will be shared upon reasonable request to the corresponding author.

[3] https://github.com/Moyoxkit/cluster-counts

References

Abbott T. M. C., et al., 2022, Phys. Rev. D, 105, 023520
Ade P., et al., 2019, J. Cosmology Astropart. Phys., 2019, 056
Akino D., et al., 2022, PASJ, 74, 175
Allen S. W., Evrard A. E., Mantz A. B., 2011, ARA&A, 49, 409
Angulo R. E., Springel V., White S. D. M., Jenkins A., Baugh C. M., Frenk C. S., 2012, MNRAS, 426, 2046
Artis E., Melin J. B., Bartlett J. G., Murray C., Euclid Consortium 2022, in mm Universe @ NIKA2 - Observing the mm Universe with the NIKA2 Camera. p. 00004 (arXiv:2111.05432), doi:10.1051/epjconf/202225700004
Bahé Y. M., et al., 2022, MNRAS, 516, 167
Barnes D. J., et al., 2017, MNRAS, 471, 1088
Black W. K., Evrard A., 2022, MNRAS, 516, 1170
Bleem L. E., et al., 2022, ApJS, 258, 36
Bleem L. E., et al., 2024, The Open Journal of Astrophysics, 7, 13
Bocquet S., Saro A., Dolag K., Mohr J. J., 2016, MNRAS, 456, 2361
Bocquet S., et al., 2019, ApJ, 878, 55
Bocquet S., Heitmann K., Habib S., Lawrence E., Uram T., Frontiere N., Pope A., Finkel H., 2020, ApJ, 901, 5
Bocquet S., et al., 2024, arXiv e-prints, p. arXiv:2401.02075
Booth C. M., Schaye J., 2009, MNRAS, 398, 53
Borrow J., Schaller M., Bower R. G., Schaye J., 2022, MNRAS, 511, 2367
Braspenning J., et al., 2023, arXiv e-prints, p. arXiv:2312.08277
Chaikin E., Schaye J., Schaller M., Benítez-Llambay A., Nobels F. S. J., Ploeckinger S., 2022a, arXiv e-prints, p. arXiv:2211.04619
Chaikin E., Schaye J., Schaller M., Bahé Y. M., Nobels F. S. J., Ploeckinger S., 2022b, MNRAS, 514, 249
Chen K.-F., et al., 2024, arXiv e-prints, p. arXiv:2406.11966
Chiu I. N., Klein M., Mohr J., Bocquet S., 2023, MNRAS,
Costanzi M., et al., 2019, MNRAS, 488, 4779
Cui W., Borgani S., Murante G., 2014, MNRAS, 441, 1769
Cui W., et al., 2018, MNRAS, 480, 2898
Cusworth S. J., Kay S. T., Battye R. A., Thomas P. A., 2014, MNRAS, 439, 2485
Dalla Vecchia C., Schaye J., 2008, MNRAS, 387, 1431
Debackere S. N. B., Schaye J., Hoekstra H., 2021, MNRAS, 505, 593
Debackere S. N. B., Hoekstra H., Schaye J., Heitmann K., Habib S., 2022a, MNRAS, 515, 3383
Debackere S. N. B., Hoekstra H., Schaye J., 2022b, MNRAS, 515, 6023
Diemer B., 2018, ApJS, 239, 35
Diemer B., Joyce M., 2019, ApJ, 871, 168
Elahi P. J., Cañas R., Poulton R. J. J., Tobar R. J., Willis J. S., Lagos C. d. P., Power C., Robotham A. S. G., 2019, PASA, 36, e021
Elbers W., Frenk C. S., Jenkins A., Li B., Pascoli S., 2021, MNRAS, 507, 2614
Elbers W., Frenk C. S., Jenkins A., Li B., Pascoli S., 2022, MNRAS, 516, 3821
Euclid Collaboration et al., 2023, A&A, 671, A100

Euclid Collaboration et al., 2024, arXiv e-prints, p. arXiv:2405.13491
Evrard A. E., Arnault P., Huterer D., Farahi A., 2014, MNRAS, 441, 3562
Foreman-Mackey D., 2016, The Journal of Open Source Software, 1, 24
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Ghirardini V., et al., 2024, arXiv e-prints, p. arXiv:2402.08458
Hahn O., Rampf C., Uhlemann C., 2021, MNRAS, 503, 426
Hamana T., Shirasaki M., Lin Y.-T., 2020, PASJ, 72, 78
Han J., Jing Y. P., Wang H., Wang W., 2012, MNRAS, 427, 2437
Han J., Jing Y. P., Wang H., Wang W., 2017, HBT: Hierarchical Bound-Tracing, Astrophysics Source Code Library, record ascl:1711.022
Heitmann K., et al., 2016, ApJ, 820, 108
Heymans C., et al., 2021, A&A, 646, A140
Hilton M., et al., 2018, ApJS, 235, 20
Hirschmann M., Dolag K., Saro A., Bachmann L., Borgani S., Burkert A., 2014, MNRAS, 442, 2304
Huško F., Lacey C. G., Schaye J., Schaller M., Nobels F. S. J., 2022, MNRAS, 516, 3750
Ivezić Ž., et al., 2019, ApJ, 873, 111
Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, MNRAS, 321, 372
Klein M., et al., 2024, MNRAS,
Kugel R., et al., 2023, MNRAS, 526, 6103
Kugel R., Schaye J., Schaller M., McCarthy I. G., Braspenning J., Helly J. C., Forouhar Moreno V. J., McGibbon R. J., 2024, arXiv e-prints, p. arXiv:2406.03180
Le Brun A. M. C., McCarthy I. G., Schaye J., Ponman T. J., 2014, MNRAS, 441, 1270
Liu A., et al., 2022, A&A, 661, A2
Liu A., et al., 2024, arXiv e-prints, p. arXiv:2402.08454
Lovisari L., Reiprich T. H., Schellenberger G., 2015, A&A, 573, A118
Mantz A. B., 2019, MNRAS, 485, 4863
McCarthy I. G., Schaye J., Bird S., Le Brun A. M. C., 2017, MNRAS, 465, 2936
McCarthy I. G., et al., 2023, MNRAS, 526, 5494
Melin J. B., Bartlett J. G., Delabrouille J., 2006, A&A, 459, 341
Melin J. B., et al., 2012, A&A, 548, A51
Melin J. B., Bartlett J. G., Cai Z. Y., De Zotti G., Delabrouille J., Roman M., Bonaldi A., 2018, A&A, 617, A75
Miyatake H., et al., 2023, arXiv e-prints, p. arXiv:2304.00704
Murray S. G., Power C., Robotham A. S. G., 2013, Astronomy and Computing, 3, 23
Nelson D., Pillepich A., Ayromlou M., Lee W., Lehle K., Rohr E., Truong N., 2023, arXiv e-prints, p. arXiv:2311.06338
Pacaud F., et al., 2018, A&A, 620, A10
Pakmor R., et al., 2023, MNRAS, 524, 2539
Pillepich A., et al., 2018, MNRAS, 473, 4077
Planck Collaboration et al., 2014a, A&A, 571, A20
Planck Collaboration et al., 2014b, A&A, 571, A21
Planck Collaboration et al., 2016a, A&A, 594, A24
Planck Collaboration et al., 2016b, A&A, 594, A27
Planck Collaboration et al., 2020a, A&A, 641, A6
Planck Collaboration et al., 2020b, A&A, 641, A8
Ploeckinger S., Schaye J., 2020, MNRAS, 497, 4857
Pratt G. W., Arnaud M., Biviano A., Eckert D., Ettori S., Nagai D., Okabe N., Reiprich T. H., 2019, Space Sci. Rev., 215, 25
Riess A. G., et al., 2022, ApJ, 934, L7
Rozo E., Bartlett J. G., Evrard A. E., Rykoff E. S., 2014, MNRAS, 438, 78
Salvati L., et al., 2022, ApJ, 934, 129
Schaller M., et al., 2024, MNRAS, 530, 2378

Schaye J., Dalla Vecchia C., 2008, MNRAS, 383, 1210
Schaye J., et al., 2010, MNRAS, 402, 1536
Schaye J., et al., 2015, MNRAS, 446, 521
Schaye J., et al., 2023, MNRAS, 526, 4978
Springel V., Di Matteo T., Hernquist L., 2005, MNRAS, 361, 776
Sunyaev R. A., Zeldovich Y. B., 1972, Comments on Astrophysics and Space Physics, 4, 173
Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709
Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, ApJ, 724, 878
Velliscig M., van Daalen M. P., Schaye J., McCarthy I. G., Cacciato M., Le Brun A. M. C., Dalla Vecchia C., 2014, MNRAS, 442, 2641
Wiersma R. P. C., Schaye J., Theuns T., Dalla Vecchia C., Tornatore L., 2009, MNRAS, 399, 574
Zubeldia Í., Chluba J., Battye R., 2023, MNRAS, 522, 5123

5 Cosmic Web Dynamics: Forces and Strains

Authors: **Roi Kugel** and Rien van de Weygaert
submitted to MNRAS

Abstract

This study concerns an inventory of the gravitational force and tidal field induced by filaments, walls, cluster nodes and voids on Megaparsec scales and how they assemble and shape the Cosmic Web. The study is based on a $N_{Part} = 512^3$ ΛCDM dark matter only N-body simulation in a $(300h^{-1}$ Mpc$)^3$ box at $z = 0$. We invoke the density field NEXUS+ multiscale morphological procedure to assign the appropriate morphological feature to each location. We then determine the contribution by each of the cosmic web components to the local gravitational and tidal forces. We find that filaments are, by far, the dominant dynamical component in the interior of filaments, in the majority of underdense void regions and in all wall regions. The gravitational influence of cluster nodes is limited, and they are only dominant in their immediate vicinity. The force field induced by voids is marked by divergent outflowing patterns, yielding the impression of a segmented volume in which voids push matter towards their boundaries. Voids manifest themselves strongly in the tidal field as a cellular tapestry that is closely linked to the multiscale cosmic web. However, even within the interior of voids, the dynamical influence of the surrounding filaments is stronger than the outward push by voids. Therefore, the dynamics of voids cannot be understood without taking into account the influence of the environment. We conclude that filaments constitute the overpowering gravitational agent of the cosmic web, while voids are responsible for the cosmic web's spatial organisation and hence of its spatial connectivity.

5.1 Introduction

This study concerns a systematic inventory of the gravitational force and tidal field on Megaparsec scales and its role in determining the structure of the Cosmic Web. We assess the force field and tidal field induced by filaments, walls, cluster nodes and voids, and assess in how far they contribute, and dominate, the gravity and tides in the various regions of the cosmic web. It allows us to

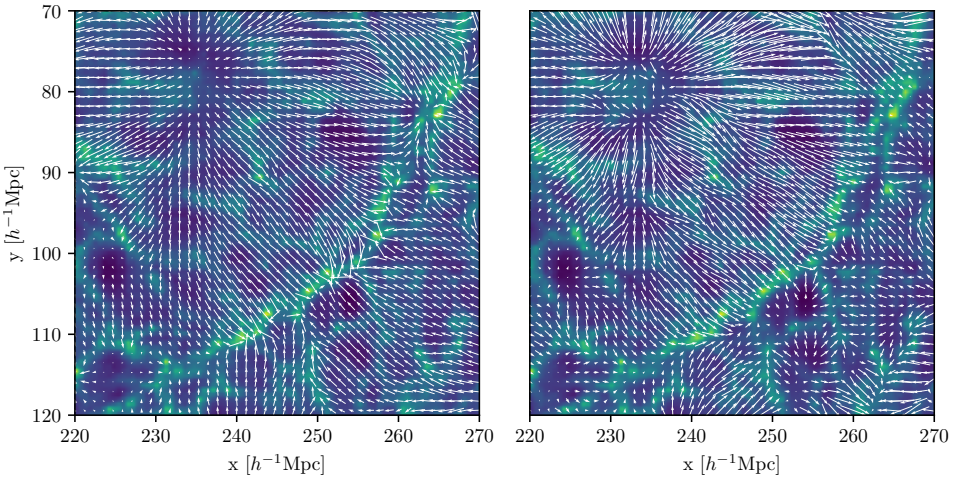


Figure 5.1: Gravitational force field in and near a void. Left: the force vectors representing the local direction and amplitude of the total gravitational force, ie. the force exerted by the total cosmic mass distribution. Right: the local force vectors resulting from the combined influence of mass elements in void regions. The vector lengths are normalised with respect to the local total velocity and show the relative contribution of the field at that location.

5

investigate the question which morphological features in the large scale universe dominate - and drive - the gravitationally driven formation and evolution of the largest structure in the universe. Also, as this will depend to a considerable extent on location, we include a systematic inventory of the identity of the regions over which voids, filaments and clusters dominate the the gravitational and tidal force. In an accompanying study, we specifically focus on the dynamical influence of cosmic voids in the large scale matter distribution, which represent the major share of the cosmic volume and who - along with filaments - dominate the dynamics of the large scale Universe.

5.1.1 The Cosmic Web: structure and detection

The *Cosmic Web* is the intricate multiscale network defined by the matter and galaxy distribution on Megaparsec scale (Zeldovich, 1970; Einasto, 1977; Bond et al., 1996; van de Weygaert & Bond, 2008a; Cautun et al., 2014). It represents the fundamental spatial organisation of matter on scales of a few up to a hundred Megaparsec. Galaxies, intergalactic gas and dark matter arrange themselves in a salient wispy pattern of dense compact clusters, long elongated filaments, and sheetlike tenuous walls surrounding near-empty void regions. Filaments are the most visually outstanding features of the Megaparsec Universe, in which around

50% of the mass and galaxies in the Universe resides. On the other hand, almost 80% of the cosmic volume belongs to the interior of voids (see e.g. Cautun et al., 2014; Ganeshaiah Veena et al., 2018). Together, they define a complex spatial pattern of intricately connected structures, displaying a rich geometry with multiple morphologies and shapes. This complexity is considerably enhanced by its intrinsic multiscale nature, including objects over a considerable range of spatial scales and densities. The connectivity of this rich pallet of features, the nature of how the various structures connect to establish the pervasive network, has only recently been recognised as an important defining - topological - aspect (Aragón-Calvo et al., 2010; Shim et al., 2021; Wilding et al., 2021; Feldbrugge & van de Weygaert, 2023). For a recent up-to-date report on a wide range of relevant aspects of the cosmic web, we refer to the volume by van de Weygaert et al. (2014).

In the observational reality, the existence and structure of the Cosmic Web has been revealed in the most detail by maps of the nearby cosmos produced by large galaxy redshift surveys. Starting from first revelation of the web-like arrangement of galaxies by the CfA2 survey (e.g. de Lapparent et al., 1986), subsequent surveys such as 2dFGRS, the SDSS, the 2MASS and GAMA redshift surveys (Colless et al., 2003; Tegmark et al., 2004; Huchra et al., 2012; Liske et al., 2015) established the web-like arrangement of galaxies as a fundamental characteristic of cosmic structure. Maps of the galaxy distribution at larger cosmic depths, such as VIPERS (de la Torre et al., 2013), showed its existence over a sizeable fraction of cosmic time.

Also the intergalactic gaseous medium, closely follows the web-like structure defined by the dark matter, the principal component of the cosmic web. A range of observational probes have detected the web-like structure over which intergalactic gas, in a range of thermodynamic states (see Meiksin, 2009, for a review), has diffused itself. Lyα absorption lines in the spectra of bright background sources such as QSOs are piercing through the web-like assembly of neutral hydrogen gas in the cosmic web at high redshifts (Ostriker & Cen, 1996; Cen, 1997). The combination of sufficiently close linear probes even allows a reconstruction of the full three-dimensional intergalactic hydrogen lanes (Pichon & Bernardeau, 1999).

It has already led the Clamato survey (Lee et al., 2018) to successfully produce fascinating maps of the full three-dimensional gaseous cosmic web at high redshifts. Recent observations by the MUSE integral field unit on the very large telescope, even managed to see the Lyα emission from the filamentary gaseous extensions around clusters directly. At lower redshifts, most of the intergalactic gas has heated up as it settled in the deepening potential wells of the dark matter cosmic web. This warm gas, the so called WHIM, is assumed to represent the major share of baryons in the current Universe. As such is a prime target for detection and mapping (Nicastro et al., 2018; Macquart et al., 2020), although it

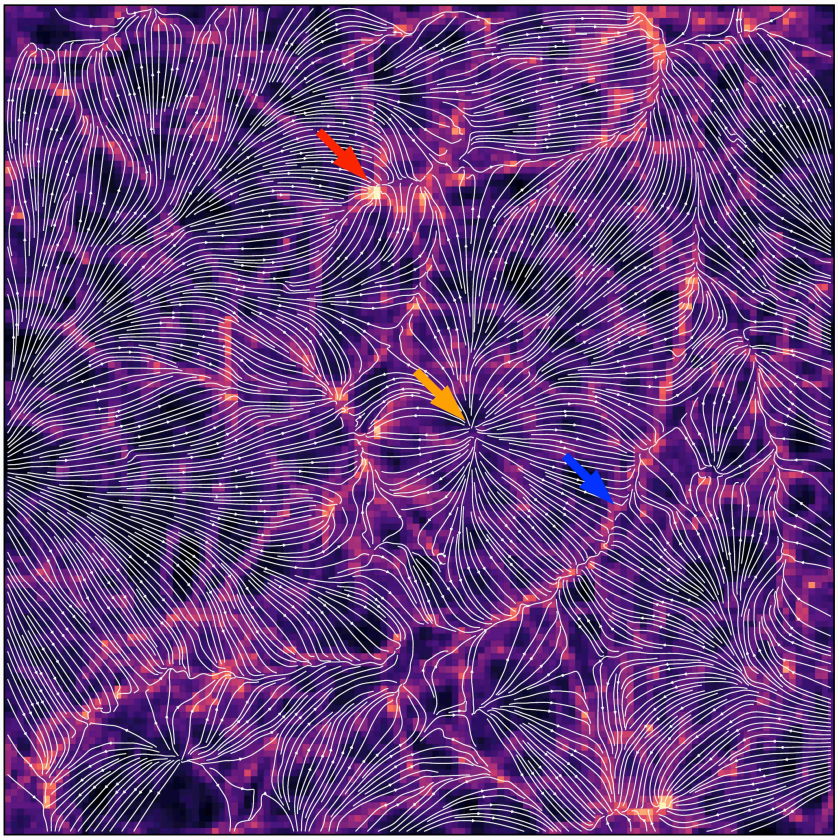


Figure 5.2: Flows of matter in the cosmic web. Superimposed on the (logarithmic) map of the cosmic density field in a slice of the simulation box are the streamlines of the corresponding cosmic velocity field. Arrows indicate three different aspects of the flow field: The red arrow highlights velocity inflow into a cluster node (negative divergence), the orange arrow highlights velocity outflow from a void (positive divergence) and the blue arrow highlights shear flow along a filament.

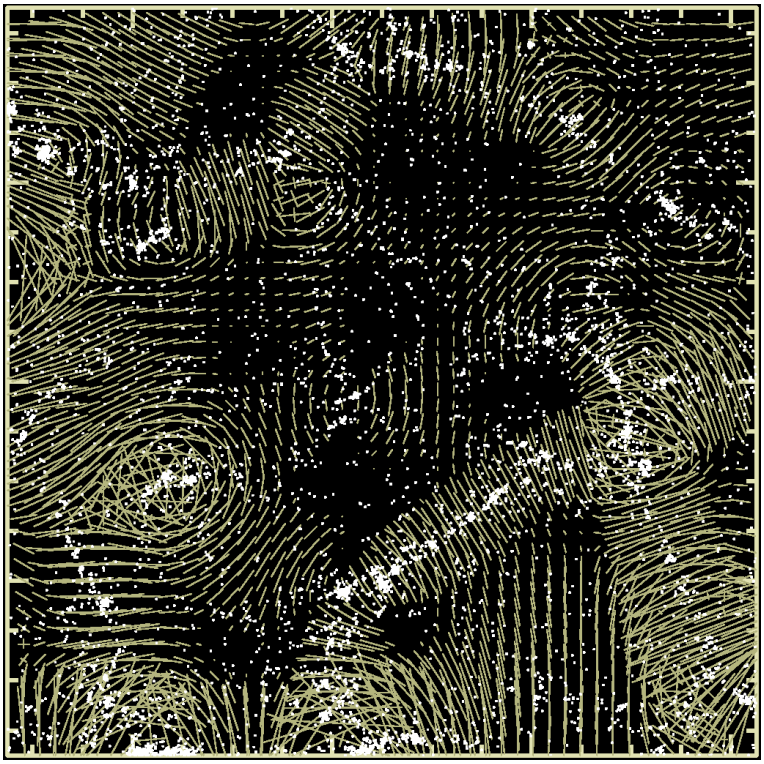


Figure 5.3: Cosmic web and tidal force field: the relation between the cosmic web and the spatial pattern of the corresponding compressional tidal field components. The image shows an LCDM matter distribution at the present cosmic epoch, along with the (compressional component) tidal bars in a a $5h^{-1}$ Mpc thin central slice. The simulation is a realization of a dark matter only CDM based scenario (in an open, $\Omega = 0.3$ Universe. The tides are determined on a scale $R_G = 2h^{-1}$ Mpc). The matter distribution, displaying a pronounced weblike geometry, is clearly intimately linked with the characteristic coherent compressional tidal bar pattern.

had proven to be notoriously hard to detect. It has been more straightforward to detect the hot gas residing in the strongest filamentary features in the cosmic web at high redshifts, filling the short dense bridges between two adjacent clusters. The hot gas reveals itself through the Sunyaev-Zeldovich upscattering of CMB photons. It has even allowed the detection of a few individual high-redshift gaseous filaments, where their ubiquitous presence has been revealed by stacking numerous cluster pair Sunyaev-Zeldovich observations (Bonjean et al., 2018; de Graaff et al., 2019).

5.1.2 The Cosmic Web: dynamics, formation and evolution

The origin of the cosmic web is to be found in the tiny inhomogeneities in the primordial spatial matter distribution. These induce an inhomogeneous and anisotropic gravitational force field. It sets off the migration of matter, culminating in the continuous growth of density and velocity inhomogeneities (Peebles, 1980). Ultimately, as the matter fluctuations become nonlinear, it leads to the contraction and collapse of structures and expansion of underdense regions. At the transition from the initial long phase of linear growth of the primordial (Gaussian) density field and density field to the complex structured nonlinear mass distribution, we find the emergence of the intriguing and complex spatial pattern of the *Cosmic Web*. It marks the cosmic web as a key phase in the dynamical buildup of structure in the Universe.

A plethora of cosmological computer simulations (White et al., 1987; Springel, 2005; Pillepich et al., 2018; Bocquet et al., 2020; Angulo et al., 2021), including more elaborate cosmological hydrodynamics simulations (e.g. Hirschmann et al., 2014; Schaye et al., 2015; McCarthy et al., 2017; Pillepich et al., 2018; Schaye et al., 2023) revealed that through gravitationally driven evolution the Gaussian initial density and velocity perturbations morph into an intricate web-like pattern that resembles that seen in the spatial distribution of galaxies. The simulations show that the cosmic web is a fundamental aspect of structure formation in the standard ΛCDM cosmology, although this also turns out to be true for variations around the standard ΛCDM model (for a review see Frieman et al., 2008).

Within the context of the gravitational buildup of the cosmic web, the gravitational forces induced by the inhomogeneous mass distribution are the agent for the buildup of structure. They are instrumental in directing the cosmic migration streams that transport matter from low density areas to emerging matter concentrations. The accompanying tidal force field is the key towards shaping the mass distribution into a weblike pattern. This had already been recognized and accurately described in the mildly nonlinear stage by the Zel'dovich formalism (Zeldovich, 1970; Shandarin & Zeldovich, 1989). The formation and evolution of the characteristic anisotropic structures, ie. the filaments and walls, are the product of the anisotropic tidal strains and resulting anisotropic flow field and deformations (Bond et al., 1996; Hahn et al., 2007; van de Weygaert & Bond, 2008a; Lee et al., 2009; Lee & Springel, 2010; Hahn et al., 2010; Wang et al., 2014; Feldbrugge et al., 2018a; Feldbrugge & van de Weygaert, 2024). It establishes a close relation between the weblike structures and the anisotropy of the induced migration flows, which has been the focus of a few insightful studies (Kitaura et al., 2012a; Hoffman et al., 2012; Wang et al., 2014). Recent work has also revealed the extent to which the spatial structure of the tidal force field determines the connections between the components of the cosmic web. The connectivity

can even already be recognised in the primordial tidal and deformation field, showing the extent towards which the primordial anisotropic force field is steering and shaping the structure of the emerging cosmic web (see eg. Wilding, 2022; Feldbrugge et al., 2023; Feldbrugge & van de Weygaert, 2024). It even leads to the realization that embryonic outline of the cosmic web, in particular its filamentary network, can already be seen in the primordial tidal eigenvalue field (Wilding, 2022; Feldbrugge et al., 2023) (see figure 5.3).

For insight into the nature and origin of the characteristic properties of the cosmic web and for identifying the dependence of these on cosmology and cosmological parameters, theoretical understanding of the physical mechanisms and processes behind the emergence of the cosmic web. Analytical approximation and model have been essential in interpreting the results of surveys, as well as of simulations (e.g. Zeldovich, 1970; Hidding et al., 2012; Wang et al., 2014; Hidding et al., 2016; Feldbrugge et al., 2018a; Feldbrugge & van de Weygaert, 2024). An important aspect of the formation process are the forces and strains that shaped cosmic structure. For example, the phase-space based *Caustic Skeleton* model by Feldbrugge et al. (2018a) demonstrated that a full understanding of the cosmic web structure is obtained through the spatial characteristics of the *eigenvalue* and *eigenvector* fields of the cosmic tidal force field. Motivated by these consideration, the intention of the present investigation and inventory of the dynamics of the cosmic web is to establish the role of the various morphological features of the cosmic web in its formation and dynamical evolution. It involves the assessment of in how far the various elements of the cosmic web are formed and have evolved, and how they connect up in the complex, intricate pervasive network of the cosmic web.

While many structural aspects of the cosmic web have been addressed by numerous studies, its dynamics and dynamical evolution has only been - superficially - explored within a purely theoretical or simulation context. Over recent years, we have seen that the dynamics behind the formation and evolution of the cosmic web is becoming increasingly accessible to observational investigation (see eg. Kitaura et al., 2012b). The forces and tides that shape the complex spatial pattern induce nonlinear migration currents, marked by distinct divergent and shear-like flows, have recently been traced in new, densely probed, galaxy peculiar velocity surveys. The most notable examples of this are the Cosmicflows-3 and Cosmicflows-4 (Tully et al., 2016; Kourkchi et al., 2020) surveys, a view that will be considerably extended by DESI (Schlafly et al., 2023). Including the peculiar velocity information even allowed to identify the impact of some individual components of the cosmic web is also found in the local universe. This concerns in particular the gravitational influence of voids: Tully et al. (2008) concluded that the Local Void has a large contribution of no less than $\sim$ 240km/s to the peculiar velocity of the Local Group.

Also the subtle morphing influence of the tidal forces has become suscepti-

ble to observational scrutiny. The distinct anisotropic shape of the filaments and walls in the cosmic web is the direct outcome of the gravitationally driven formation of the cosmic web by the large scale anisotropic force field (van de Weygaert & Bond, 2008a). Even the population of voids on Megaparsec scales is noticeably shaped and aligned by the large scale tidal force field (Park & Lee, 2007; Platen et al., 2008).

Arguably the most prominent manifestation of the (tidal) anisotropic force field induced by the inhomogeneous matter distribution in the Universe is that of gravitational lensing. Hence, at a global level the tidal dynamics of the cosmic web is reflected in the corresponding deformation of galaxy images. In case the distortions are linear, inversion allows the study of the generating (projected) mass distribution. At present, gravitational lensing has become one of the most powerful probes of the global Universe and the cosmological parameters characterising it. With the availability of the upcoming powerful and accurate cosmological surveys, such as enabled by the Vera Rubin observatory and Euclid, the tedious lensing studies may even resolve the web-like nature of dark matter distribution. A few gravitational lensing studies have even already managed to detect and resolve filamentary dark matter bridges between nearby massive clusters (Dietrich et al., 2012). Amongst the most massive representatives amongst the filament population (Cautun et al., 2014), they leave a rather accessible and detectable lensing imprint.

Also at galaxy scales, the large scale tidal forces induce noticeable signatures (see e.g. van de Weygaert & Babul, 1994; Paranjape, 2021; Alam et al., 2024). Perhaps the most outstanding manifestation is that of the alignment of the spin axis of collapsing dark matter halos with the filaments in which they are embedded, and hence that of the corresponding rotation axis of galaxies. It is clear from observations (see Jones et al., 2010; Tempel et al., 2013; Welker et al., 2020) that the rotation axis of galaxies preferentially align with the components of the cosmic web. Largely the result of the imparted tidal torques on the collapsing halos (Hoyle, 1949; Peebles, 1969; Efstathiou & Jones, 1980; White, 1984; Lee & Pen, 2000; Porciani et al., 2002a,b; Schäfer, 2009), additional secondary effects responsible for the mass dependence of the spin orientation (Aragón-Calvo et al., 2010; Hahn et al., 2007, 2010; Codis et al., 2012; Tempel et al., 2013; Ganeshaiah Veena et al., 2018, 2019; López et al., 2021; Zhang et al., 2023) may be the result of the induced filamentary inflow of mass (Ganeshaiah Veena et al., 2021). The latter has become a key issue of attention, given the implications for gravitational lensing studies.

5.1.3 Morphology of the Cosmic Web

While the cosmic web has four morphologically well defined features, there are many ways to identify the different components. Over the past decades, a range

of methods and formalisms have been put forward for the detection and classification of filaments. A review and comparison of more than a dozen formalisms can be found in Libeskind et al. (2018). They distinguish at least five classes of formalisms to classify and analyse the cosmic web. Geometric filament finders are usually based on the Hessian of the density or gravitational potential at each location. It includes the Tweb and Vweb formalisms (Hahn et al., 2007; Forero-Romero et al., 2009; Hoffman et al., 2012), which emphasize and exploit the intimate link between the tidal foce field and induced anisotropic velocity flows and the spatial structure and connectivity of the weblike pattern that is emanating as a result of these. The most sophisticated ones explicitly take into account the multiscale nature of the mass distribution, of which MMF/Nexus is a particular example (Aragón-Calvo et al., 2010; Cautun et al., 2013). Topological methods address the connections between structural singularities in the mass distribution. They are amongst the most widely used formalisms in the studies of the cosmic web, in particular, that of Disperse (Sousbie, 2011; Sousbie et al., 2011). Other representatives are Spineweb (Aragón-Calvo et al., 2007a) and Felix (Shivashankar et al., 2016).

In addition to the geometric and topological formalisms, several alternative methods have played a substantial role in the study of the cosmic web. Bisous is a well-known stochastic formalism, involving Bayesian exploration based on stochastic geometric modelling of filaments (Tempel et al., 2014a). It forms the basis for the widely used filament catalogue extracted from the SDSS survey (Tempel et al., 2014b). More recent developments often incorporate machine learning codes (see *e.g.* Awad et al., 2023). Perhaps the oldest representatives for a systematic analysis of filamentary patterns are graph-based methods. The Minimal Spanning Tree (MST) is a prime example and is figuring prominently in the cosmic web analysis of the GAMA survey (Alpaslan et al., 2014). There is even a class of cosmic web identifiers that exploit the resemblance of the cosmic web to biological branching networks. The Monte Carlo Physarum Machine, inspired by the growth of Physarum polycephalum 'slime mold', has been successfully applied to the structural analysis of the cosmic web in both simulations and observations (Elek et al., 2020, 2022; Wilde et al., 2023).

Possibly the most profound techniques for classification of the cosmic web are those emanating from the analysis of the 6D phase-space structure of the cosmic mass distribution (Shandarin, 2011; Shandarin et al., 2012; Neyrinck & Shandarin, 2012; Abel et al., 2012). Restricted to situations in which the initial conditions are known, they yield the identification of the matter streams constituting the migration of mass in the buildup of structure. It allows the definition of an objective physical criterion for what constitutes the various structural elements of the cosmic web. The recent study by Feldbrugge & van de Weygaert (2024) on the phase-space based dynamical specification of the nature of cosmic filaments emanates from a detailed phase-space based assessment of cosmic structure for-

mation, within the context of the Caustic Skeleton model for the formation of the cosmic web (Feldbrugge et al., 2018b).

The present study is based on the MMF/Nexus morphology classification formalism(Aragón-Calvo et al., 2007b,a, 2010; Cautun et al., 2013, 2014), that is unique in addressing both the geometric nature as well as the multiscale nature of the cosmic matter distribution. By means of the Hessian of the density field NEXUS to find the geometric/morphological signature of the individual components. Of particular interest for our purpose is the NEXUS+ flavour of NEXUS. It applies log-space filtering to the density field before identification. Like shown by Cautun et al. (2014) this allows NEXUS+ to find even the more tenuous elements of the cosmic web.

5.1.4 Cosmic Web Dynamical Inventory

The present study extends the previous cosmic web inventory of Cautun et al. (2014) to that of the corresponding force and tidal field influence as a function of cosmic web environment. The distinct structure of the morphological elements of the cosmic web implies that each component will also be dynamically distinct, both in its local dynamics, but also in the influence it will have on its surroundings through the force and tidal field. Bond et al. (1996) argue that the main origin of filaments in the cosmic web is due to the distribution of nodes in the cosmic web, and not much work has been done on the topic since. In this work we aim to extend our knowledge about the dynamical influence and origin of the cosmic web.

To quantify the dynamical influence, we analyse a ΛCDM (dark matter only) simulation, and asses at each location in a cosmic volume the contribution by filaments, walls, voids and nodes to the local gravitational and tidal force (see eg. fig. 5.1 for the force field contribution by voids). This allows a statistical inventory of the dynamical dominance of the various morphological elements of the cosmic web. To this end we apply NEXUS+ to identify for each point in the simulation whether it belongs to a filaments, wall, void or node. Using direct summation over the gridcells, we are then able to construct the force and tidal field originating from each component.

The obtained fields are full vector fields, for the force field, or matrix fields, for the tidal field. This allows us to study both the morphology and the strength of the fields belonging to each component. For each field we look at both the (relative) amplitude and the direction. In addition, the spatial coherence and pattern of the induced flow field is studied on the basis of streamline maps (see fig. 5.2). We are also able to zoom in on a few specific regions to study the interplay between the different components. In this way we are able to find where, and how the different components contribute to the total force field. By looking at the

tidal field we are also able assess which component is responsible for creating the an-isotropic nature of the cosmic web (see fig. 5.3).

5.1.5 Outline

This paper is structured as follows. In section 5.2 we describe the simulation we use, how we obtain the identification of the cosmic web components and how we calculate the force and tidal fields. In section 5.4 we describe the force field of the individual components for the complete $(300\text{Mpc})^2$ slice. In section 5.5 we investigate two smaller regions of the full slice, and look at the force fields around a void and a node. In section 5.6 we describe our results for the tidal field. We finish by summarising our conclusions in section 5.8.

5.2 Data and Analysis: N-body simulations

For the dynamical inventory of the cosmic web in the present study, we analyze a LCDM dark matter only cosmological N-body simulation. The force and tidal inventory in the present study largely pertains to the current cosmic epoch, redshift $z = 0$. The evolution of the force and tidal force fields will be analyzed in an accompanying upcoming publication. In this section we describe the specific methods and formalisms used to get to the inventory of the gravitational force and tidal field within this simulation.

The simulation is processed and analyzed with a set of instruments that allow the optimal identification of the multiscale structure and dynamics of the cosmic web. These include the translation from the discrete set of particle position and velocities towards continuous density and flow fields that optimally retains the anisotropic pattern of the cosmic web as well as its multiscale structure. This is accomplished through the use of the DTFE formalism (Schaap & van de Weygaert, 2000; van de Weygaert & Schaap, 2009; Cautun & van de Weygaert, 2011).

Instrumental in our study is the classification and identification of the morphological components of the cosmic web. The morphological identification and classification of the geometric components of the cosmic web is based on the use of the MMF/Nexus formalism, specifically that of NEXUS+ version of the MMF/Nexus pipeline (Aragón-Calvo et al., 2007a; Cautun et al., 2013).

The second major aspect of our analysis concerns the calculation of the (peculiar) gravity and tidal fields, globally as well specific for those induced by the individual components of the cosmic web. The computational details are described in section 5.3.1. To appreciate the spatial structure of the force and tidal fields better, and to interpret the results obtained, in section 5.3.1 we include a description of the various aspects of visualisation of these fields.

The necessary details of the Nbody simulation and applied analysis tools are outlined below.

5.2.1 N-body Simulation

The study is based on the analysis of the N-body, dark matter only simulation described in Bos (2016). The cosmological context is that of a ΛCDM cosmology with WMAP 3-year parameter values: $\Omega_{\rm m} = 0.268$, $\Omega_\Lambda = 0.732$, $\Omega_{{\rm b}Zel} = 0.044$, $h = 0.704$, $\sigma_8 = 0.776$ and $n = 0.947$. The simulation contains 512^3 dark matter particles in a simulation box with a boxlength of 300 h^{-1} Mpc. The simulations start at $z = 60$, following the initial displacement and velocity set by the Zeldovich approximation (Zeldovich, 1970). The analysis described in the present study focuses on the on the inventory of the force and tidal field at the current epoch, $z = 0$.

5.2.2 DTFE Density and Velocity maps

To translate the spatial particle distribution into density and velocity maps, we convert the particle locations and velocities into a space-filling DTFE density and velocity maps. The density Δ,

$$\Delta = \frac{\rho}{\rho_u} = \delta + 1 \tag{5.1}$$

and velocity field $\mathbf{v}_{\rm DTFE}$ are sampled on a 512^3 grid. It corresponds to a resolution of $0.59h^{-1}$Mpc for each grid cell.

The Delaunay tessellation field estimator (DTFE) (Schaap & van de Weygaert, 2000; van de Weygaert & Schaap, 2009; Cautun et al., 2013), uses the Voronoi and Delaunay tessellation of the spatial particle distribution to get an unbiased volume-weighted estimate of the local density including a natural interpolation over the entire volume using the dual Delaunay tessellation as natural interpolation grid. Application of DTFE to the velocity field, traced by the simulation particles, not only yields a volume-weighted space covering reconstruction of the velocity field but also yields maps of the first order gradient of the flow field, ie. the divergence and shear (and even vorticity) of the flow field (Bernardeau et al., 1997; van de Weygaert & Bernardeau, 1998; Romano-Díaz & van de Weygaert, 2007; van de Weygaert & Schaap, 2009). THE DTFE density and velocity fields are sampled on a 512^3 grid.

5.2.3 MMF/NEXUS+

To dissect the cosmic matter distribution into the various morphological components of the cosmic web we use the MMF/NEXUS pipeline, specifically its NEXUS+ version (Cautun et al., 2013, 2014). It is the highest dynamic range version of the NEXUS library of routines for pattern classification based on the MMF/Nexus Multiscale Morphology Filter formalism (Aragón-Calvo et al., 2007a,b, 2010; Cautun et al., 2013, 2014). For a short review see appendix 5.A, as

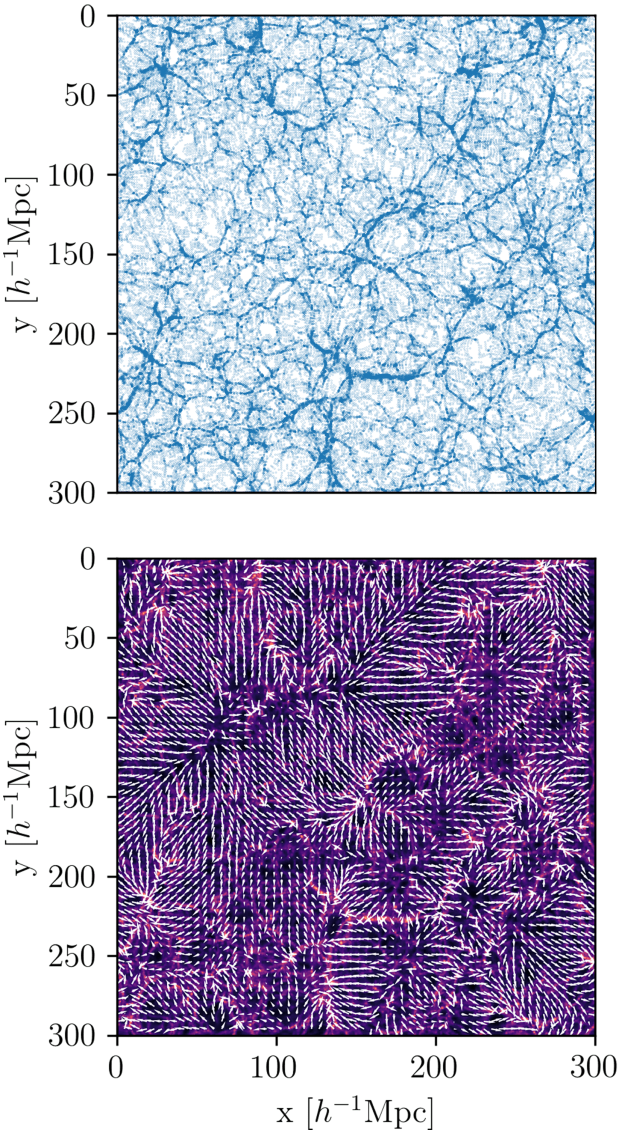


Figure 5.4: Cosmic Mass distribution and force field. Top panel: the particle distribution, at $z = 0$ in a 0.59 h^{-1}Mpc thick slice from a LCDM N-body simulation in a 300 h^{-1}Mpc by 300 h^{-1}Mpc box. Bottom panel: the corresponding gravity force vector plot in the same slice.

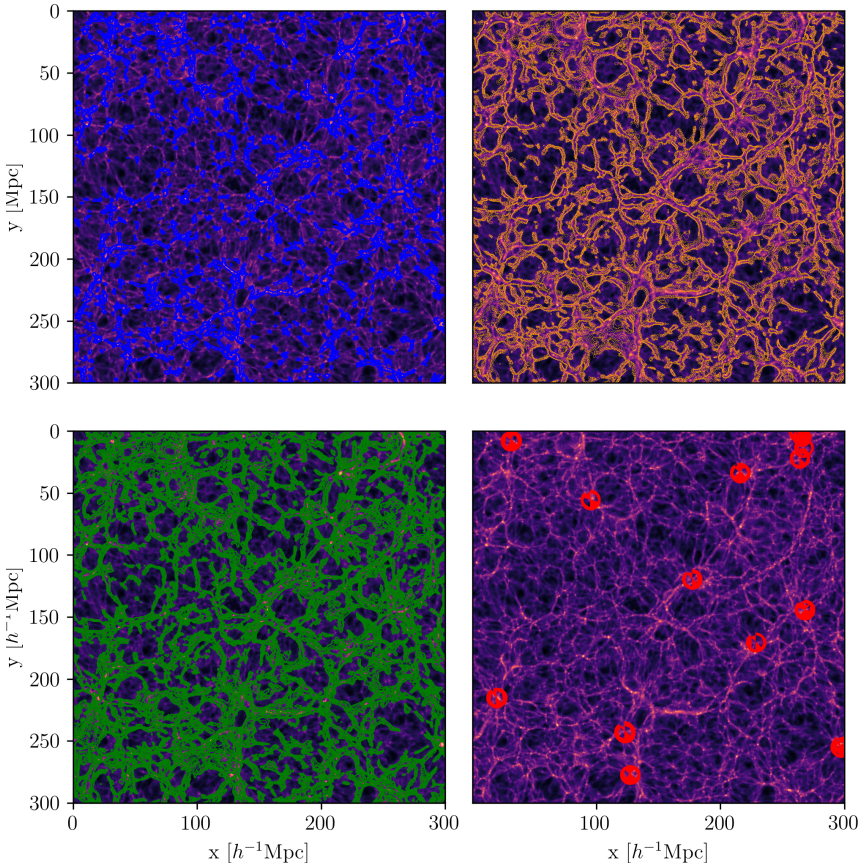


Figure 5.5: Cosmic web components identified by Nexus+: filaments, voids, walls and cluster nodes. Each panel shows a representation of a 300 h^{-1}Mpc by 300 h^{-1}Mpc by 0.59 h^{-1}Mpc slice from the N-body simulation used. Top left panel: Filaments. The NEXUS+ filament contours, in blue, are superimposed on the log density field. Top right panel: Voids. The NEXUS+ void boundary contours, in orange, are superimposed on the log density field. Bottom left panel: Walls. the NEXUS+ wall contours, in green, are superimposed on the log density field. Bottom right panel: Cluster nodes. The cluster nodes, indicated by red symbols, are superimposed on the log density field.

well as Libeskind et al. (2018). We specifically chose to use the NEXUS+ version as it is optimally suited to take into account the multiscale nature and the wide dynamic range of cosmic web density field (Cautun et al., 2014). Compared with the other NEXUS methods it provides the sharpest rendering of the cosmic web, including the more tenuous walls and filaments.

Instrumental and unique for the class of MMF cosmic web identification methods is that it simultaneously pays heed to two principal aspects characterising the web-like cosmic mass distribution. The first aspect is that of the different geometric shapes of the various structural components of the cosmic web, in particular the anisotropic shapes of filaments and walls. The Hessian of the corresponding fields translates into the local geometry and local anisotropy on the basis of the ratios of its eigenvalues. That is, it determines whether the mass distribution has a roundish, flattened or elongated geometry. The second aspect is that of the multiscale character of the cosmic mass distribution, the product of the hierarchical evolution and buildup of structure in the Universe. To this end, MMF/Nexus invokes a *Scale-Space* analysis to enable it to probe the scale dependence of the local geometry of the mass (or velocity or gravity) distribution. It allows the detection, and complete and unbiased characterisation, of features present at all scales, from the prominent structures present in overdense regions to the tenuous networks pervading the cosmic voids.

The MMF/Nexus formalism results in a scale adaptive framework that classifies the matter distribution on the basis of local spatial variations in either the density field, velocity field or gravity field. Their geometry and anisotropy are encoded in the Hessian matrix determined at each of the spatial scales at which the fields are assessed. Subsequently, a set of morphological filters is used to classify the spatial matter distribution into three basic components, the clusters, filaments and walls that constitute the cosmic web. It produces a map in which for each location in the analysed volume the morphological identity is specified. The end product of the pipeline is a multiscale identification of the cosmic web into its structural components. To this end, each location and mass element of the matter distribution is assigned a unique morphological identity as either belonging to a filament, cluster node, wall or void.

For more details on the MMF/Nexus formalism we refer to appendix 5.A.

NEXUS+: practical implementation

While in principle the NEXUS *Scale Space* formalism involves an infinite number of scales, in our practical implementation we use a finite number of filter scales, restricted to the range of $[1.0, 4.0]h^{-1}$Mpc. We tested a few other filter ranges, and concluded that the used range provides a robust population of filaments. This includes the major dynamically dominant filaments, along with a fair fraction of the tenuous tendrils. We also assessed this for the higher redshift simulation

Table 5.1: Mass and volume fractions of the different components of the cosmic web as identified by NEXUS+.

Component	Mass fraction [%]	Volume fraction [%]
Filaments	46.2	7.25
Voids	21.4	72.09
Nodes	4.6	20.65
Walls	27.8	0.01

snapshots, to find that for $z > 3.8$ most filaments are to be found in this scale range.

The morphological signature is evaluated on a 512^3 grid, ie. for each of the gridpoints on this grid NEXUS+ provides us with the identity of the morphological features to which it belongs.

5.3 Force and Strain Field: Procedure

For the computation of the diverse gravitational force and tidal fields, we use a brute force approach. At each grid location we compute the contributions by the various morphological elements of the cosmic web by summing the gravitational and tidal force contributions from the grid points located in the regions that have been identified with this morphology. In short, the force and tidal influence of filaments is the sum of the mass elements residing in filaments, of voids the sum of mass elements residing in voids, etc. The total (peculiar) gravitational force and tidal strain at each location is the sum of the contributions by all mass elements in the volume.

The stated force and tidal force calculations involve the following practical issues:

1. The gravitational force field and tidal strain field are evaluated on the same 512^3 grid as the one on which we have computed the DTFE density field, and assessed the NEXUS+ identity (see previous section).

2. The forces and tidal strains are computed by brute force, i.e. we compute these at a given location by summing the force and tidal contribution.

3. Following a sheer brute force summation, our computational resources allow us to carry out the computation for a restricted number of locations, instead of for all 512^3 gridpoints.

4. For the brute force computation, we assume that a spherical surrounding volume of radius $150h^{-1}$Mpc, corresponding to the largest sphere that fits

inside the simulation box, is sufficient to include all gravitationally relevant influences. In other words, force and tide contributions are considered to be negligible beyond a distance of $r = 150h^{1-}$Mpc. This is certainly true for the tidal contributions, and turns out to be valid to very good approximation for the gravitational force itself.

5. A more efficient Fourier space based formalism is under development, but has not yet been applied to the analysis of the simulation in this study. A Fourier code would automatically guarantee periodic boundary conditions pertaining to the simulation box, and hence by implication include the force and tidal influences throughout the entire simulation volume.

Because the brute force calculations are computationally intensive, we restricted the force and tidal evaluations to a few slices of the simulation box, yielding a total of 512^2 force evaluations on the corresponding 512^2 grid. The force at each point is computed using the full three dimensional grid, yielding a fully three dimensional force and tidal tensor at each gridpoint. For the statistical results presented in our study, we established that the force and tides on a 512^2 two-dimensional grid is representative and hence sufficient for our discussion. Figure 5.4 shows the resulting total force field. Figure 5.5 shows the NEXUS+ morphological identifications in the same slice. The volume and mass fractions for the different components of the cosmic web are given in Table 5.1.

5.3.1 Gravitational Force field

The peculiar gravitational force field for a continuous matter distribution field, specified in terms of the density perturbations $\delta_m(\mathbf{x})$,

$$\Delta_m(\mathbf{x}) = \frac{\rho(\mathbf{x},t) - \rho_u(t)}{\rho_u(t)}, \tag{5.2}$$

The resulting peculiar gravity field, following Peebles (1980), is given by

$$\mathbf{g}(\mathbf{x}) = \frac{3\Omega_m H_0^2}{8\pi} \int \mathrm{d}\mathbf{x}' \Delta_m(\mathbf{x}') \frac{\mathbf{x}' - \mathbf{x}}{|\mathbf{x}' - \mathbf{x}|^3}. \tag{5.3}$$

As we represent this field on a discrete grid, the force integral expression is converted into a sum over the mass-weighted contribution to the gravitational force by each gridpoint. Evaluating the force field at the gridpoint i, at location $\mathbf{x_i}$, the corresponding grid expression for the full force field $\mathbf{g}_i(\mathbf{x_i})$ is the sum over all N gridpoints j, at locations $\mathbf{x'}_j$,

$$\mathbf{g}(\mathbf{x_i}) = \frac{3\Omega_m H_0^2}{8\pi} \sum_j^N \Delta_m(\mathbf{x'}_j) \frac{(\mathbf{x_i} - \mathbf{x'}_j)}{|\mathbf{x}_i - \mathbf{x'}_j|^3}. \tag{5.4}$$

The summation is done up to a distance of $150h^{-1}$Mpc.

The gravitational force induced by the different cosmic web morphological components is obtained by the straightforward mass-weighted summation over all gridpoints that are located within the NEXUS+ identified regions, ie. the regions identified by NEXUS+ with either cluster node, filament, wall or void. The four cosmic web morphological components are assigned by index CWM = $node, fil, wall, void$. The (peculiar) gravitational force induced by the mass residing in any of these four morphological components is then the sum of the gravitational pull induced by the M mass elements residing in the corresponding areas,

$$\mathbf{g}^{\rm CWM}(\mathbf{x_i}) = \frac{3\Omega_m H_0^2}{8\pi} \sum_k^{\rm M} \Delta_m(\mathbf{x'}_k) \frac{(\mathbf{x_i} - \mathbf{x'}_k)}{|\mathbf{x}_i - \mathbf{x'}_k|^3}. \tag{5.5}$$

where $\mathbf{x}_i$ is the location for which the force is calculated, $\mathbf{x}_k$ one of the locations identified as belonging to morphological feature CWM and M the total number of grid-points identified with that morphological feature.

The net result of the force calculation is a representation of the force field $\mathbf{g}(\mathbf{x})$ at every gridpoint by five force vectors, the full gravitational force $\mathbf{g}(\mathbf{x_i})$, and the four morphological contributions, $\mathbf{g}^{\rm node}(\mathbf{x_i})$, $\mathbf{g}^{\rm fil}(\mathbf{x_i})$, $\mathbf{g}^{\rm wall}(\mathbf{x_i})$ and $\mathbf{g}^{\rm void}(\mathbf{x_i})$. Evidently, the total force $\mathbf{g}(\mathbf{x_i})$ is the sum of the four morphological contributions,

$$\mathbf{g}(\mathbf{x_i}) = \mathbf{g}^{\rm node}(\mathbf{x_i}) + \mathbf{g}^{\rm fil}(\mathbf{x_i}) + \mathbf{g}^{\rm wall}(\mathbf{x_i}) + \mathbf{g}^{\rm void}(\mathbf{x_i}). \tag{5.6}$$

Gravitational Force field: visualization

One of the principal aspects of the present study is to study and assess the gravity, velocity and tidal tensor fields. Before any quantitative and statistical assessment of these fields, the most direct impression of their characteristics and spatial properties is obtained by visual inspection.

To obtain an impression of the spatial structure and properties of the gravity and velocity fields, we invoke a few different visualisations:

- *gravitational force amplitude* maps
  of the amplitude of the gravitational force, and of the amplitude contributions by the various individual cosmic web morphological components.

- *gravity vector field* maps
  This includes maps of the full gravity vector field, as well as the vector plots of the gravity contributions by the various individual cosmic web morphologies.
  In the various gravity vector plots, we depict the component of the gravity vector in the plotted (2-D) box slice. The vectors are oriented towards the direction of the gravity vector component in the box slice. The length of the gravity vectors is normalized in one of two different options:

- *total* gravitational force:
  vector length proportional to the amplitude of the total gravitational force $\mathbf{g}(\mathbf{x})$, in units of the mean gravitational force amplitude, ie. in units of the dispersion $\sigma(|\mathbf{g}|)$.

$$\mathbf{g}_{\rm tot}^{\rm CWM}(\mathbf{x}) = \frac{\mathbf{g}^{\rm CWM}(\mathbf{x})}{\sigma(|\mathbf{g}|)}. \tag{5.7}$$

- *relative* gravitational force:
  local relative contribution to the gravitational force by each of the cosmic web components, by expressing the amplitude of the gravitational force in terms of the (local) total gravitational force $\mathbf{g}(\mathbf{x})$,

$$\mathbf{g}_{\rm rel}^{\rm CWM}(\mathbf{x}) = \frac{\mathbf{g}^{\rm CWM}(\mathbf{x})}{|\mathbf{g}(\mathbf{x})|} = \frac{\mathbf{g}_{\rm tot}^{\rm CWM}(\mathbf{x})}{|\mathbf{g}_{tot}(\mathbf{x})|}. \tag{5.8}$$

  In a sense, $\mathbf{g}_{\rm rel}^{\rm CWM}(\mathbf{x})$ adheres to the concept of the fractional contribution by the cosmic web components CWM to the local gravitational force.

- *velocity streamline* map. The velocity streamline field provides a transparent view of the spatial structure of the velocity field, and allows the identification of regions dominated by divergence, ie. inflow or outflow, and those by shear flow. To this end, the gravity field is transformed into an *equivalent* linear velocity field (Peebles, 1980),

$$\mathbf{v}_{\rm lin} = \frac{2f}{3H_0\Omega_m}\mathbf{g}, \tag{5.9}$$

  in which the linear growth rate

$$f(\Omega_m) = \frac{a}{D}\frac{dD}{da} \approx \Omega_m^{\gamma}, \gamma = 0.55 + 0.05(1+w), \tag{5.10}$$

  Streamlines are everywhere tangent to the velocity vectors in the flow field, and represent the direction of velocity at each point in the flow. On the basis of the linear velocity field, the streamline map is inferred using the matplotlib (Hunter, 2007) function pyplot.streamplot.

5.3.2 Tidal field

To calculate the tidal field at $\mathbf{r} = (x_1, x_2, x_3)$ we use the traceless tidal tensor defined as (van de Weygaert & Bond, 2008b)

$$T_{ij}(\mathbf{r}) = \frac{3\Omega_m H_0^2}{8\pi}\int {\rm d}\mathbf{r}'\Delta_m(\mathbf{r}')\frac{3(x_i'-x_i)(x_j'-x_j)-|\mathbf{r}'-\mathbf{r}|^2\delta_{ij}}{|\mathbf{r}-\mathbf{r}'|^5} - \frac{1}{2}\Omega H^2\Delta_m(\mathbf{r})\delta_{ij}, \tag{5.11}$$

In order to calculate this numerically we turn the integral into a discrete sum over the grid with the DTFE density field values. In our evaluations, we smooth the density field with a Gaussian on a scale of 5 h^{-1}Mpc. This is done to smooth out the regions in which nonlinear behaviour would be dominant. Representing this field on a discrete grid, the tidal force integral expression is converted into a sum over the mass-weighted contribution to the tidal force by each gridpoint. Evaluating the tidal field at the gridpoint (**r**), the corresponding grid expression for the full tidal field $\tilde{T}_{ij}(\mathbf{r})$ is the sum over all N gridpoints ($\mathbf{r}'$)

$$\tilde{T}_{ij}(\mathbf{r}) = \frac{3\Omega_m H_0^2}{8\pi} \sum_k^N \Delta_m(\mathbf{r}') \frac{3(x_i'-x_i)(x_j'-x_j) - |\mathbf{r}'-\mathbf{r}|^2\delta_{ij}}{|\mathbf{r}-\mathbf{r}'|^5}. \tag{5.12}$$

In order to separate the tidal field contribution by the different components CWM - (node,filament, wall or void) - of the cosmic web we sum only over the M gridpoints that are located within a region identified as of morphology CWM. Only the mass allocated at those gridpoints ia taken along in the tidal force calculation.

$$\tilde{T}_{ij}^{\rm CWM}(\mathbf{r}) = \frac{3\Omega_m H_0^2}{8\pi} \sum_k^M \Delta_m(\mathbf{r}') \frac{3(x_i'-x_i)(x_j'-x_j) - |\mathbf{r}'-\mathbf{r}|^2\delta_{ij}}{|\mathbf{r}-\mathbf{r}'|^5}. \tag{5.13}$$

The sum uses all cells up to a radius of 150h^{-1}Mpc. The traceless tidal tensor T_{ij} follows from

$$T_{ij} = \tilde{T}_{ij} - \frac{1}{3}(\tilde{T}_{11} + \tilde{T}_{22} + \tilde{T}_{32})\delta_{ij}. \tag{5.14}$$

The eigenvalues of tensor T_{ij} are

$$\begin{aligned} &T_1 > T_2 > T_3 \\ &|T| = \sqrt{T_1^2 + T_2^2 + T_3^2} \end{aligned} \tag{5.15}$$

where T_1, T_2 and T_3 are the sorted eigenvalues of the T_{ij} matrix and |T| signifies the strength of the tidal field. The accompanying eigenvectors $\hat{T}_1, \hat{T}_2, \hat{T}_3$ are sorted accordingly. The value of the different eigenvalues signify the level of elongation or contraction that the corresponding mass element undergoes due to tidal forces.

Tidal field: normalization & visualization

In the analysis of the tidal field, we use two different normalization options:

- *total* tidal force:
  the tidal force induced by cosmic web component CWM, $T_{ij}^{\rm CWM}(\mathbf{x})$, in units of the mean tidal force amplitude, ie. the dispersion $\sigma(|T|)$

$$T_{ij,\rm tot}^{\rm CWM}(\mathbf{x}) = \frac{T_{ij}^{\rm CWM}(\mathbf{x})}{\sigma(|T|)}. \tag{5.16}$$

- *relative* tidal force:
  local relative contribution to the tidal force by each of the cosmic components, by expressing the tidal tensor components and amplitude in terms of the local total gravitational force $|T|_{\rm tot}(\mathbf{x})$,

$$T^{\rm CWM}_{ij,\rm rel}(\mathbf{x}) = \frac{T^{\rm CWM}_{ij}(\mathbf{x})}{|T|(\mathbf{x})|}. \tag{5.17}$$

  In a sense, $T^{\rm CWM}_{ij,\rm rel}(\mathbf{x})$ adheres to the concept of the fractional contribution by the cosmic web components CWM to the local tidal field.

To visualize orientation, coherence and strength of the tidal force field, we use two different kinds of maps:

- *tidal force amplitude* maps
  maps of the amplitude of the tidal field and tidal field contributions by the various cosmic web morphological components. For the amplitude of the tidal field we use $|T|$ as defined in Eq. 5.15.

- *tidal bar* maps
  Tidal bars indicate the direction of the principal orientation of the tidal force field, in combination with their strength. They have the orientation of the eigenvectors of the tidal force, with a length proportional to the absolute value of the corresponding eigenvector,

$$T_{i,\rm plot} = \hat{T}_i |T_i|. \tag{5.18}$$

  In the tidal bar maps in the present study we depict the compressional component of the tidal force.

From these maps we may hence deduce the directions along which the mass elements get compressed.

5.4 Gravitational Force Field

The gravity field contributions by the various cosmic web components, for the specific case of the slice in figure 5.4, are shown in figure 5.6 and figure 5.7. The four panels in figure 5.6 show the gravity vectors for the combined force impact by filaments (top lefthand frame), voids (top righthand frame), walls (bottom lefthand frame) and cluster nodes (bottom righthand panel). We should note that the vector arrows for the filament force field are rescaled, as they are by far the most substantial contribution to the gravity force field (which should perhaps not surprise us, given filaments contain more than 50% of the cosmic matter content, see Cautun et al. (2014)). With the gravity vector plot providing insight

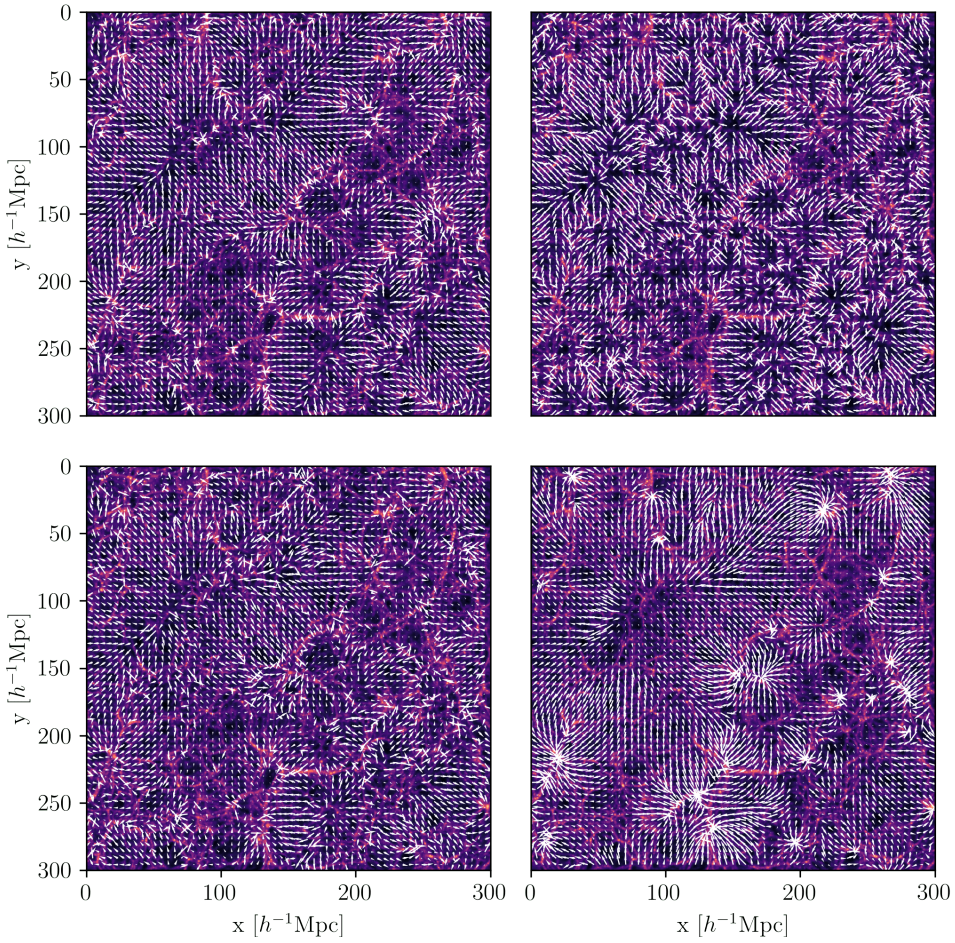


Figure 5.6: Gravitational Force field by Cosmic Web component: vector field. The force fields originating of the different components for the complete field, normalised by the total field (see Section 5.3.1). Top left has filaments, Top right has voids, bottom left has walls and bottom right has nodes. **Note:** For this figure, filaments are rescaled with respect to the other components. In comparison the filament arrows should be longer. Voids, walls and nodes are one the same scale.

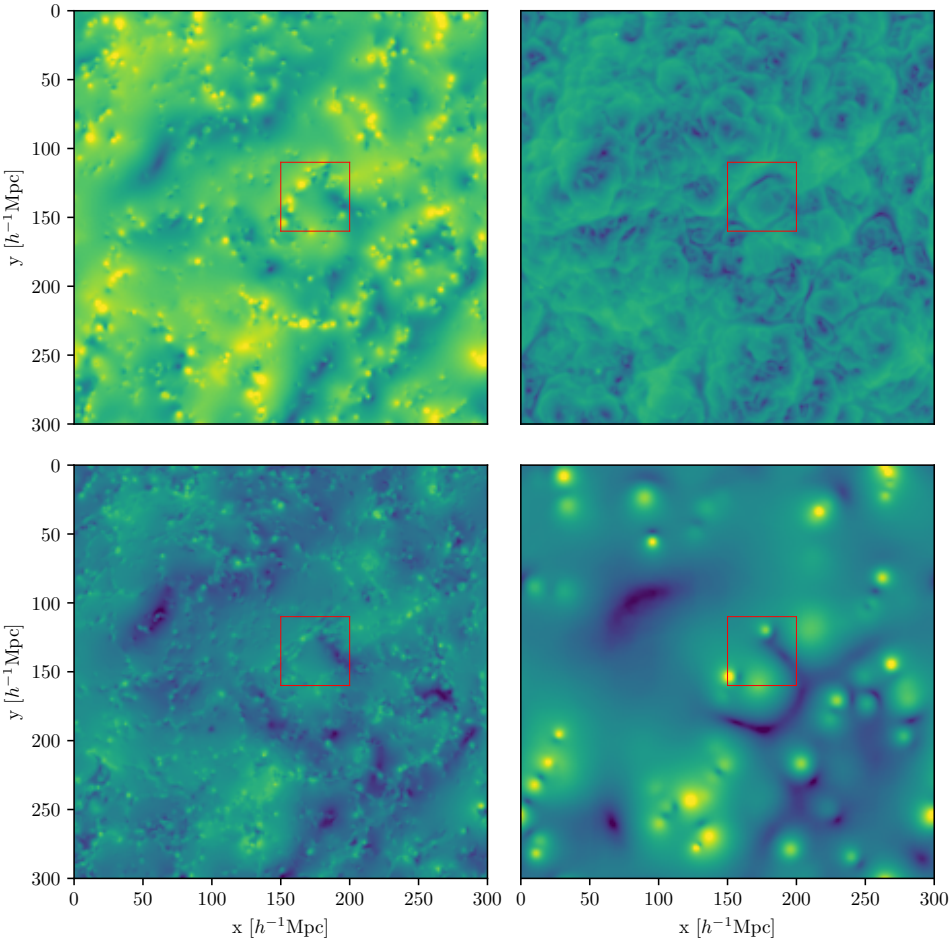


Figure 5.7: Gravitational Force field by Cosmic Web component: amplitude field. The amplitude of the peculiar force field $\log_{10}|\mathbf{g}|$ for the different components of the cosmic web for a 300 h^{-1}Mpc by 300 h^{-1}Mpc by 0.59 h^{-1}Mpc slice. The top left panel shows the component originating from filaments, the top right panel shows the component originating from voids, the bottom left panel shows the component originating from walls and the bottom right panel shows the component originating from nodes. The red box indicates the region used for the zoom in Figure 5.10.

into the structural patterns of the induced gravitational force fields, maps of the strength of the gravitational force yield the needed information on the relative strength of the cosmic web morphological components and the spatial extent of their influence. Figure 5.7 contains the four corresponding panels with force amplitude level maps for the filament population (top lefthand), void population (top righthand), wall population (bottom lefthand) and cluster nodes (bottom righthand).

One immediate observation is the significant differences between the gravitational influence of the different morphological components of the cosmic web. Each contributes uniquely to the total force field, yielding force fields that differ substantially in shape, coherence, scale and amplitude. Most outstanding is the dominant gravitational influence of the filamentary network. Their overarching strength can be directly inferred from the gravity amplitude maps in figure 5.7. The towering influence of filaments is expressed in the fact that throughout the simulation box the filament force is factors higher than that exerted by voids and walls. Only cluster nodes reach comparable strengths, but this is only restricted to the immediate vicinity of the cluster nodes. Perhaps the most interesting finding is that of the significant and coherent gravitational impact of voids, marked by a typical signature in terms of characteristic bubble shaped regions in and around minima in the matter distribution, the manifestation of a typical - effectively repulsive - force field pattern. Moreover, comparison between the cluster node and void force field reveals the perhaps surprising conclusion that voids are the more dominant component over far larger regions of a cosmic volume.

5

5.4.1 Filaments

From the filament force vector field map (top lefthand panel, fig. 5.6), we see that the gravitational influence of filaments acts coherently over a large range of distances. In fact, its overall structure and pattern is quite close to that of the total gravitational force field (bottom panel, fig. 5.4, a telling testimony of the fact that most of the large scale gravitational force field is due to the combined influence of the filamentary network. We should also note that, as expected, the force field only traces the largest structures. Small scale features hardly induce a noticeable effect. The overpowering dynamical influence of the filamentary network should not come as a surprise: the filamentary network of the cosmic web represents more than 50% of the cosmic matter content of the universe (Cautun et al., 2014). As important is the pervasive spatial character of the multiscale filamentary network, with filaments spreading throughout cosmic volumes and their tendrils branching out into even the most remote and desolate realms (voids).

Inspection of the force amplitude map shows that over nearly the entire cosmic volume filaments are responsible for high force amplitudes, their overdense nature guaranteeing a consistent and coherent cumulative attractive force contri-

bution (top lefthand panel, fig. 5.7). We find that over almost 97% of the cosmic volume, the filament induced forces dominate over that induced by any of the other morphological components (table 5.2, 3rd column). The largest and densest filaments create coherent force fields on scales up to $\sim 50\ h^{-1}$Mpc. In fact, the coherence scale of the force field is much larger than that of the weblike structure of the underlying matter distribution.

5.4.2 Voids

The picture is considerably different for the void force field. The top righthand panel of figure 5.6 shows a spatial pattern marked by individual divergent patches, indicating individual expanding void regions. Inside these individual void regions, the force field increases monotonically - and usually almost linearly - from the center of a void to their boundary of surrounding walls and filaments. It produced the characteristic superhubble expanding void regions (see e.g. Icke, 1984; van de Weygaert & van Kampen, 1993). Due to the rapidly rising mass density near their boundary, the strength of the void force field drops at the void edge, resulting in the conspicuous bubble-like configurations that stand out in the void force field in figure 5.6.

Overall the void force field appears to be one in which individual superhubble expanding voids (Icke, 1984) produce a segmented volume with clearly distinguishable individual void regions. Indicative for the latter are the small voids near the center, whose influence appears to be mainly limited to their own interior. Nonetheless, we may observe assemblies of voids that seem to operate collectively in pushing out the surrounding mass, such as the the void agglomeration at the top left of the field. Its force fields adds up to create an expanding region on a scale of $\sim 50\ h^{-1}$Mpc.

In all, it suggests that in general voids hold sway over in particular their local environment, in which large scale effects are significant but less prominent than those seen in the filament force field. On large scales voids do produce significant residual large scale effects emanating due to their spatial clustering, be it of a lower amplitude than that characteristic for the forces in individual voids. It confirms the finding by Platen et al. (2008), who showed that voids even induce noticeable tidal effects over scales in excess of $\sim 30\ h^{-1}$Mpc).

5.4.3 Walls

The force field induced by the walls is the least conspicuous one of the web morphologies. Throughout the cosmic volume it is very weak, as testified by the map of its force strength (see bottom lefthand panel fig. 5.7. The walls do not contribute much to the overall force field. Also, as they populate moderate density realms as well as the interior of voids (see e.g. Cautun et al., 2014), their effect may be effectively either be attractive or repulsive. To some extent, it leads to the

moderate and low density walls compensating each others force contributions. It results in a force field that only faintly reflects the cosmic web pattern seen in the filament and void force fields.

Globally, the spatial structure of the wall force map bears resemblance to that of the filament strength map (top lefthand panel fig. 5.7. It appears to be a faint reflection of the filament strength map. One may recognise only a part of the most prominent features that are seen in the filament force field, others are conspicuously absent: e.g. the strong filament force field at the top lefthand corner of the box does not seem to have a wall equivalent. Also the overall large scale pattern of the wall force vector map does resemble that of the filament force vector map. However, the wall force map appears far less coherent and organised, marked by a considerable number of randomly scattered small scale patches of irregular oriented force vectors.

5.4.4 Cluster Nodes

5

A rather surprising finding is the fact that cluster nodes only appear as highly localised force centers, and they are far from the dominant gravitational component of the cosmic web that we had expected them to be.

In the force strength map (bottom righthand panel, fig. 5.7) cluster nodes stand out as compact high amplitude peaks in the force field. They act as a set of randomly clustered monopole attractors. The cluster nodes are only dominant over the rather small scales of their immediate environment. At this range, they overshadow all other force contributions. In most cases, this is only over distances of a few h^{-1}Mpc, in an occasional exception out to $\sim 20h^{-1}$Mpc. The latter concerns an agglomerate, “superclusters", such as that at the bottom centre of the map (bottom righthand panel, fig. 5.7).

Their spatial connection to the cosmic web is less clear, and it is not easy to recognize the global weblike mass distribution in their own spatial arrangement or the direct dynamical relation between the cluster nodes and the spatial intricacies of the cosmic web. At large scales there are substantial regions were the influence of multiple nodes is one in which they effectively cancel each other, resulting in a contribution that is less than - or at best comparable to - that by the voids.

5.4.5 Gravity Field strength: statistical analysis

To quantify the visual impressions of the force field contributions by the different cosmic web components, discussed extensively above, we assess the statistical (volume-weighted) distribution of the fractional gravitational force contributions by each of the cosmic web components CWM. The fractional gravitational force

Table 5.2: Gravitational Force Statistics. The second column shows the mean force ratio for the different components of the cosmic web, Note the direction dependence in the ratio $|\mathbf{g}|^{\rm CWM}/|\mathbf{g}|^{\rm tot}$. The third column shows the percentage of the total volume where each component has a contribution that is larger than the other components.

Component	Mean [%]	Largest in [%]
Filaments	60.4	97.2
Voids	20.3	0.9
Nodes	17.0	1.5
Walls	13.7	0.4

$\mathcal{F}(\mathbf{r})$ at each location $\mathbf{r}$ is defined by the ratio,

$$\mathcal{F}^{\rm CWM}(\mathbf{r}) = \frac{|\mathbf{g}^{\rm CWM}(\mathbf{r})|}{|\mathbf{g}_{\rm tot}(\mathbf{r})|}, \tag{5.19}$$

where $\mathbf{g}^{\rm CWM}(\mathbf{r})$ is the gravitational force exerted by the components CWM - i.e. filaments, voids, walls and cluster nodes - at location $\mathbf{r}$. The total sum of these is

$$\mathbf{g}_{\rm tot}(\mathbf{r}) = \sum_{\rm CWM} \mathbf{g}^{\rm CWM}(\mathbf{r}). \tag{5.20}$$

Note that the ratio is independent of direction, while the sum of the forces does take into account the orientation of each force contribution. As a result, the fractional force $\mathcal{F}(\mathbf{r})$ may have a value larger than unity. The average ratios for the fractional force contributions are listed in table 5.2. There is a clear hierarchy of force field contributions: filaments $\gg$ voids $>$ nodes $>$ walls. Averaged over all volume elements, the filament impact is on average no less than 60%, with voids as a surprising second at on average contributing $\sim 20\%$ of the local gravitational force.

The distribution for the complete force field is shown in the top panel of Figure 5.8. The bottom four panels show the same distribution, but split up by morphology. They show for each of the four cosmic web environments, the fractional force contribution by each of the morphologies. In other words, these tell you what the relative impact is of filaments in voids, of voids in filaments, or of cluster nodes in filaments.

The top panel of Figure 5.8 reveals the major differences in force impact by the different cosmic web components. Filaments stand out as by far the most dominant element of the cosmic web. Their force impact ranges over a wide range of values $\mathcal{F}^{fil}(\mathbf{r})$, centering around 60% of the total gravitational force. At the vast majority of locations the filamentary cosmic web is the leading gravitational influence.

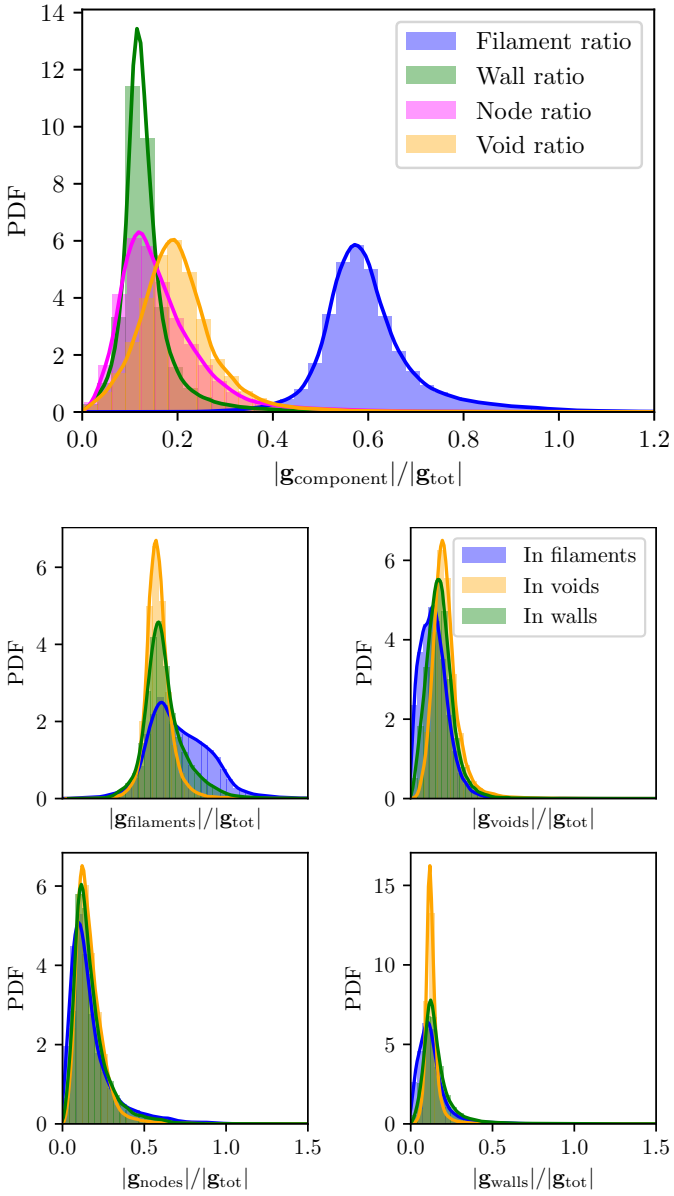


Figure 5.8: Cosmic Web Gravitational Force Field: Strength Inventory. Top panel: The distribution of $|\mathbf{g}^{\rm CWM}|/|\mathbf{g}_{\rm tot}|$ for the different components of the cosmic web. Each using the complete field. Bottom four panels: the distribution of $|\mathbf{g}^{\rm CWM}|/|\mathbf{g}_{\rm tot}|$ for the different components of the cosmic web. Per panel the pdf for cosmic web component CWM is plotted, within three different environments: filaments (blue), voids (yellow), walls (green). Centre left: filament induced force field. Centre right: void induced force field. Bottom left: wall induced force field. Bottom right: cluster induced force field.

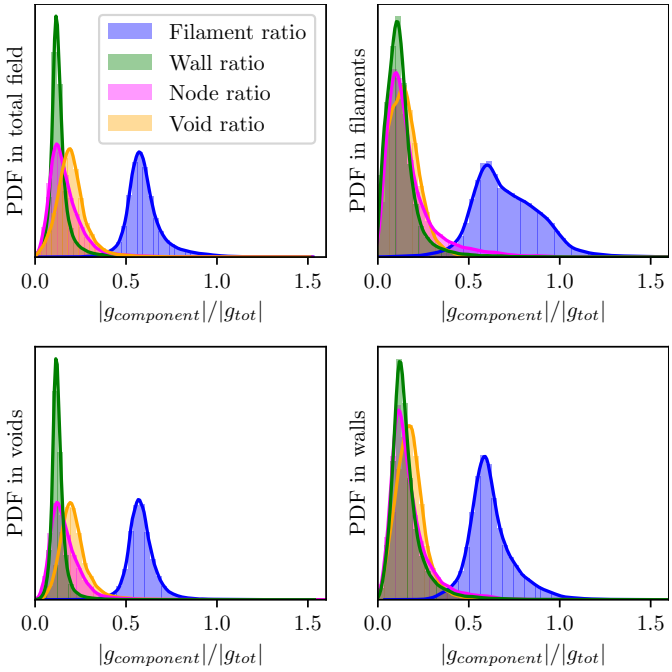


Figure 5.9: Cosmic Web Gravitational Force Field: Strength Inventory by Cosmic Web environment. The panels show the pdf of the induced gravitational force field $|\mathbf{g}^{\rm CWM}|/|\mathbf{g}_{\rm tot}|$ by each of the four cosmic web components CWM in the different cosmic web environments. The different colours represent the different components: filament induced force field (blue), void induced force field (yellow), wall induced force field (green) and cluster node induced force field (magenta). Top left panel: total gravitational field. Top right panel: gravitational field in filaments. Bottom left panel: gravitational field in voids. Bottom right panel: gravitational field in walls.

The role of voids is one of the most interesting aspects of the gravitational force inventory. They appear to be the second most dominant source of gravitational force. They account for more than 20% of the local force at far more locations than e.g. walls, and even than cluster nodes. At many locations voids assume an even stronger fraction of the force budget. The fact that the void distribution function has a long high end tail implies that voids are the dominant influence at various regions of the cosmic web. Nearly without exception this concerns the interior of large voids. A similar long high end tail is found for the cluster node gravitational influence $\mathcal{F}^{node}(\mathbf{r})$, reflecting the overpowering gravitational influence of nodes in and around their own location. We find that the cluster node distribution has a mode near $10-12\%$ of the total force, substantially less than that of voids. It means that over most of the cosmic volume, voids

have a stronger gravitational influence than cluster nodes !

As expected, the distribution of the wall force fraction $\mathcal{F}^{walls}(\mathbf{r})$ expresses their weak role in the cosmic web force budget. Not only does the distribution peak around 10%, quite similar to the mode for cluster nodes, but it also concerns a narrow distribution without an outstanding tail towards higher values. It shows that walls, over nearly the entire cosmic volume, are only a minor fiddle in the force concert.

Assessing the force impact differentiated by cosmic web environment provides additional information on the nature of the dynamical impact of filaments, voids, walls and nodes. These are provided by the bottom panels of figure 5.8 and by the panels of figure 5.9.

Differentiating between the impact of morphological components in different environments, the four bottom panels of figure 5.8 reveal a few interesting aspects. Filaments are clearly the dominant dynamical component over the entire cosmic volume, and this is true in voids, walls as well as in filaments themselves (voids: bottom lefthand panel fig. 5.9, walls: bottom righthand panel fig. 5.9, filaments: top righthand panel fig. 5.9). In the case of the filament interior, the shoulder of the filament pdf reveals that it is in particular inside filaments themselves that they display the strongest impact: filaments hold sway inside filaments. A minor detail is that we see that $\mathcal{F}^{fil}(\mathbf{r})$ is higher in walls than in voids. This surely is a manifestation of the spatial proximity of filaments to these structures: voids are large, and in the central interior the filament's force may be somewhat weaker than in and around the walls to which they are connected.

For the impact of voids in different cosmic web environments, we see that they hold their strongest influence over the void regions themselves, while they are relatively stronger in walls than in the interior of filaments (centre righthand panel fig. 5.8). Inside void interiors, the force inventory quantified in the bottom lefthand panel of fig. 5.9 shows how important external influences are in the dynamics and evolution of voids. The filament forces dominate the force field of voids. This is certainly true for the outer regions of voids, but may even be so for their inner regions. The latter will be most evident for small voids, where these external forces may even induce their collapse, an essential process - called void-in-cloud - in the buildup of the void population (Sheth & van de Weygaert, 2004). However, the filaments represents the principal gravitational influence for even the large voids. This implies that any analysis of void dynamics should include the mass distribution surrounding the void (see e.g van de Weygaert, 2016, for a recent review), while any exploitation of the characteristics of the void population for cosmological purposes cannot ignore this fact and base their analysis on simplistic isolated void dynamics (see e.g. Pisani et al., 2019). A recent study of the dynamics of a sample of voids from SDSS, including velocity flow information from the Cosmicflows-4 galaxy peculiar velocity survey, did also indicate

this finding in the observational reality (Courtois et al., 2023).

Meanwhile, less outstanding are the observations with respect to the gravitational influence of nodes and walls. Their impact appears relatively indifferent to the environment. The node force fraction decreases slightly going from voids to walls to filaments. This might be somewhat counter intuitive, but is an expression of the fact that filaments are so much stronger than walls and voids. Relatively speaking, nodes will then assume a smaller force fraction $\mathcal{F}^{node}(\mathbf{r})$ in a filament environment.

5.5 Gravitational Force Field: individual structures

In this section we zoom in on the force structure in and around a few representative individual features. This with the intention to develop an intuitive and visual image of the role and gravitational influence of the different cosmic web components with respect to the various large scale environments, and to appreciate the complex interplay between them. We concentrate on the force field in and around a single void, around a filament, and one in and around a cluster node. In each case we zoom in on $50 \times 50 h^{-1}$ Mpc segment. They confirm the observations and conclusions that we reached in the previous section on the basis of its structural and statistical assessment analysis of the cosmic web gravitational force field.

5.5.1 Case study 1: void force field

Our first study case is the force field in and around a void. Against the backdrop of the corresponding density field, with the red contours indicating the NEXUS+ identified region, the four panels in figure 5.10 show the decomposition of the gravitational force field in and around the void into its individual morphological contributions. The force field is decomposed into the filament induced force field (top lefthand panel), that by voids themselves (top righthand panel) by walls (bottom lefthand panel) and by cluster nodes (bottom righthand panel).

Evidently, the filament induced force field dominates over nearly the entire region. Two of the most characteristic aspects of the filament induced force field is its large scale reach and coherence, and the easily recognisable signature shear pattern in the gravity vector field. The latter involves the splitting of the gravity vector field in opposite directions at saddle points of the corresponding gravitational potential. Two of the most outstanding ones are the one near righthand edge of the void (at $(x, y) \approx (190, 140) h^{-1}$Mpc, and the one near $(x, y) \approx (170, 120) h^{-1}$. The latter is situated in between two dense concentrations within the filamentary network. These are responsible for stronger attractive cores along the filaments. The two dense mass concentrations along the filament

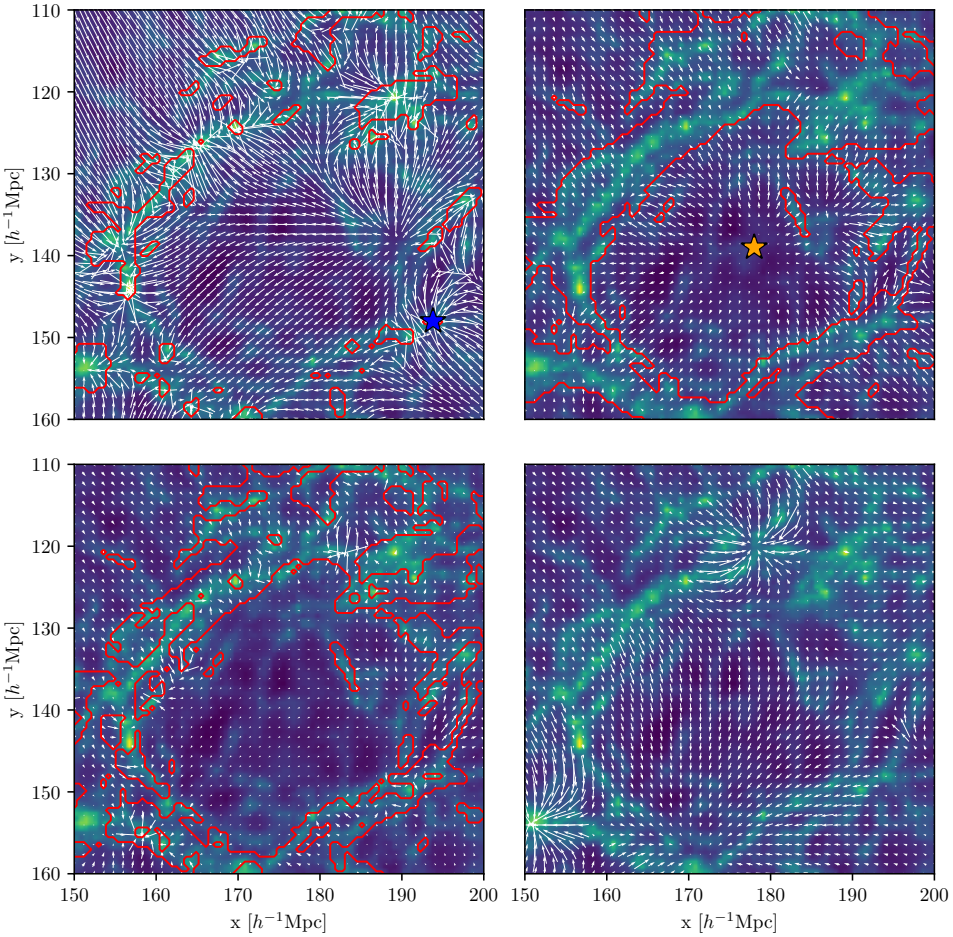


Figure 5.10: Gravitational force vector field in and around void region. The region is a 50 h^{-2}Mpc^2 region centered on a void (void center: orange star, top lefthand panel). The panels represent the force vector field in this region generated by the different morphological components of the cosmic web. Top left panel: filament induced force field; top right panel: void induced force field; bottom left panel: wall induced panel; borrom right panel: cluster node induced field. The arrows show the magnitude and direction of the corresponding (normalized) gravitational force, $\mathbf{g}_{tot}^{\rm CWM}$. The vector field is superimposed on the (log) density field map. In the different panels the red contours indicate the regions belonging to the corresponding NEXUS+ identified cosmic web component. The orange star (top righthand panel) indicates the void center, the blue star (top left panel) the filament location, around which the radial force profiles in figures 5.11 and 5.12 are determined.

turn out to be massive filamentary branches connecting to a cluster node, forming the connections of the cluster with the filamentary skeleton of the cosmic web. Over this specific void, the cluster node represents a sizeable external influence over the entire realm of the void. The cluster induced force field is comparable in magnitude to that of the void force field itself.

The coherence of the filament induced gravity field over the entire reach of the central void is the source for a large scale bulk flow (eq. 5.9), superimposed on the iconic superhubble divergent flow induced by the void(s) themselves. The induced force field by the voids themselves is easily recognisable, consisting of an almost pure divergent vector field centered around the density minimum (top righthand panel fig. 5.10. To first approximation, the superhubble force field in the interior of the void involves a radially outward directed force that increases with distance to the center (see eg. Icke, 1984; van de Weygaert, 2016). It reflects the almost uniform (underdense) mass distribution in the interior of the void (see e.g van de Weygaert & van Kampen, 1993; Sheth & van de Weygaert, 2004). In reality, also voids display a substantial level of substructure. The presence of a small expanding subvoid superimposed on the large central void reflects the multiscale structure of both the interior void density and velocity field. It is the product of the hierarchical buildup of the void population (Sheth & van de Weygaert, 2004; Aragon-Calvo & Szalay, 2013).

While the filamentary force field displays a coherence over nearly the entire volume, the void induced field appears to be more restricted in spatial extent. The void gravitational force increases radially outward up to its boundary, at a radius of around $r \approx 10 - 15h*-1$Mpc. At the edge it plummets to an almost negligible influence, as the gravity by the higher density outer realms of the voids kick in to compensate for the influence underdense interiors. Minor residual void forces may be traced in the interstitial boundary regions, yielding the suggestion of the voids pushing matter into the surrounding weblike filamentary and planar structures. Similar patterns are seen to emanate from the surrounding voids. As a result, we get the impression of a void force field segmented into separate void regions. This was also seen on the more global scale of the simulation volume in the void force field in figure 5.6 (top righthand panel). In all, we may conclude that while voids may have some minor effect on large scale flows, their main impact seems to be that of the local pushing around of matter into the interstitial elements of the cosmic web.

Quantitatively the above impressions are supported by the radial profile of the different force amplitudes around the center of the void. The top panel of Figure 5.13 plots the radially averaged force amplitude as seen from the center of the void. It confirms the dominance of the filament induced gravity (blue line), over the entire interior of the void (and beyond). The void induced force field (orange line) reveals the expected characteristic (almost) superhubble expansion. Near the edge of the void, the outward directed void force represents a substan-

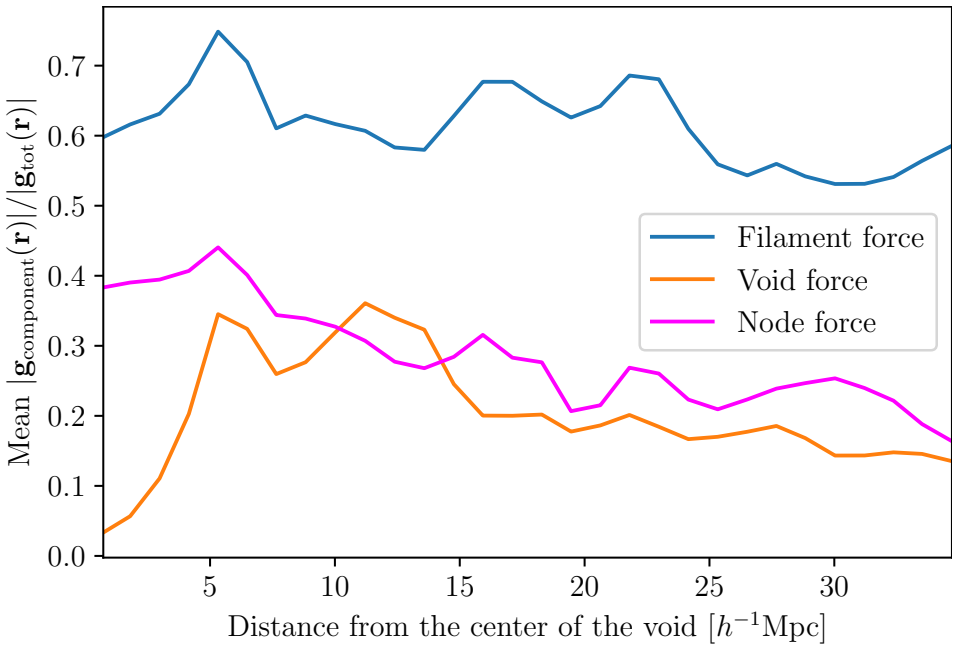


Figure 5.11: Radial gravitational force (amplitude) profile for the cosmic web components CWM, around the (orange) void center (fig.5.10 , top righthand panel). Plotted is the force ratio $|\mathbf{g}^{\rm CWM}|/|\mathbf{g}_{\rm tot}|$ for the CWM induced gravitational force amplitude, as a function of distance from the void centre. Blue: filament force amplitude; yellow: void force amplitude; magenta: cluster node force amplitude.

tial fraction of the complete gravitational force, but never more than around 40% of the filament force. A full account of the force field for this particular voids also includes a substantial influence from the massive cluster node on its north side, whose magnitude is comparable to that of the void itself. The corresponding force profile (magenta line) confirms the visual impression offered by the bottom lefthand panel of fig. 5.10.

Seen from the perspective of the filament, a complementary view of the sizeable void influence is obtained. The middle panel of Figure 5.13 shows the radial force profile around the dense filament location at the bottom right of the region (marked by a blue star, top left panel fig. 5.10). At small distances from this location, still inside the filament, the filament induced force field is overwhelmingly dominant, responsible for more than 80% of the full force field. As we move towards the edge of the filament, and enter into the void's realm, we observe a rapid decline in the filament's influence, going along with a corresponding rise of the void's influence. Further afield, as the influence of overdense surroundings at the other side of the void becomes noticeable, we see that the force contribution

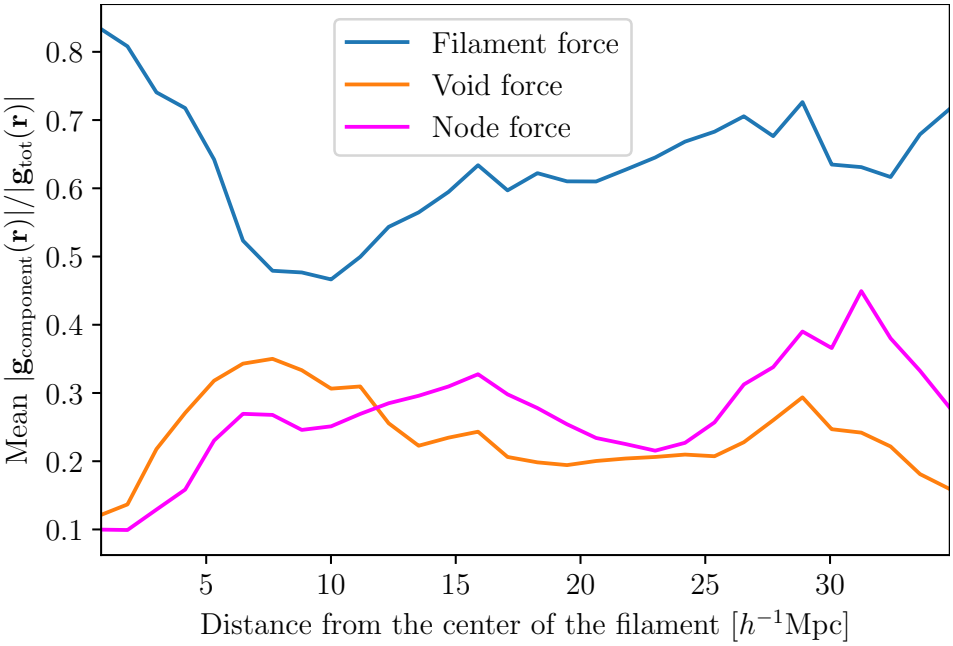


Figure 5.12: Radial gravitational force (amplitude) profile for the cosmic web components CWM, around the (blue) filament location (fig.5.10 , top lefthand panel). Plotted is the force ratio $|\mathbf{g}^{\rm CWM}|/|\mathbf{g}_{\rm tot}|$ for the CWM induced gravitational force amplitude, as a function of distance from the central filament location. Blue: filament force amplitude; yellow: void force amplitude; magenta: cluster node force amplitude.

by the void diminishes while that of filaments remains more or less constant.

5.5.2 Case study 2: cluster node force field & flowlines

The second case study concerns the force field around an isolated cluster node. To get a visual appreciation of its impact on the environment, we assess the induced flowfield in and around the node. To this end, we use the linear relation between gravity and velocity (Eq. 5.9). Formally, it is only valid in the linear regime, but to reasonable approximation may also be used in the quasi-linear configurations we are investigating. It yields the implied velocity fields induced by the four different cosmic web components, filaments, voids, walls and cluster nodes. These velocity fields are visualized by means of the corresponding streamlines. It yields a flowline representation of the velocity field that provides a direct insight into the structure of the cosmic migration flows. Figure 5.14 shows the streamline fields for the induced flowfield by filaments (top lefthand panel), the voids (top righthand panel), the walls (bottom lefthand panel) and the cluster

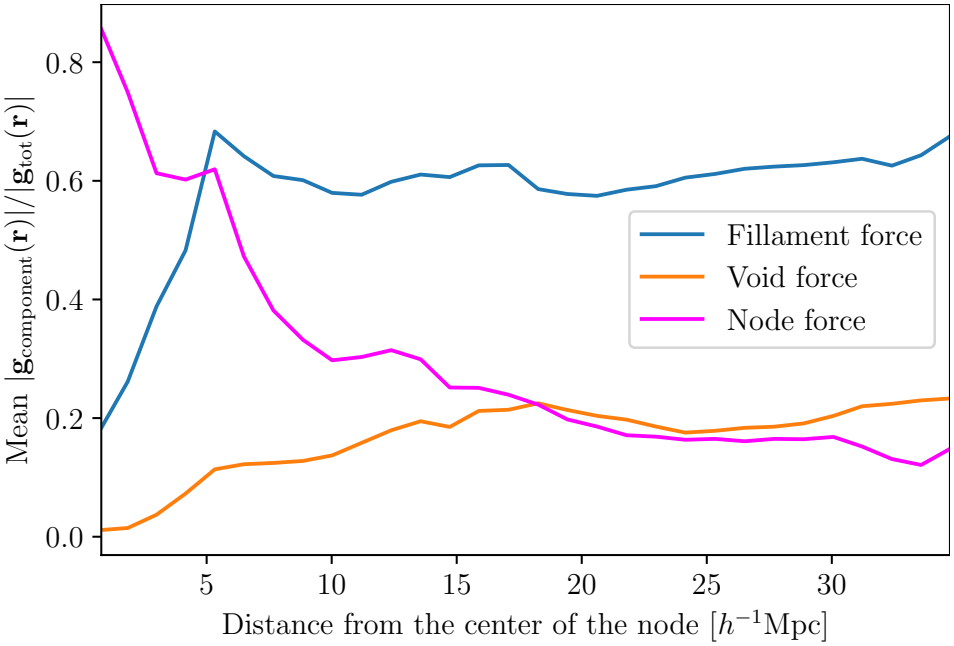


Figure 5.13: Radial gravitational force (amplitude) profile for the cosmic web components CWM, around the prominent massive cluster in fig. 5.14. Plotted is the force ratio $|\mathbf{g}^{\rm CWM}|/|\mathbf{g}_{\rm tot}|$ for the CWM induced gravitational force amplitude, as a function of distance from the cluster node location. Blue: filament force amplitude; yellow: void force amplitude; magenta: cluster node force amplitude.

nodes (bottom righthand panel).

The most prominent aspect of the flowfield is the cluster node induced flowfield (bottom righthand panel). The node induced a massive inflow pattern over the entire entire region. The dominant impact of the cluster on its environment is confirmed by the radial force profile centered on the node (bottom panel of fig. 5.13). In the inner sphere of $\approx 5h^{-1}$Mpc around the cluster node, it dictates the gravitational force field, taking care of even more than 80% at its center. Outside of this immediate vicinity, its sway rapidly declines, taken over by the surrounding filamentary web. To a considerable extent, the filament streamlines offer a similar spatial flow pattern, augmented by the presence of shear signature. It concerns the impact of the massive filamentary branches connecting to the node, and as such strongly related to this peak in the mass distribution.

By contrast, the void induced flow field offer a more localized pattern. In the overall pattern we recognize the typical segmented nature of the field, carrying the imprint of individual voids. More so than we have seen in the force vector field, the flow field also reveals the tendency for a large void scale induced bulk flow, running from the lefthand to the righthand side of the region. It is a man-

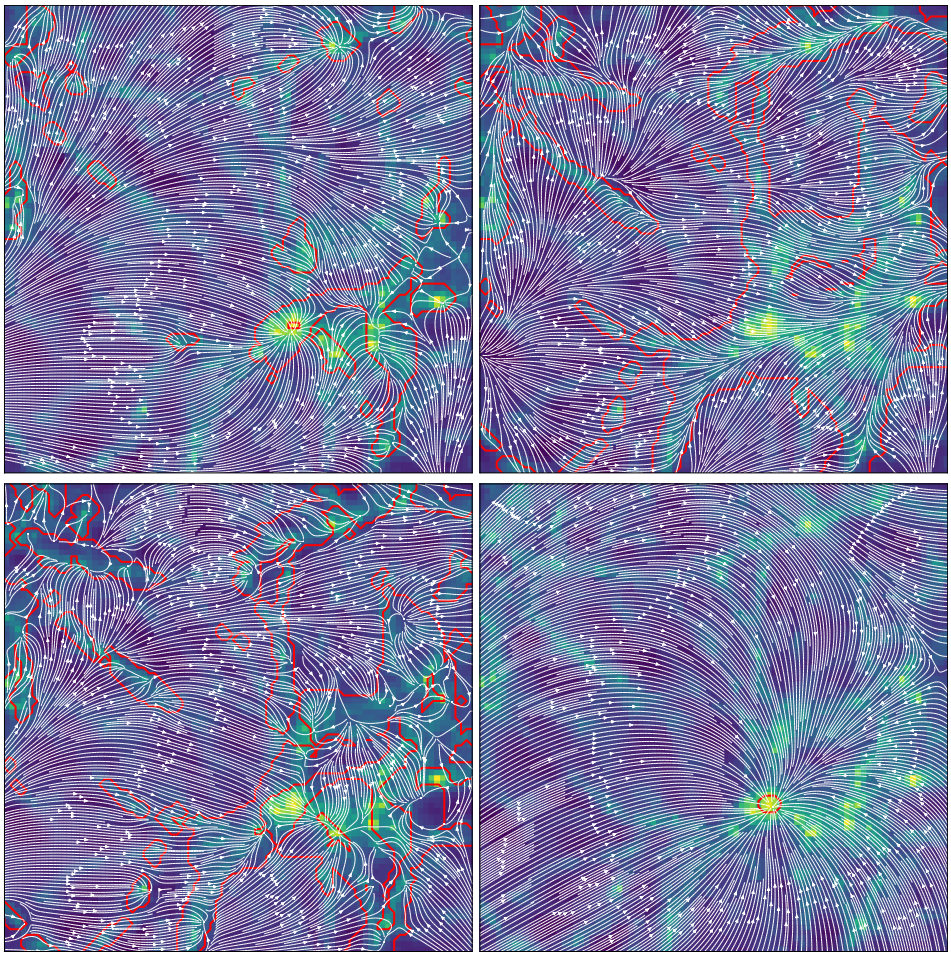


Figure 5.14: Velocity flow field in and around cluster node. The region is a 50 h^{-2}Mpc^2 region, containing a massive cluster node (at bottom righthand corner). The velocity vector field is represented by the corresponding streamlines, highlighting the overall spatial structure of the flowfield. The four panels show the streamline field for the induced velocity flow field by each of the cosmic web morphological components. The velocity field streamlines are superimposed on the (log) density field map. In the different panels the red contours indicate the regions belonging to the corresponding NEXUS+ identified cosmic web component. Top left panel: filament induced velocity field; top right panel: void induced velocity field; bottom left panel: wall induced velocity field; borrom right panel: cluster node induced velocity field. In this indivual case, the velocity field is clearly dominated by the cluster.

ifestation of the multiscale nature of the void population, a direct consequence of its hierarchical buildup (Sheth & van de Weygaert, 2004; Aragon-Calvo & Szalay, 2013). Voids tend to be embedded in larger (less) underdense regions, whose dynamical influence shows up in the superposition of the individual void outflows as a residual large scale (bulk) outflow (see Aragon-Calvo & Szalay, 2013, for telling illustrations). As seen from the center of the cluster node, the void induced flow field becomes noticeable at a distance of around $\sim 16h^{-1}$Mpc, beyond which distance it represents a stronger contribution than that of the cluster itself.

5.6 Tidal field

To understand the dynamics underlying the formation and shaping of the cosmic web, tidal gravitational forces and the induced deformation of mass elements play an instrumental role (see e.g. Zeldovich, 1970; Bond et al., 1996; van de Weygaert & Bond, 2008a; Hahn et al., 2010; Feldbrugge & van de Weygaert, 2023, 2024). The present study therefore needs to complement the inventory of the gravitational force in the previous section by an analysis of the corresponding tidal force field.

Following the definitions in section 5.3.2, in the present study we first investigate the tidal force field generated by filaments, voids, walls and cluster nodes by means of tidal force amplitude maps, i.e. maps of the (traceless) tidal field amplitude $|T|$,

$$|T| = \sqrt{T_1^2 + T_2^2 + T_3^2}\,, \tag{5.21}$$

with $T_1 > T_2 > T_3$ the tidal field eigenvalues. Figure 5.16 presents the tidal amplitude maps for the tidal fields generated by the filament population (top lefthand panel), the void population (top righthand panel), the wall population (bottom lefthand panel) and cluster nodes (bottom righthand panel). The first direct observation from the four frames is the difference between the tidal field morphology and patterns to that of the force field. The tidal force field is intimately linked to the weblike spatial pattern of the cosmic mass distribution, and also reflects more closely its smaller scale structure than the force field. Recent work has also revealed the close similarity between the structure of the primordial tidal field, specifically its eigenvalues, and that of the emerging structure of the cosmic web (see e.g Feldbrugge & van de Weygaert, 2023, 2024).

5.6.1 Tidal Field illustrated: tidal impact voids

To appreciate the intricacies of the cosmic web tidal force field, it is most insightful to focus on the component induced by the voids and the void population. To this end, we first zoom in on a few interesting and telling regions. Figure 5.15

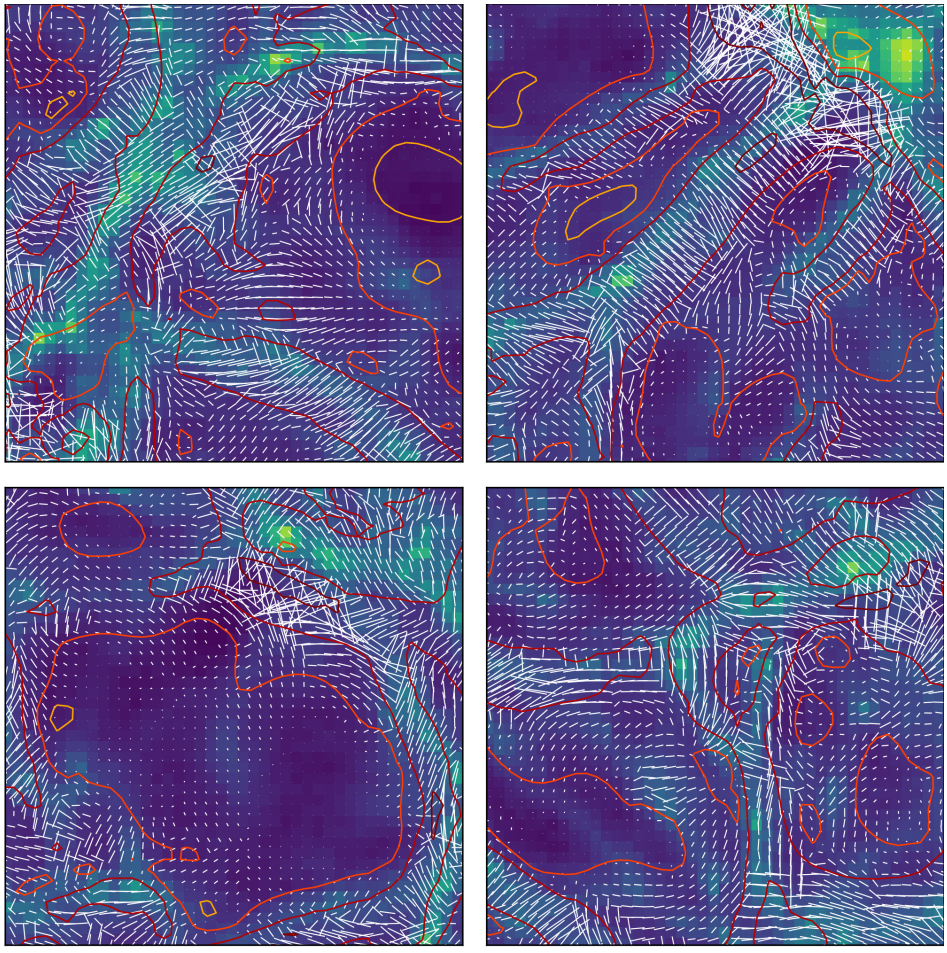


Figure 5.15: Tidal field induced by voids, in four different 50 h^{-2}Mpc^2 regions. The tidal force field is represented by tidal bars and contours. All four panels show the tidal field induced **exclusively** by the void population in the simulation box. The tidal bars and contours are superimposed on the (log) density field map. The tidal bars indicate strength and direction of the compressional component of the tidal field. The definition and construction of the tidal bars is explained in section 5.3.1.

shows zoom-ins on four different regions of size $25 \times 25 h^{-1}$Mpc. The four panels show the tidal bars of the compressional components of the tidal field induced by the void population in the matter distribution. The direction of the bars is along the corresponding tidal tensor eigenvectors, their size proportional to that of the corresponding eigenvalues.

The panels illustrate the close connection between the weblike spatial pattern of the matter distribution, and its multiscale nature, and the tidal force field induced by voids, which we find - as discussed in more detail below - to be a surprisingly striking aspect of the large scale tidal field. Even more so than the filament induced tidal force field, we find that the deformation of the matter distribution is strongly correlated to the spatial pattern and connectivity seen in the void induced tidal force field. The void tidal field traces out structures at all scales and over a wide range of densities. It causes compression along even the smaller, more tenuous, structures seen in the density field.

The clearest example of a void's tidal impact is offered by the bottom left-hand panel, centered at the interior of a void. It has the typical characteristics of a tidal force field expected in and around voids. A typical aspect of the void induced tidal force field is that it is weak in the interior of voids. It steeply increases in amplitude as it enters the overdense boundaries, a reflection of the strong differential gravitational force between interior and surroundings. In this sense, it adheres closely to that expected for isolated spherical voids. It induces a compressional component T_{rr} along the radial direction, while the dilational components T_{tt} are oriented along the tangential direction. At a radial distance r from the void center, for a void with density profile $\Delta(r)$,

$$\begin{aligned} T_{rr} &= \Omega H^2 \Big[\Delta(r) - \langle\Delta(r)\rangle\Big], \\ T_{tt} &= -\frac{1}{2} T_{rr}(r), \end{aligned} \tag{5.22}$$

in which $\langle\Delta\rangle(r)$ is the mean interior density of the void. Hence, in the near uniform - bucket shape - mass distribution in the interior of voids (see e.g. van de Weygaert, 2016) the tidal force field will be negligible. Meanwhile, the fact that $\Delta(r) > \langle\Delta(r)\rangle$ as we enter the boundary of voids implies into a compressional tidal force along the radial direction of the void. It translates into the compression of mass elements in the radial direction of the void, reflecting the formation of an overdense boundary around the void.

The variety and extent of the tidal impact of voids on their environment can be appreciated from the other three panels in figure 5.15. The top righthand panel is a telling illustration of their influence on the dynamics of filaments: the tidal bars trace the three filamentary structures that are connecting to the node at the top righthand corner of the panel. Evidently, the implication is that voids are instrumental in effecting further compression of the filaments. Overall, we see that the void induced tidal force field follows the underlying mass distribution

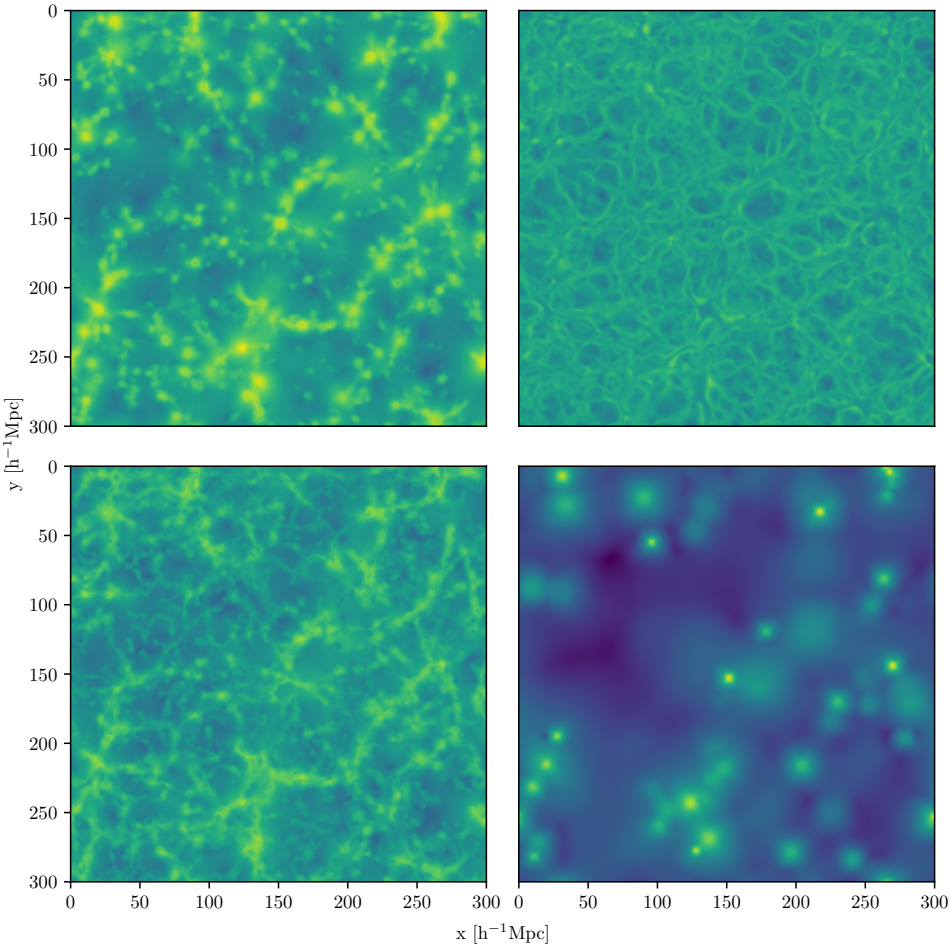


Figure 5.16: Tidal field by Cosmic Web component: amplitude field. The amplitude of the tidal field $\log_{10}|T|$ for the different components of the cosmic web for a 300 h^{-1}Mpc by 300 h^{-1}Mpc by 0.59 h^{-1}Mpc slice. The top left panel shows the component originating from filaments, the top right panel shows the component originating from voids, the bottom left panel shows the component originating from walls and the bottom right panel shows the component originating from nodes.

in meticulous detail, which we find to be an exclusive property of the void population. Even on scales smaller then a few Mpc the voids are able to closely trace out the underlying matter distribution. The tidal fields of the other components do not display this level of detail.

5.6.2 Tidal Field Structure

For comparison between the tidal influence of the various cosmic web components, figure 5.16 shows the tidal amplitude maps for each individual component of the cosmic web. The top left panel shows the filament tidal field, the top right panel shows the void tidal field, the bottom right panel shows the wall tidal field and the bottom right panel shows the tidal field for nodes. The first global impression offered by these maps is that the differences are starker than for the corresponding force fields.

On large scales, the amplitude maps for the filaments and walls share a similar pattern, be it that the tides generated by filaments is substantially stronger than that for walls. The fact that they share a similar large scale spatial structure is an expression of the fact that these anisotropic features are intimately related aspects of the weblike network of which they are the principal constituent elements. A clear difference with the corresponding force amplitude maps (fig. 5.7) is that the tidal fields rapidly drop to a negligible level directly outside the filaments and walls, a direct reflection of the more localised nature of the tidal field. In terms of the relative amplitudes the filament induced tidal forces are still dominant (see fig. 5.17), in particular within the filament and wall network of the cosmic web, but far less so than in the gravitational force field. Within the interior of the filaments, the filament induced tides stand out as by far the strongest influence. Wall induced tides are marginally stronger inside walls, somewhat stronger than those induced by (nearby) filaments and voids (see fig. 5.18).

The filament tidal amplitude field delineates a large scale pattern of massive elongated filamentary structures. With the elongated shape of the filaments is an outstanding fundamental aspect of the filament tidal field, this goes along with a rather substantial level of inhomogeneity along the filaments, with high amplitude peaks marking the immediate high-density environment of massive clusters or the branching connections with other filaments in the weblike network. It is a direct reflection of the wide range of densities of filaments (see e.g. Cautun et al., 2014) and of the large density variations along the ridges of filaments in the cosmic web(see Cautun et al., 2014), which manifest themselves strongly in the more localised nature of the tidal field (as opposed to the more large-scale nature of the force field).

The wall tidal amplitude map shares the dominant large scale features seen in the filament map, but also includes several different and significant characteristics. While the same large scale structures in the wall tidal field can be seen, their contrast is substantially less than that seen in the filament induced tidal field. Also the tidal strength induced by the walls shows far less internal variation, yielding a more coherent tidal field within the outlined structures. This is a direct consequence of the mass density in walls being more uniform and spanning a much narrower range than that of the filament population (see e.g. Cautun et al.,

2014). By far the most outstanding difference to the filament tidal field, in addition to the degree of coherence, is the presence of far more small-scale structure filling up the space between the dominant large scale structures. Throughout the entire volume, it outlines a more intricate weblike network marked by the tidal footprint of small walls. Often they surround or connect with each other to form the boundaries of small void regions. In implies the multiscale nature of the cosmic web to be more readily visible in the wall induced tidal field than that in the overpowering filament tidal field, the latter more dominated by the major arteries of the cosmic web.

It is the void induced tidal field that reveals a surprisingly different pattern than that seen in the filament tidal amplitude map. In a sense it extrapolates the trend seen in the wall tidal amplitude maps of the dominant presence of small scale weblike structure. It yields a rich spatial pattern marked by small scale voidlike regions, with their boundaries connected into a pervasive network, yielding an ordered assembly of small-scale voids surrounded by wall-like boundaries. The pattern is also coherent and has a rather uniform amplitude, a direct reflection of the narrow density range of these cosmic underdensities (Cautun et al., 2014). It suggests that voids are largely responsible for defining and outlining the small-scale structure of the cosmic web, and appears to imply that voids play a crucial role in the formation, development and spatial organisation of the multiscale cosmic web.

By contrast, the node tidal field offers the least significant contribution to the tidal field. Even more so than in the case of the corresponding force field, the node induced tides define a spatial distribution that is highly localised in the immediate vicinity of the cluster nodes. The field does not contain a clearly defined structure, and has the appearance of a random set of tidal monopoles, with a field strength falling off like r^{-3}.

5.6.3 Tidal field strength: statistical analysis

For a quantitative judgement over the relative and absolute role of the induced tidal fields in the cosmic web, we turn to a similar assessment as in the case of the cosmic web force field, in terms of the probability functions for the total and relative tidal force field. Overall, we find a similar hierarchy in tidal force dominance as that in the case of the force field. However, filaments are far less dominant in the tidal field than in the force field. Walls, and in particular voids, have a considerably larger impact in setting the tides in the Universe. By contrast, nodes do not have any influence outside their immediate vicinity they , despite their high density contrast sometimes in excess of a thousand.

The top panel of figure 5.17 compares the pdf of the amplitude of the tides generated by the individual cosmic web components. The amplitude $|T|^{\rm CWM}_{tot}$ is given in units of the average amplitude of the total field, $\sigma(|T|)$. The distribution

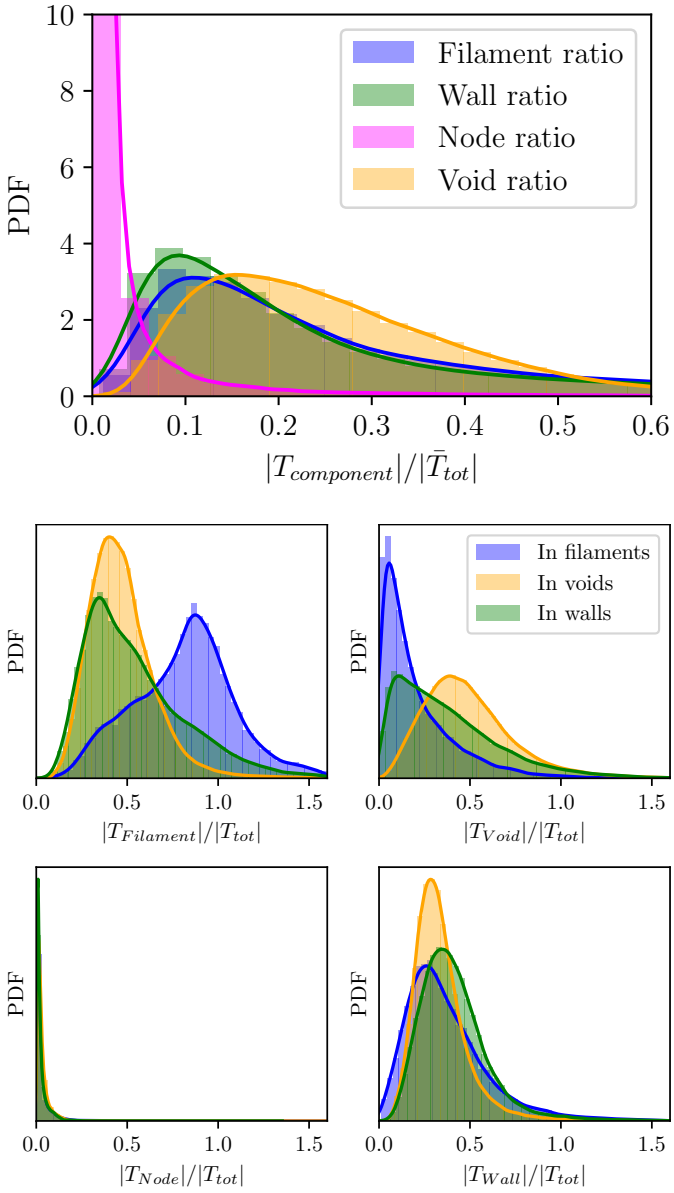


Figure 5.17: Cosmic Web Tidal Field: Strength Inventory. Top panel: The distribution of $|T|^{\mathrm{CWM}}|/|T|_{\mathrm{tot}}|$ for the different components of the cosmic web. This is the distribution for the entire field. Bottom four panels: the distribution of $|T|^{\mathrm{CWM}}|/|T|_{\mathrm{tot}}|$ for the different components of the cosmic web. Per panel the pdf for cosmic web component CWM is plotted, within three different environments: filaments (blue), voids (yellow), walls (green). Centre left: filament induced tidal field. Centre right: void induced tidal field. Bottom left: wall induced tidal field. Bottom right: cluster induced tidal field.

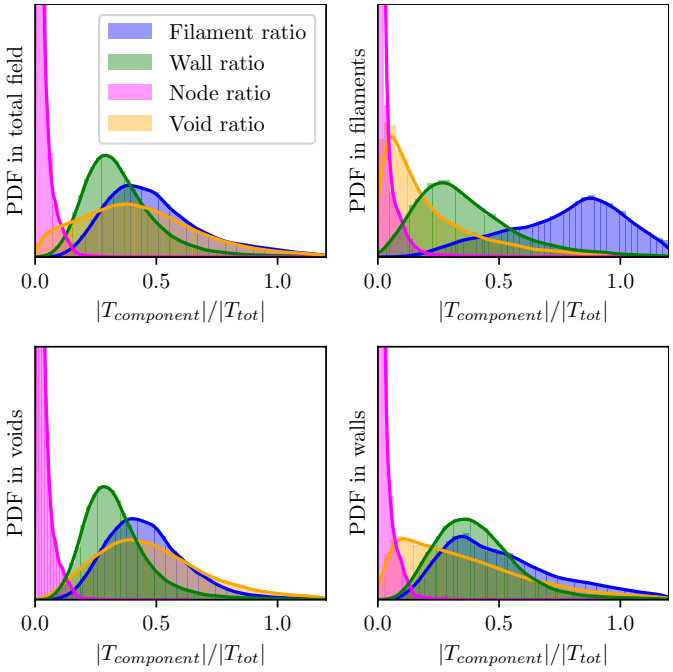


Figure 5.18: Cosmic Web Tidal Field: Strength Inventory by Cosmic Web environment. The panels show the pdf of the induced tidal field $|T|^{\mathrm{CWM}}/|T|_{\mathrm{tot}}$ by each of the four cosmic web components CWM in the different cosmic web environments. The different colours represent the different components: filament tidal field (blue), void induced tidal field (yellow), wall induced tidal field (green) and cluster node induced tidal field (magenta). Top left panel: total tidal field. Top right panel: tidal field in filaments. Bottom left panel: tidal field in voids. Bottom right panel: tidal field in walls.

of the amplitude of the induced tidal fields shows a substantial difference between the various components. Voids are by far the most ubiquitous components in the medium range of tidal values, in between $0.15 < |T|_{norm} < 0.5$. This is a clear reflection of their substantial tidal imprint, as we have already noticed in the amplitude maps, and the fact that they occupy a major fraction of the cosmic volume.

Walls and filaments have a major presence at low tidal amplitudes, with a mode near $T_{norm} \approx 0.1$, a reflection of the rapid falloff towards low tidal strengths outside of their own realm. However, both have a long tail towards high tidal values, ie. for $T_{norm} > 0.5$. It even leads to an average value substantially higher than that for voids, as may be seen in table 5.3. The filament tail reaches the highest tidal values, confirming the visual impression of figure 5.16, with the walls remaining at more moderate tidal values. The high tidal values for these

Table 5.3: Tidal field statistics. The middle two columns show the mean tidal force ratio for the different components of the cosmic web using either global or local normalisation. Note the direction dependence in the ratio $|T|^{\rm CWM}/|T|^{\rm tot}$ and $|T|^{\rm CWM}/|T|^{\rm tot}$. The fourth column shows the percentage of the volume where each component is larger than all other components.

Component	Global [%]	Local [%]	Largest in [%]
Filaments	65.3	51.5	48.5
Voids	26.9	45.2	41.0
Nodes	3.5	3.0	0.1
Walls	34.9	36.2	10.4

structures are a direct reflection and manifestation of the corresponding mass concentration in lower dimensionless geometric structures, i.e. of their elongated and flattened geometry. In turn, these are the result of their formation driven by the gravitational contraction induced by these anisotropic forces.

Telling is also the conclusion that quantitatively cluster nodes play only a minor role in setting the overall tidal force field. In most of the cosmic volume the nodes take care of only minor tidal values. The fact that their tidal pdf has a very long but low tail hints at their dominance in the minute regions of their immediate neighbourhood, but nowhere else.

Upon assessing in which environments the various component tidal influences hold sway, we notice substantial differences in the case of voids and filaments, far less so for walls and cluster nodes. Overall, the wall and node contributions to the tidal inventory of the cosmic web appear to be far less dependent on location than that of filaments and voids. For the nodes this is largely because of their spatially severely limited influence.

Evidently, the relative filament and void influence are highly dependent on their location in the cosmic web. The four panels in figure 5.17 show the distribution of the relative filament, void, cluster and wall tidal contribution in either filaments (purple), void (orange) or wall (green) environments. The top right-hand panel reveals the dominance of the filamentary induced tides within the realm of the filaments themselves, with the related panels for voids and walls indicating only a minor influence of these structures within filaments. Meanwhile, the tidal influence of voids appears to be largely restricted to that in voids themselves. Walls hardly ever are a dominant tidal source, even not within walls themselves, although they always appear to represent an omnipresent minor influence.

The latter implies that the dynamics of filaments is largely propelled by the filamentary network itself. This is emphasised in the accompanying panels in figure 5.18, which for each of the cosmic web environments shows the distribution

of the relative tidal impact by filaments, walls, voids and cluster nodes. Within filaments (top righthand panel), tidal forces are mostly due to the influence of filaments themselves, with a moderate but substantially weaker influence by walls, and far less by voids and cluster nodes. The moderate influence by walls is an expression of the close spatial and geometric relation between the anisotropic filamentary and wall-like elements in the cosmic web.

Void regions show a more chequered tidal image. From the bottom lefthand panel of figure 5.18, we see that within voids, the tidal forces induced by the voids themselves represent a major influence, usually taking care of up to 40 – 50% of the tidal force field. Nonetheless, the impact of the surrounding filaments remains important and often is even dominant (the long tail of the pdf). To a lesser extent this is also true for the walls. This often concerns the outer parts and boundary regions of the voids, where the dynamical influence of the higher density of filaments rapidly takes over the gravitational influence of the underdense voids. It emphasises the observation seen earlier with respect to the gravitational force field, the fact that the dynamics and dynamical evolution of voids cannot be understood without taking into account the external influence by in particular filaments.

5.6.4 Tidal field alignment

Given the notion that the elongated filamentary ridges and flattened walls in the cosmic web are the result of the tidally induced deformation of primordial matter concentrations, we should recognise this in the existence of an alignment of filaments and walls with respect to the tidal force field. To assess to what extent structures are aligned, we include a rough appraisal of the orientation of structures in the cosmic web with the tidal eigenvectors.

At each location we measure the alignment between the local geometry of the mass distribution and the tidal force. To this end we determine the orientation between the eigenvectors of the local tidal tensor and the eigenvectors of local inertia tensor. The latter is represented by the Hessian $\mathcal{H}(\mathbf{x})$ of the density field

$$H_{ij}(\mathbf{x}) = \frac{\partial^2 \Delta(\mathbf{x})}{\partial x_i \partial x_j}. \tag{5.23}$$

The eigenvectors of the Hessian are $\hat{e}_1$, $\hat{e}_2$ and $\hat{e}_3$, with corresponding eigenvalues

$$e_1 > e_2 > e_3 \tag{5.24}$$

$$|e| = \sqrt{e_1^2 + e_2^2 + e_3^2}. \tag{5.25}$$

A schematic diagram indicating the eigenvectors, of both tidal and inertia tensor, with respect to an elongated filament is shown in figure 5.19. Our alignment analysis investigates the orientation between the largest - compressional -

tidal eigenvector, $\hat{T}_1$ and the inertia eigenvector along the ridge of the filament, ie. the smallest density Hessian, $\hat{e}_3$. In a perfectly aligned setting, the compressional tidal eigenvector $\hat{T}_1$ would be expected to be perpendicular to the filament's ridge, and hence to $\hat{e}_3$, implying an inproduct

$$\cos(\theta) = \hat{T}_1 \cdot \hat{e}_3 = 0\,. \tag{5.26}$$

On the other hand, in the absence of any correlation between the shape of the structure and the tidal field, the orientation θ between $\hat{T}_1$ and $\hat{e}_3$ would be random and the distribution of $\mu = \cos\theta$) entirely uniform (flat).

While the diagram in figure 5.19 relates to filaments, similar geometric considerations also hold for walls and voids. In the case of walls, $\hat{e}_3$ is one of the two eigenvectors directed along the plane of a wall, while $\hat{T}_1$ is the compressional tidal component, ideally directed perpendicular to the wall. For voids, the compressional tidal component $\hat{T}_1$ is expected to be directed along the radial direction of the void (see eqn. 5.22). To first approximation, inside voids the density hardly varies along radial shells, so that the density Hessian eigenvector $\hat{e}_3$ is oriented along the transverse direction of the void.

The distribution of the orientation angle $\cos(\theta)$ between $\hat{T}_1$ and $\hat{e}_3$ for the different cosmic web environments is shown in the four panels of figure 5.20. For the entire field, as well as for the void, filament and wall environments, the panels depict the orientation pdf for the compressional component of the complete tidal force $\hat{T}_1$, as well as for the corresponding compressional components for each of the cosmic web components CWM, $\hat{T}_1$.

All panels reveal a substantial level of alignment between the compressional tidal force and the local geometry of the environment. The distribution functions peak strongly towards $\mu = \cos(\theta) = 0$, ie. we find a strongly perpendicular tendency. The one major exception is that for the node induced tidal force, which in all environments shows a mere weak alignment. It is a manifestation of the weak cluster node induced tidal field throughout the entire volume. With respect to filaments and voids the following observations are made:

- *F*ilaments
  Inside filaments, the alignment between tide and geometry can be almost exclusively ascribed to the the filament induced tidal force. To a lesser extent, the filament induced tide is responsible for a major share of alignment in walls, confirming yet again the strong structural bond between filaments and walls. Hence, the filament induced tides are not only dominant with respect to their amplitude and strength, but also with respect to their orientation and dynamical impact. On the other hand, we also see that even within filamentary environments, the filament induced tides do not account for all alignments. It hints at the still significant influence by the other cosmic web components, in particular voids and walls.

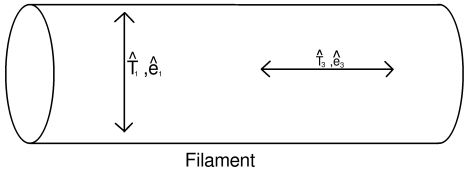


Figure 5.19: Tidal field & Cosmic Web aligment. Diagram illustration of (idealized) orientations inertial (Hessian density field) and tidal eigenvectors with respect to an elongated filament. Indicated is the orientation of the larges tidal eigenvector (T_1) and that of the smallest inertia eigenvector $\hat{e}_3$, which is aligned along the principal ridge of the filament.

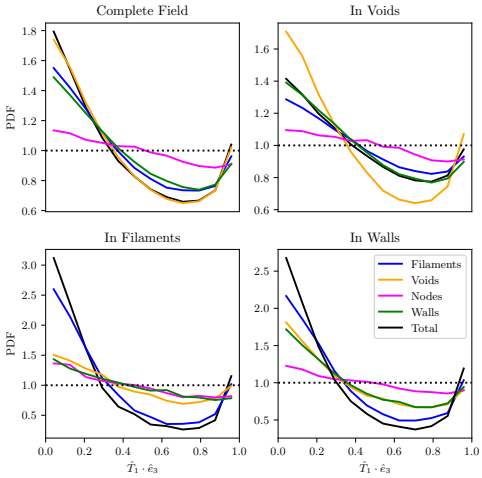


Figure 5.20: Tidal field & Cosmic Web aligment. The pdf of the alignment angle $\cos\theta = \hat{T}_1 \cdot \hat{e}_3$ between the largest (compressional) tidal eigenvector and the inertia tensor direction along the filament ridge, $\hat{e}_3$. The four panels show the pdf curves for the tidal field induced by the entire mass distribution and the specific tidal fields induced by the four morphological components CWM in four different cosmic web environments. Black curve: tidal field induced by entire mass distribution; blue curve: filament induced tidal field; orange curve: void induced tidal field; green curve: wall induced tidal field; magenta curve: cluster node tidal field. The environments in the four panels: entire field (top lefthand panel); void regions (top righthand panel); filament regions (bottom lefthand panel); wall region (bottom righthand panel). Note the different range for the vertical axis of the four panels.

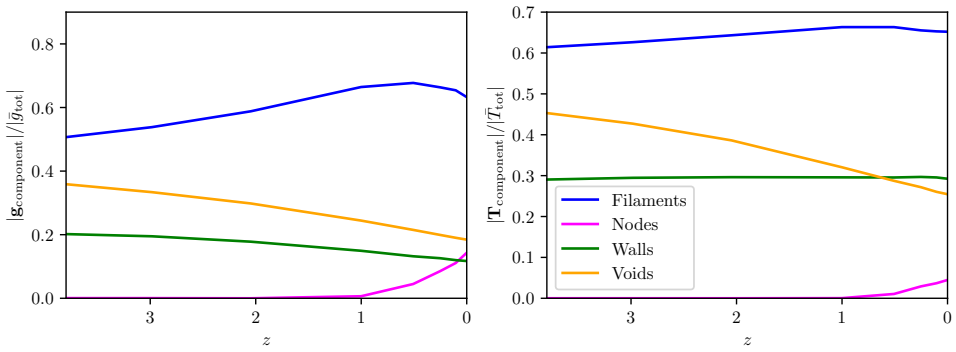


Figure 5.21: Redshift Evolution Forces and Strains in the Cosmic Web. Left frame: redshift evolution of the global average gravitational force fraction $\mathcal{F}^{\rm CWM} = |\mathbf{g}|^{\rm CWM}/|\mathbf{g}|^{\rm total}$ for the force induced by cosmic web component CWM. Right frame: redshift evolution of the global average tidal field fraction $\mathcal{F}^{\rm CWM} = |T|^{\rm CWM}/|T|^{\rm total}$ for the tide induced by cosmic web component CWM.

5

- *V*oids
  A most interesting finding is that the mass distribution inside voids is more strongly aligned with the tidal field than that in the overall field (top right-hand frame fig:tidangd). To a large extent this is to be ascribed to the strong alignment with the tidal force induced by the void itself. Also, we find that the latter has only a weak alignment with the ridge of filaments (bottom righthand frame fig. 5.20), and a mere moderate alignment with the plane of walls (top righthand frame fig:tidangd). In all, it confirms the earlier impression of the voids tidally dominating the more tenuous parts of the cosmic web.

5.7 Cosmic Web Force & Tidal Evolution

The present study has concentrated predominantly on the dynamical structure of the cosmic web at the current cosmic epoch, $z = 0$. In order to develop more definitive conclusions on the buildup of structure by the force and tidal influence of the various cosmic web components, it will be imperative not only to study the dynamics at the current time, but also to investigate their behaviour over time. While in an upcoming study we address the dynamical evolution of the cosmic web in detail, we may get a global impression by evaluating global averages of the force and tidal amplitudes as a function of redshift z.

We follow the evolution of the mean (normalized) force and tide contribution by the various cosmic web components - filaments, walls, voids and cluster nodes. To this end, for each redshift z, at each location $\mathbf{x}$ we determine the fractional contribution $\mathcal{F}^{\rm CWM}(\mathbf{x}, z)$ by a component CWM to the total force or tide

amplitude $\mathcal{A}^{\rm total}(\mathbf{x},z)$,

$$\mathcal{F}^{\rm CWM}(\mathbf{x},z) = \frac{\mathcal{A}^{\rm CWM}(\mathbf{x},z)}{\mathcal{A}^{\rm total}(\mathbf{x},z)}. \tag{5.27}$$

By normalizing with respect to the average (total) force or tide amplitude at redshift z, we compensate for the time evolution of the general force and tidal fields. We average over all locations $\mathbf{x}$ to obtain a global fractional average $\bar{\mathcal{F}}(z)$ at redshift z.

Figure 5.21 plots the redshift evolution of the fractional force contributions by cosmic web components CWM (lefthand panel) and corresponding fractional tidal amplitude contributions (righthand panel), over a redshift range $0 < z < 4$. On the basis of these rudimentary global averages, a few generic conclusions may be drawn with respect to the evolution of cosmic web dynamics.

Over the entire redshift range, filaments are consistently by far the most dominant cosmic web component, with respect to both force and tidal fields. Up to a redshift $z \approx 0.5-1$, the influence of filaments is increasing, partially the reflection of the growth of the filament population and mass as cosmic structure develops (see e.g. Cautun et al., 2014). The influence of filaments stabilises after this redshift, and may even decrease somewhat. It is tantalising to identify this with the universe transiting from a matter-dominated to a dark energy dominated regime, and the corresponding global stemming of structure growth as the Universe enters an exponential de Sitter expansion phase. Another factor may be that of the steep emergence and growth of massive cluster nodes, whose fractional force and tidal influence we see rising steeply after $z \approx 1$ (see fig. 5.21, magenta curve). This goes along with the accretion of large amounts of mass along dense filamentary branches connecting the filamentary arteries with the cluster nodes, and the subsequent increasing force influence of the nodes. At high redshifts the influence of cluster nodes is almost negligible, as their hierarchical buildup has not yet formed virialized clusters of comparable mass.

Arguably the most interesting evolutionary trend is seen for the void population. The force and tidal field influence of voids turns out to be even more prominent at higher redshifts than it is today. To some extent this may seem surprising. Because voids are smaller and less empty at higher redshifts, at first one might expect them to have a weaker influence over the cosmic mass distribution. On the other hand, the filamentary and wall-like skeleton of the cosmic web is also developing rapidly at high redshifts, and represents less mass at those redshifts. Relatively speaking, voids may therefore be more prominent at high redshifts. The details of the dynamical influence of voids, of the induced mass migration out of their interior, within the context of the hierarchically evolving void populations (Sheth & van de Weygaert, 2004; Aragon-Calvo & Szalay, 2013), therefore need further investigation. This is the subject of a related upcoming study.

Finally, walls offer a chequered view of force and tidal evolution. Their tidal influence appears to remain almost constant throughout time. Their force evo-

lution, however, appears to follow a similar downward trend as that seen for the void population. It may reflect a proportionally higher mass content at higher redshift, a consequence of the flow of mass from void troughs, via walls towards filaments.

5.8 Summary and Discussion

The cosmic web emerges at the transition between the linearly evolving cosmic matter distribution on large cosmological scales, and the highly nonlinear realm on small scales where we find matter assembled in virialized halos(Bond et al., 1996; van de Weygaert & Bond, 2008a; Cautun et al., 2014; Libeskind et al., 2018). At this transition scale, we see the emergence of complex spatial structure in the shape of an intricate weblike pattern. Long elongated filaments and flattened tenuous planar walls form the boundaries of near-empty void regions and defined a pervasive interconnected network, whose dense compact nodes are the sites where we find massive virialized halos.

With the cosmic web forming the manifestation of this key dynamical transition in the organisation of the cosmic matter distribution, it is essential to develop a more profound insight and understanding of the dynamics driving the buildup of cosmic structure, of the induced migration flows, and the gravitational influence of the various constituents of the cosmic web on these processes. This will bear upon a range of major cosmological issues. Prime is a more fundamental understanding of the formation and assembly of the cosmic web itself, and its relation to the primordial mass distribution (see e.g. Feldbrugge et al., 2018b; Feldbrugge & van de Weygaert, 2024). Of major current interest is the influence of the large scale environment on the formation and evolution of galaxies, of which the most widely recognised factor is that of the generation of angular momentum of galaxies by tidal torqueing by the same tidal forces that shape the filaments and walls in the cosmic web (e.g. Hoyle, 1949; White, 1984; Porciani et al., 2002a,b; Jones & van de Weygaert, 2009; Codis et al., 2012, 2015; Ganeshaiah Veena et al., 2019). Of increasing interest is also the potential information content of the intricate cosmic web structure on the value of cosmological parameters and the properties of dark matter and dark energy. Over the recent years in particular the structure and kinematics of cosmic voids have received ample attention for the potentially high precision with which they may reflect the character and value of dark energy (see e.g. Lee & Park, 2009; van de Weygaert & Platen, 2011; Lavaux & Wandelt, 2012; Bos et al., 2012; van de Weygaert, 2016; Pisani et al., 2019).

The present study is the first stage in a systematic investigation of the detailed dynamics involved in the emergence and assembly of the cosmic web and its various morphological components. We assess the force field and tidal field induced

by filaments, walls, cluster nodes and voids, and investigate in how far they contribute, and dominate, the gravity and tides in the various regions of the cosmic web. It facilitates an in-depth investigation of the role of the different morphological components of the cosmic web, and its environmental dependence, in driving the gravitational evolution driven formation and evolution of the largest structures in the universe.

The dynamical inventory study is based on the redshift $z = 0$ structure in a 512^3 large ΛCDM dark matter only N-body simulation in a $(300h^{-1}\text{Mpc})$ box. At each location in the simulation box the gravitational force and tidal tensor is computed for the entire mass content in the simulation volume, as well as separate for the force and tides generated by filament regions, walls, void regions and cluster nodes separately. In the current study this is based on a brute-force direct summation over the mass elements residing in the various identified morphological structures. Given the morphological identity of each mass element, we are able to do the gravitational force and tidal field analysis separately for the various cosmic web regions. Hence, it allows us to study in how far the dynamics of voids is influenced by the surrounding filaments or, reversely, in how far the gravitational force and tidal forces in filaments are feeling the presence of the surrounding voids. In addition, it enables an evaluation of the reach of void induced forces and tides, the scales over which filaments dominate, and a wide range of related issues.

For the identification of the cosmic web structures, we invoke the density field NEXUS+ multiscale morphological formalism (Aragón-Calvo et al., 2007b,a, 2010; Cautun et al., 2013, 2014), which delineate the filament, void, wall and node regions. NEXUS+ is the highest dynamic range version of the Scale Space MMF/Nexus pipeline, unique in its combination of geometric and dynamical criteria for the morphological classification and that of taking into account of the multiscale nature of the cosmic matter distribution. It enables us to identify structures over a wide spectrum of scales and densities, and hence assures an optimal assessment of the multiscale aspects of cosmic web dynamics. The NEXUS formalism assigns to each location in the simulation volume the appropriate morphology - filaments, walls, void or cluster node - following a parameter-free evaluation at the proper physical scale of the locally dominant feature.

The systematic inventory of cosmic web dynamics yielded insights that confirm existing assumptions. More tantalising are new and surprising insights with respect to the role of voids and cluster nodes. Amongst the conclusions based on the presented force and tide inventory are:

- The gravitational force fields generated by the distinct cosmic web components differ significantly from each other. The gravitational fields induced

by filaments, voids, walls and cluster nodes have markedly and systematically dissimilar characters. The overall amplitude of the gravitational force, their range, and spatial structure and pattern are quite dissimilar. In terms of the amplitude of the induced forces, there is a clear hierarchy: filaments are by far the most dominant and powerful force, followed by voids, only then - surprisingly - by cluster nodes, followed by walls.

- At nearly every location, the total gravitational force is the combination of the contributions by several cosmic web components. The important implication is that it is hardly possible to describe the dynamics and development of a structure as an isolated object. In general the gravitational influence is the result of the interplay between the several contributions. The only exceptions may be the force field in the interior of the most prominent filaments and in and around the immediate vicinity of massive cluster nodes. Yet, over most of the filamentary network also the influence of nearby voids should be taken into account, while the dynamical evolution of walls cannot be understood at all without taking into account the impact by the nearly filaments and voids. It certainly is true without any exception for all void regions, whose dynamics cannot be understood on the basis of their underdense interior but should also take into account the even more dominant exterior influence by the nearby filaments.

- Filaments dominate the gravitational force field over nearly the entire cosmic volume, with the exception of the immediate vicinity of the massive cluster nodes. On average, they are responsible for more than 50% of the exerted gravitational force at any location. Filaments' exert their influence over a large spatial range.

- While voids turn out to represent a sizeable gravitational influence over nearly the entire cosmic volume, they hardly dominate anywhere, even within their own interior. Instead, at a large fraction of space they are the second gravitational influence after filaments.

- Voids have a limited spatial range over which they wield their gravitational influence. The spatial pattern defined by the void generated gravitational field is one of a segmented volume, each segment surrounding a locally strong repulsive density trough (with radii up to 50 h^{-1}Mpc. Zooming in on individual void regions shows their superhubble expanding effect over their interior and up to their overdense boundary. Within the interstitial regions between voids, i.e. for their filamentary and planar boundaries, their gravitational impact declines rapidly to very low levels.

- Despite the mostly segmented nature of the void induced force field, still we find a residual large scale void force field. This is a manifestation of

the multiscale nature of the void population (see Sheth & van de Weygaert, 2004; Aragon-Calvo & Szalay, 2013), a direct result of its hierarchical evolution and which has recently been detected in the observational Cosmicflows-3 peculiar velocity field (Courtois et al., 2023).

- A complete dynamical description of voids in terms of isolated expanding underdense regions is not representative, The evolution of voids depends strongly on the exterior mass distribution. Void dynamics is predominantly a combined effect Of the induced superhubble expansion by voids themselves and the force and tidal influence of the the surrounding filaments (and, if in their near vicinity, the nearby cluster nodes). It shows that any consideration of void dynamics in terms of isolated almost spherically expanding regions needs a fundamental revision: the dynamics of voids cannot be understood without taking into account the gravitational and tidal influence of the environment.

- A surprising finding is that of the rather low dynamical influence by massive cluster nodes. Both the gravitational force induced by cluster nodes, as well as the corresponding tidal force, are only significant in and in the immediate surroundings of the cluster nodes, out to a radius of $\approx 5h^{-1}$Mpc. Because of their rather limited gravitational influence, they only wield a minor impact on the structure of the cosmic web.

- With respect to the corresponding gravitational tides, the situation in terms of strength is less outspoken than in the case of the gravitational force. Still, filaments are the strongest source of gravitational tides. However, voids, and even walls, yield a significant additional and alternative tidal influence. While the tidal impact of filaments has a predominantly large scale character, strongly correlated with the main elongated arteries in the cosmic web, walls and in particular voids add an essential small-scale aspect of the tidal field that is to be identified with the multiscale nature and structure of the cosmic web.

- The tidal field generated by filaments is largely a large-scale phenomenon, reflecting the outline of the most massive and prominent arteries in the cosmic web. The multiscale structure of the cosmic web can hardly be recognised in the filament induced tidal field. Filaments are the predominant source for the tidal forces inside the filaments themselves, as a result of their overwhelming strength within the filamentary ridges of the cosmic web.

- Along the filamentary ridges, the tidal field generated by filaments is rather inhomogeneous and patchy, marked by a fluctuating trend between high

amplitude and low amplitude regions along a filament's ridge. This is a direct manifestation of the rather strong mass density variations along a filament's ridge (see e.g. Cautun et al., 2014). With respect to the orientation of the tidal force field along filaments, we see a coherent pattern of aligned compressional tidal bars along the ridge of a filament.

- Voids are the dominant source of tidal influence within their interior. As voids represent the major share of a cosmic volume, void assume a major organising role in the shaping of the cosmic matter distribution.

- An interesting finding is the remarkable spatial pattern in the void induced tidal amplitude map: the void tidal field defines a coherent rather uniform cellular pattern over the entire volume, dominated by the small-scale voids in the multiscale buildup of the cosmic web. The void tidal field appears to trace the cosmic web down to even the most tenuous structures found inside voids (also see Aragon-Calvo & Szalay, 2013, with respect to a similar observation wrt. the interior void velocity field).

5

- The finding that the spatial organisation of the cosmic web is most readily expressed in the void tidal field suggests voids to be seen as the organisers of the cosmic web. It is confirmed by the orientation of the void induced tidal field (see the case studies in fig. 5.15): the compressional component appears to trace the more tenuous parts of the cosmic web, suggesting them to be responsible for the shaping of the tenuous walls and filaments that form the boundaries of the small voids. Analysis of the alignment between tidal field and mass distribution underlines this conclusion, revealing a stronger alignment within the interior of voids than in the overall cosmic web.

- The tidal force field induced by the wall population is more noticeable than its gravitational force field, arguably the reflection of their intimate relationship with the filamentary spine of the cosmic web. Both are the product of the anisotropic force field emanating from the inhomogeneous mass distribution, i.e. from the tidal force field. The wall tidal field follows the large scale spatial outline seen in the filament tidal field, tracing the heavy arteries and connecting sheets in the cosmic web. Yet, it also reflects - more faintly - the small-scale cellular pattern seen in the void induced tidal field.

While most of our analysis concerns the current cosmic matter distribution, a global statistical overview shows that the overall conclusions concerning strength of the gravitational force and tidal forces generated by the various cosmic web components pertains over a rather wide range of redshifts, at least from $z = 4$ down to $z = 0$.

In summary, we find that filaments are the main drivers of the dynamical evolution of the cosmic web. This may not be surprising, given filaments represent at least $\approx 50\%$ of the mass in the universe, concentrated in rather compact elongated ridges pervading the entire universe. More revealing is the finding of the prominent role of voids in the dynamics of the cosmic web. When looking at the cause of anisotropic structure, of which tidal fields are the main agent, we find that voids play an important role. In other words, they can be seen as the main spatial organisers of the mulitscale cosmic web.

This work has the intention to open the window on to the intricate interplay of the various gravitational influences in the cosmic web. It is the starting point of a more extensive program. Several immediate and evident practical and computational improvements of the current study are foreseen. These include a more efficient force and tidal tensor computation that would enable a more detailed statistical study and the extension of our study to large cosmological simulations including gas dynamics, baryonic physics and galaxy formation. More fundamental will be the astrophysical and cosmological issues that should be addressed on the basis of the new insights gained from the present inventory of cosmic web dynamics. Amongst a wide range of related issues, two prominent aspects are the following.

The most outstanding issue to be addressed is that of the dynamics and evolution of cosmic voids. In a follow-up study the multiscale structure of the void force and tidal field, as well as the generated mass flows, will be analysed armed with the insights obtained in the current study. The low-density environment in voids imply the major influence of external gravitational and tidal influences. As demonstrated in the present study, for most void configurations it is essential to include the role of the surrounding filamentary cosmic web if we seek to understand and exploit their structure and flow fields. This will alter our understanding of galaxy formation and evolution within void environments, given for example the impact on matter accretion. It will substantially impact on the use of voids as cosmological probes, given the fact that their internal structure and dynamics will even more highly altered by the surrounding cosmic web structure than by cosmological factors.

Another important observable effect of the cosmic web is that on the shape and rotation of galaxies. While it has long been recognised that especially the strong tidal forces exerted by (proto)filaments are the principal source of rotation for many galaxies, many recent studies have uncovered additional processes such as the anisotropic accretion along filaments that may seriously affect the outcome. In a slightly different context, the tidal forces induced by the cosmic web yield intrinsic alignments of gravitationally lensed background galaxy images and their interpretation within the cosmological context. Given the current study's finding that the tidal force field is highly sensitive to the small-scale as-

pects of the cosmic web, highlighting the presence of small voids and walls that are observationally difficult to detect, it may be necessary to focus in more detail on the intricate dynamical structure of the cosmic web and its effect on to compensate for its impact.

Acknowledgements

We are grateful to Marius Cautun and Bernard Jones for their willingness to share the Nexus code for this project, and to Joop Schaye for helpful suggestions. Many decades ago Vincent Icke taught RvdW about the importance of voids, the present study is a token of gratitude for this wise lesson. RvdW owes Dick Bond for emphasizing the key role of tidal forces to understand the formation of the cosmic web. Finally, RvdW also acknowledges Simon White for a highly motivating remark that inspired a major share of the current study. This work is partly funded by research programme Athena 184.034.002 from the Dutch Research Council (NWO).

Data Availability

All data presented in this paper will be shared upon reasonable request to the corresponding author.

References

Abel T., Hahn O., Kaehler R., 2012, MNRAS, 427, 61
Alam S., Paranjape A., Peacock J. A., 2024, MNRAS, 527, 3771
Alpaslan M., et al., 2014, MNRAS, 438, 177
Angulo R. E., Zennaro M., Contreras S., Aricò G., Pellejero-Ibañez M., Stücker J., 2021, MNRAS, 507, 5869
Aragon-Calvo M. A., Szalay A. S., 2013, MNRAS, 428, 3409
Aragon-Calvo M. A., Yang L. F., 2014, MNRAS, 440, L46
Aragón-Calvo M. A., Jones B. J. T., van de Weygaert R., van der Hulst J. M., 2007a, A&A, 474, 315
Aragón-Calvo M. A., van de Weygaert R., Jones B. J. T., van der Hulst J. M., 2007b, ApJ, 655, L5
Aragón-Calvo M. A., van de Weygaert R., Jones B. J. T., 2010, MNRAS, 408, 2163
Awad P., et al., 2023, MNRAS, 520, 4517
Bernardeau F., van de Weygaert R., Hivon E., Bouchet F. R., 1997, MNRAS, 290, 566
Bocquet S., Heitmann K., Habib S., Lawrence E., Uram T., Frontiere N., Pope A., Finkel H., 2020, ApJ, 901, 5
Bond J. R., Kofman L., Pogosyan D., 1996, Nature, 380, 603
Bonjean V., Aghanim N., Salomé P., Douspis M., Beelen A., 2018, A&A, 609, A49
Bos E., 2016, PhD thesis, University of Groningen
Bos E. G. P., van de Weygaert R., Dolag K., Pettorino V., 2012, MNRAS, 426, 440
Cautun M. C., van de Weygaert R., 2011, The DTFE public software: The Delaunay Tessellation Field Estimator code
Cautun M., van de Weygaert R., Jones B. J. T., 2013, MNRAS, 429, 1286
Cautun M., van de Weygaert R., Jones B. J. T., Frenk C. S., 2014, MNRAS, 441, 2923
Cen R., 1997, ApJ, 479, L85
Codis S., Pichon C., Devriendt J., Slyz A., Pogosyan D., Dubois Y., Sousbie T., 2012, MNRAS, 427, 3320
Codis S., Pichon C., Pogosyan D., 2015, MNRAS, 452, 3369
Colless M., et al., 2003, arXiv e-prints, pp astro–ph/0306581
Courtois H. M., et al., 2023, A&A, 673, A38
Dietrich J. P., Werner N., Clowe D., Finoguenov A., Kitching T., Miller L., Simionescu A., 2012, Nature, 487, 202
Efstathiou G., Jones B. J. T., 1980, Comments on Astrophysics, 8, 169
Einasto J., 1977, in Problems of Observational and Theoretical Astronomy. pp 26–43
Elek O., Burchett J. N., Prochaska J. X., Forbes A. G., 2020, arXiv e-prints, p. arXiv:2009.02441
Elek O., Burchett J. N., Prochaska J. X., Forbes A. G., 2022, arXiv e-prints, p. arXiv:2204.01256
Feldbrugge J., van de Weygaert R., 2023, J. Cosmology Astropart. Phys., 2023, 058
Feldbrugge J., van de Weygaert R., 2024, arXiv e-prints, p. arXiv:2405.20475
Feldbrugge J., van de Weygaert R., Hidding J., Feldbrugge J., 2018b, J. Cosmology Astropart. Phys., 2018, 027
Feldbrugge J., van de Weygaert R., Hidding J., Feldbrugge J., 2018a, J. Cosmology Astropart. Phys., 2018, 027
Feldbrugge J., Yan Y., van de Weygaert R., 2023, MNRAS, 526, 5031

Florack L., ter Haar Romeny B., Koenderink J., Viergever M., 1992, Image and Vision Computing, 10, 376
Forero-Romero J. E., Hoffman Y., Gottlöber S., Klypin A., Yepes G., 2009, MNRAS, 396, 1815
Frieman J. A., Turner M. S., Huterer D., 2008, ARA&A, 46, 385
Ganeshaiah Veena P., Cautun M., van de Weygaert R., Tempel E., Jones B. J. T., Rieder S., Frenk C. S., 2018, MNRAS, 481, 414
Ganeshaiah Veena P., Cautun M., Tempel E., van de Weygaert R., Frenk C. S., 2019, MNRAS, 487, 1607
Ganeshaiah Veena P., Cautun M., van de Weygaert R., Tempel E., Frenk C. S., 2021, MNRAS, 503, 2280
Hahn O., Carollo C. M., Porciani C., Dekel A., 2007, MNRAS, 381, 41
Hahn O., Teyssier R., Carollo C. M., 2010, MNRAS, 405, 274
Hidding J., van de Weygaert R., Vegter G., Jones B. J. T., Teillaud M., 2012, arXiv e-prints, p. arXiv:1205.1669
Hidding J., van de Weygaert R., Shandarin S., 2016, in van de Weygaert R., Shandarin S., Saar E., Einasto J., eds, IAU Symposium Vol. 308, The Zeldovich Universe: Genesis and Growth of the Cosmic Web. pp 69–76 (arXiv:1611.01221), doi:10.1017/S1743921316009650
Hirschmann M., Dolag K., Saro A., Bachmann L., Borgani S., Burkert A., 2014, MNRAS, 442, 2304
Hoffman Y., Metuki O., Yepes G., Gottlöber S., Forero-Romero J. E., Libeskind N. I., Knebe A., 2012, MNRAS, 425, 2049
Hoyle F., 1949. eds. Burgers, J.M. and van de Hulst, H.C. (Central Air Documents Office, Dayton)
Huchra J. P., et al., 2012, The Astrophysical Journal Supplement Series, 199, 26
Hunter J. D., 2007, Computing in Science & Engineering, 9, 90
Icke V., 1984, MNRAS, 206, 1P
Jones B., van de Weygaert R., 2009, Astrophysics and Space Science Proceedings, 8, 467
Jones B. J. T., van de Weygaert R., Aragón-Calvo M. A., 2010, MNRAS, 408, 897
Kitaura F.-S., Angulo R. E., Hoffman Y., Gottlöber S., 2012a, MNRAS, 425, 2422
Kitaura F.-S., Erdoğdu P., Nuza S. E., Khalatyan A., Angulo R. E., Hoffman Y., Gottlöber S., 2012b, MNRAS, 427, L35
Kourkchi E., et al., 2020, ApJ, 902, 145
Lavaux G., Wandelt B. D., 2012, ApJ, 754, 109
Lee J., Park D., 2009, ApJ, 696, L10
Lee J., Pen U.-L., 2000, ApJ, 532, L5
Lee J., Springel V., 2010, J. Cosmology Astropart. Phys., 2010, 031
Lee J., Hahn O., Porciani C., 2009, ApJ, 705, 1469
Lee K. G., et al., 2018, VizieR Online Data Catalog, p. J/ApJS/237/31
Libeskind N. I., et al., 2018, MNRAS, 473, 1195
Lindeberg T., 1994, Journal of Applied Statistics, 21, 225
Liske J., et al., 2015, MNRAS, 452, 2087
López P., Cautun M., Paz D., Merchán M., van de Weygaert R., 2021, MNRAS, 502, 5528
Macquart J. P., et al., 2020, Nature, 581, 391
McCarthy I. G., Schaye J., Bird S., Le Brun A. M. C., 2017, MNRAS, 465, 2936
Meiksin A. A., 2009, Reviews of Modern Physics, 81, 1405
Neyrinck M. C., Shandarin S. F., 2012, arXiv e-prints, p. arXiv:1207.4501
Nicastro F., et al., 2018, Nature, 558, 406
Ostriker J. P., Cen R., 1996, ApJ, 464, 27
Paranjape A., 2021, MNRAS, 502, 5210
Park D., Lee J., 2007, Phys. Rev. Lett., 98, 081301
Peebles P. J. E., 1969, ApJ, 155, 393
Peebles P. J. E., 1980, The large-scale structure of the universe
Pichon C., Bernardeau F., 1999, A&A, 343, 663
Pillepich A., et al., 2018, MNRAS, 473, 4077
Pisani A., et al., 2019, BAAS, 51, 40

Platen E., van de Weygaert R., Jones B. J. T., 2008, MNRAS, 387, 128
Porciani C., Dekel A., Hoffman Y., 2002a, MNRAS, 332, 325
Porciani C., Dekel A., Hoffman Y., 2002b, MNRAS, 332, 339
Romano-Díaz E., van de Weygaert R., 2007, MNRAS, 382, 2
Schaap W. E., van de Weygaert R., 2000, A&A, 363, L29
Schäfer B. M., 2009, International Journal of Modern Physics D, 18, 173
Schaye J., et al., 2015, MNRAS, 446, 521
Schaye J., et al., 2023, arXiv e-prints, p. arXiv:2306.04024
Schlafly E. F., et al., 2023, arXiv e-prints, p. arXiv:2306.06309
Shandarin S. F., 2011, J. Cosmology Astropart. Phys., 2011, 015
Shandarin S. F., Zeldovich Y. B., 1989, Reviews of Modern Physics, 61, 185
Shandarin S., Habib S., Heitmann K., 2012, Phys. Rev. D, 85, 083005
Sheth R. K., van de Weygaert R., 2004, MNRAS, 350, 517
Shim J., Codis S., Pichon C., Pogosyan D., Cadiou C., 2021, MNRAS, 502, 3885
Shivashankar N., Pranav P., Natarajan V., van de Weygaert R., Bos E. G. P., Rieder S., 2016, IEEE Transactions on Visualizations and Computer Graphics. 2016. Vol. 22(6, 22, 1745
Sousbie T., 2011, MNRAS, 414, 350
Sousbie T., Pichon C., Kawahara H., 2011, MNRAS, 414, 384
Springel V., 2005, MNRAS, 364, 1105
Tegmark M., et al., 2004, Phys. Rev. D, 69, 103501
Tempel E., Stoica R. S., Saar E., 2013, MNRAS, 428, 1827
Tempel E., Libeskind N. I., Hoffman Y., Liivamägi L. J., Tamm A., 2014a, MNRAS, 437, L11
Tempel E., Stoica R. S., Martínez V. J., Liivamägi L. J., Castellan G., Saar E., 2014b, MNRAS, 438, 3465
Tully R. B., Shaya E. J., Karachentsev I. D., Courtois H. M., Kocevski D. D., Rizzi L., Peel A., 2008, ApJ, 676, 184
Tully R. B., Courtois H. M., Sorce J. G., 2016, AJ, 152, 50
Wang X., Szalay A., Aragón-Calvo M. A., Neyrinck M. C., Eyink G. L., 2014, ApJ, 793, 58
Welker C., et al., 2020, MNRAS, 491, 2864
White S. D. M., 1984, ApJ, 286, 38
White S. D. M., Frenk C. S., Davis M., Efstathiou G., 1987, ApJ, 313, 505
Wilde M. C., Elek O., Burchett J. N., Nagai D., Prochaska J. X., Werk J., Tuttle S., Forbes A. G., 2023, arXiv e-prints, p. arXiv:2301.02719
Wilding G., 2022, PhD thesis, University of Groningen, doi:10.33612/diss.250010290
Wilding G., Nevenzeel K., van de Weygaert R., Vegter G., Pranav P., Jones B. J. T., Efstathiou K., Feldbrugge J., 2021, MNRAS, 507, 2968
Zeldovich Y. B., 1970, A&A, 500, 13
Zhang B., Lee K.-G., Krolewski A., Shi J., Horowitz B., Kooistra R., 2023, ApJ, 954, 49
de Graaff A., Cai Y.-C., Heymans C., Peacock J. A., 2019, A&A, 624, A48
de Lapparent V., Geller M. J., Huchra J. P., 1986, ApJ, 302, L1
de la Torre S., et al., 2013, A&A, 557, A54
van de Weygaert R., 2016, Voids and the Cosmic Web: cosmic depression & spatial complexity. pp 493–523 (arXiv:1611.01222), doi:10.1017/S1743921316010504
van de Weygaert R., Babul A., 1994, ApJ, 425, L59
van de Weygaert R., Bernardeau F., 1998, in Proceedings of the 12th Potsdam Cosmology Workshop. pp 207–216 (arXiv:astro-ph/9803143), doi:10.48550/arXiv.astro-ph/9803143
van de Weygaert R., Bond J. R., 2008a, Clusters and the Theory of the Cosmic Web. p. 335, doi:10.1007/978-1-4020-6941-3_10
van de Weygaert R., Bond J. R., 2008b, Clusters and the Theory of the Cosmic Web. p. 335, doi:10.1007/978-1-4020-6941-3_10
van de Weygaert R., Platen E., 2011, in International Journal of Modern Physics Conference Series. pp 41–66 (arXiv:0912.2997), doi:10.1142/S2010194511000092
van de Weygaert R., Schaap W., 2009, in Martínez V. J., Saar E., Martínez-González E., Pons-Bordería M. J., eds, , Vol. 665, Data Analysis in Cosmology. pp 291–413, doi:10.1007/978-3-540-44767-2_11

van de Weygaert R., van Kampen E., 1993, MNRAS, 263, 481
van de Weygaert R., Shandarin S., Saar E., Einasto J., eds, 2014, Proceedings, IAU Symposium 308: The Zeldovich Universe: Genesis and Growth of the Cosmic Web: Tallinn, Estonia, June 23-28, 2014

Appendix

5.A MMF/Nexus & Cosmic Web Classification

The NEXUS suite of cosmic web identifiers represents an elaboration and extension of the original Multiscale Morphology Filter (Aragón-Calvo et al., 2007b, 2010) algorithm and was developed with the goal of obtaining a more physically motivated and robust method. NEXUS+ is the principal representative of the full NEXUS suite of cosmic web identifiers (see Cautun et al., 2013). These include the options for multiscale morphology identifiers on the basis of the linear density, the logarithmic density, the velocity divergence, the velocity shear and tidal force field. NEXUS has incorporated these options in a versatile code for the analysis of cosmic web structure and dynamics following the realisation that they are significant physical influences in shaping the cosmic mass distribution into the complexity of the cosmic web.

5.A.1 Hessian Geometry and Morphological Identity

The basic setup of MMF/Nexus is that of defining a four-dimensional scale-space representation of the input field $f(\vec{x})$. In nearly all implementations this achieved by means of a Gaussian filtering of $f(\vec{x})$ over a set of scales $[R_0, R_1, ..., R_N]$,

$$f_{R_n}(\vec{x}) = \int \frac{\mathrm{d}^3 k}{(2\pi)^3} e^{-k^2 R_n^2/2} \hat{f}(\vec{k}) e^{i\vec{k}\cdot\vec{x}}, \tag{5.28}$$

where $\hat{f}(\vec{k})$ is the Fourier transform of the input field $f(\vec{x})$. The Hessian $H_{ij,R_n}(\vec{x})$ of the filtered field on the scale R_n is computed in Fourier space on the basis of the corresponding Fourier components $\hat{H}_{ij,R_n}(\vec{k})$,

$$\begin{aligned} H_{ij,R_n}(\vec{x}) &= R_n^2 \frac{\partial^2 f_{R_n}(\vec{x})}{\partial x_i \partial x_j}. \\ \hat{H}_{ij,R_n}(\vec{k}) &= -k_i k_j R_n^2 \hat{f}(\vec{k}) e^{-k^2 R_n^2/2}. \end{aligned} \tag{5.29}$$

Note that the definition for the Hessian includes the normalisation term R_n^2. The key element of the MMF/Nexus formalism is the morphological information contained in the eigenvalues of the Hessian matrix, $h_1 \leq h_2 \leq h_3$. By applying a set of

morphology filters on these scaled eigenvalues (see Aragón-Calvo et al., 2007a; Cautun et al., 2013) this is translated into a scale dependent environment signature $\mathcal{S}_{R_n}(\mathbf{x})$ that represents the geometry at the corresponding scale.

5.A.2 Scale Space and Multiscale Structure

To analyse the multiscale nature of the cosmic web, the *Scale-Space* representation of the cosmic mass distributions produces a sequence of copies of the data having different resolutions (Florack et al., 1992; Lindeberg, 1994). At each location $\vec{x}$ in the probed volume, it involves an extra dimension, scale, that yields the eigenvalues of the Hessian filtered at the corresponding (Gaussian) scale and the scale dependent environment signature $\mathcal{S}_{R_n}(\mathbf{x})$.

A feature searching algorithm is applied to the combined set of scaled copies in order to identify the scale at which, locally, we find the strongest morphological signature. It involves the combination of the complete set of scale-dependent environmental signatures to find the maximum signature for all scales

$$\mathcal{S}(\mathbf{x}) = \max_{\text{levels n}} \mathcal{S}_{R_n}(\mathbf{x}). \tag{5.30}$$

5

5.A.3 Signature & Versions

The final step in the MMF/Nexus procedure involves the use of criteria to find the threshold signature that discriminates between valid and invalid morphological detections. Signature values larger than the threshold correspond to real structures while the rest are spurious detections. Different implementations and versions of the MMF/Nexus technique may differ in the definition of the threshold values.

The final outcome of the MMF/Nexus procedure is a field which at each location $\vec{x}$ specifies what the local morphological signature is, cluster node, filaments, wall or void. The resulting field $\delta_j^{\text{NEXUS}}(\mathbf{x})$ is zero when the volume is not identified as cosmic web element j and is one when the volume elements is identified as element j. Here j is either filaments, nodes or walls. In this identification we also intrinsically include the identification for voids which is defined as the volume elements that are neither a filament, node or wall.

Following the basic version of the MMF technique introduced by Aragón-Calvo et al. (2007a) it was applied to the analysis of the cosmic web in simulations of cosmic structure formation (Aragón-Calvo et al., 2010) and for finding filaments and galaxy-filament alignments in the SDSS galaxy distribution (Jones et al., 2010). The principal technique, and corresponding philosophy, has subsequently been branched in several further elaborations and developments Cautun et al. (2014); Aragon-Calvo & Yang (2014). Nexus (Cautun et al., 2014) has extended the MMF formalism to a substantially wider range of physical agents involved in the formation of the cosmic web, along with a substantially firmer

foundation for the criteria used in identifying the various web-like structures. MMF-2 (Aragon-Calvo & Yang, 2014) focuses with even more attention than the basic MMF formalism on the hierarchical nature of the cosmic web, by introducing and exploiting the concept of *hierarchical space*.

The NEXUS suite of cosmic web identifiers represents an elaboration and extension of the original Multiscale Morphology Filter (Aragón-Calvo et al., 2007b, 2010) algorithm and was developed with the goal of obtaining a more physically motivated and robust method. NEXUS+ is the principal representative of the full NEXUS suite of cosmic web identifiers (see Cautun et al., 2013). These include the options for multiscale morphology identifiers on the basis of the linear density, the logarithmic density, the velocity divergence, the velocity shear and tidal force field. NEXUS has incorporated these options in a versatile code for the analysis of cosmic web structure and dynamics following the realisation that they are significant physical influences in shaping the cosmic mass distribution into the complexity of the cosmic web.

5.A.4 NEXUS+

NEXUS+ is the density field NEXUS version with the highest dynamic range. As input it takes a regularly sampled density field. In a first step, the input field is Gaussian smoothed using a LOGFILTER filter that is applied over a set of scales $[R_0, R_1, ..., R_N]$, with $R_n = 2^{n/2} R_0$. It produces the logarithmic density field

$$\delta_+ = \log(1 + \delta(\mathbf{x})), \tag{5.31}$$

The logarithmic density field of NEXUS+ is better equipped to take account of the wide dynamic range of the nonlinear hierarchically evolved density field. The nonlinear field is highly non-Gaussian, with a large part of the volume having low-density values in combination with long high-density tails in the high-density cluster and filament regions. It translates into a nonlinear density field pdf that approaches a lognormal or skewed lognormal function.

For each of the included scale-space scales, NEXUS+ computes the eigenvalues of the Hessian matrix of the smoothed logarithmic density field. Using the Hessian eigenvalues of these, NEXUS+ computes an environmental signature for each volume element that characterises how close this region is to an ideal knot, filament and wall. Then, for each point, the environmental signatures computed for each scale are combined to obtain a scale independent signature.

In the last step, physical criteria are used to determine a detection threshold. All points with signature values above the threshold are valid structures. For knots, the threshold is given by the requirement that most knot-regions should be virialized. For filaments and walls, the threshold is determined on the basis of the change in filament and wall mass as a function of signature. The peak of

the mass variation with signature delineates the most prominent filamentary and wall features of the cosmic web.

The NEXUS+ algorithm performs the environment detection by applying the above steps first to knots, then to filaments and finally to walls. Each volume element is assigned a single environment characteristic by requiring that filament regions cannot be knots and that walls regions cannot be either knots or filaments. The remaining regions are classified as voids.

Nederlandse samenvatting

Wanneer men denkt aan de voorspellingen van Einsteins algemene relativiteitstheorie gaat de aandacht meestal naar het buigen van licht, zwarte gaten en recent, zwaartekrachtsgolven. In deze thesis wordt er gebruikt gemaakt van een andere voorspelling van Einstein, in dit geval voor het hele universum. Wanneer je aanneemt dat het universum op grote schaal uniform en isotroop is, in andere woorden, het universum is hetzelfde waar je je ook bevindt, en welke richting je ook op kijkt, dan voorspellen de vergelijkingen van Einstein een universum dat uitdijt of krimpt. Verder vormen er door zwaartekracht ook structuren in het Universum. De theorie die de samenhang tussen het uitdijende heelal en de structuren die vormen door zwaartekracht voorspelt heeft als naam kosmologie.

Introductie

Binnen de studie van kosmologie bevinden we ons op een interesant tijdstip. De afgelopen twee decennia bijna zijn al onze kosmologische waarneming te voorspellen aan de hand van een het model ΛCDM. Hier staat de Λ voor donkere energie en CDM voor koude donkere materie. Samen verzorgen deze twee componenten rond de 95% van de energiedichtijd van het Universum. Met behulp van ΛCDM kunnen we de uitdijing van het Universum, het ontstaan van structuur op grote schaal in het kosmische web en het ontstaan van sterrenstelsels tot in detail voorspellen.

Het kosmische web is een van de meest fascinererende structuren in het Universum. Op de schaal tussen uniformiteit op hele grote schaal, en de vorming van sterrenstelsels op kleinere schaal. Op deze schaal ontstaat er door zwaartekracht een complex, verbonden patroon. Het kosmische web kunnen we opsplitsen in vier componenten, afhankelijk van de hoeveelheid dimensies waar zwaartekracht de structuur heeft samengetrokken. De componenten zijn, compacte nodes, langwerpige filamenten, afgevlakte muren en grote, uitdijende leegtes.

Binnen ΛCDM zijn er een aantal parameters die we aan de hand van observaties kunnen meten. Tussen veel waarnemingen hoeven we deze parameters niet of amper aan te passen om de theorie de laten overeenkomen met wat we zien. Hier is recent alleen verandering in aan het komen. Als we waarnemingen van het jonge Universum vergelijken met waarnemingen van het nabije Universum

vinden we spanningen tussen de waardes van de parameters die worden gemeten. Een van deze spanningen is in de parameter σ_8. Deze parameter bescrhift de hoeveelheid klontering in het Universum. Bij een grotere waarde is materie het Universum meer klonterig en bij een lagere waarde is materie meer egaal.

Om dit soort spanningen beter te kunnen gebruiken maken we gebruiken van simulaties. In zulke simulaties kunnen we aan de hand van het kosmologisch model de evolutie van materie in het universum voorspellen. Binnen kosmologische hydrodynamische simulaties simuleren we structuur tot op de schaal van de processen in sterrenstelsels, zoals supernova en superzware zwarte gaten. Om simulaties te doen op een grote schaal moeten we de resolutie van de simulatie verlagen. Hierdoor is het onmogelijk van nature sterren en zwarte gaten in de simulatie te laten ontstaan. In plaats daarvan worden deze belangrijke processen toegevoegd aan de hand van sub-resolutiemodellen. Deze modellen bootsen het effect dat deze processen hebben op de structuur die wel natuurlijk in de simulatie voorkomt na. Vooral de energie die vrijkomt als een superzwaar zwart gat gas opslokt heeft ook een meetbaar effect op onze metingen van kosmologie. Het is dus belangrijk dat we het effect dan deze processen hebben goed begrijpen.

In dit proefschrift wordt gebruikgemaakt van de FLAMINGO simulaties. De FLAMINGO simulaties zijn een serie simulaties in hele grote volumes. Verder zijn er binnen FLAMINGO variaties van kosmologie, sub-resolutiemodel parameters en resolutie. Hiermee kan FLAMINGO gebruikt worden om te zien wat het effect is van deze processesen op onze kosmoligische waarnemingen.

Een van de kosmologische waarnemeningen waar FLAMINGO erg geschikt voor is zijn observaties van clusters van sterrenstelsels. Door het grote volume was binnen FLAMINGO wordt gesimuleerd kan je statistisch robuuste conclusies trekken over clusters. In het universum ontstaan structuren hiërarchisch, eerst ontstaan er kleinere structure en die vallen samen om steeds grotere structuren te laten ontstaan. Hierdoor zijn clusters de laatste structuren die ontstaan, en de hoeveelheid clusters is daarmee afhankelijk van de gehele evolutie van het Universum. Door het tellen van clusters in waarnemening kunnen we meer leren over de kosmologische parameters en het onderliggende kosmologische model.

Wat clusters maken lastig maakt is dat we clusters eerst moeten identificeren in de observaties. Omdat clusters zulke grote objecten zijn kunnen we ze observeren met veel verschillende type waarnemingen, bijvorbeeld X-rays, Compton-Y[1], en het tellen van de hoeveelheid sterrenstelsels die zich in de clusters bevinden. Wat zo'n soort selectie moeilijk maakt is dat we een heel goed model nodig hebben die beschrijft welke clusters we well kunnen zien, en welke clusters we niet kunnen zien. Alleen als dat model goed geijkt is kunnen we ki-

[1] Compton-Y is de meetbare eenheid uit het Sunyaev-Zeldovich effect. Wanneer fotonen uit de achtergrondstraling door een cluster van sterrenstelsels gaan komt een deel in contact met de losse electronen, hierdoor wordt het signaal van de achtergrondstraling subtiel verandered. De magnitude van de effect is afhankelijk van de elektronendruk, wat hoger is voor grotere clusters.

jken of de hoeveelheid clusters die we zien overeenkomt met het aangenomen kosmologische model.

Omdat het selecteren van clusters veelal wordt gedaan aan de hand van gas of sterren, is dit een plek waar hydrodynamische kosmologische simulaties kunnen helpen. Omdat die simulaties ook voorspellen of de clusters waarneembaar zijn, kunnen ze ons informeren of wat voor systematische verschillen er zijn tussen wat we observeren, en wat we verwachten aan de hand van onze modellen.

Dit proefschrift

Hoofdstuk 2: FLAMINGO: Het ijken van grote kosmologische hydrodynamische simulaties met machine-learning

Zoals eerder genoemd is het belangrijk om de sub-resolutiemodellen van kosmologische simulaties te ijken. Omdat deze processen niet natuurlijk voorkomen in de simulatie, zijn er grote onzekerheden over hoeveel impact deze modellen hebben op de structuur die wel natuurlijk voorkomt in de simulatie. Het grootste voorbeeld hiervan zijn de modellen van supernova en superzware zwarte gaten. Deze processen zijn belangrijk omdat ze ervoor zorgen dat sterrenstelsels niet oneindig gas kunnen opnemen. Het intergalactisch gas koelt langzaam af, en valt dan in sterrenstelsels. Als deze processen niet zouden bestaan zou dit zorgen voor enorme sterrenstelsels. Supernovae en superzware zwarte gaten zorgen ervoor dat dit gas weer verhit wordt, en uit het sterrenstelsel wordt geduwt. Dit process treedt vooral op als er veel gas in het sterrenstelsel zit. Hierdoor ontstaat er een soort feedback loop waar gas koelt, zorgt voor supernovae en energie vanuit zwarte gaten en daardoor weer uit het sterrenstelsel wordt geduwt.

De sterkte van dit process is afhankelijk van een aantal numerieke parameters die geijkt moeten worden. Voor simpele modelen zouden deze makkelijk geoptimaliseerd kunnen worden. Het probleem is dat hiervoor het model duizenden keren moet worden geevalueerd, iets wat niet mogelijk is met kosmologische hydrodynamische simulaties. Om dit op te lossen maken we gebruik van machine-learning, getrained op tweeendertig kleinere simulaties. Aan de hand van deze simulaties kan het machine-learning model leren hoe de simulatievoorspellingen veranderen als de parameters veranderen. Dit model kan dan geoptimaliseerd worden aangezien het vele duizenden keren sneller is dan de simulaties zelf.

We gebruiken deze opstelling om ervoor te zorgen dat de hoeveelheid sterrenstels als functie van stermassa en de hoeveelheid gas in clusters overeenkomt met waarnemingen. Hierdoor zorgen we ervoor dat de FLAMINGO simulaties realistische sterrenstelsels en clusters heeft. Verder gebruiken we het model om lichte variaties de ijken. In deze variaties veranderen we de parameters aan de hand van systematische veranderingen in de data. Op deze manier kunnen we de variaties gebruiken als een onzekerheidmarge voor onze sub-resolutiemodellen.

Hoofdstuk 3: Het FLAMINGO Project: Een vergelijking van de clusters van sterrenstelsels die worden gevonden met massa, X-ray lichtsterkte, Compton-Y of sterrenstelsels abudantie

Een van de grootste uitdagingen bij het modelleren van de hoeveelheid zichtbare clusters is het in kaart brengen van welke clusters wel te zien zijn en welke niet te zien zijn. Dit noemen we de selectiefunctie. Als de selectiefunctie niet goed in kaart word gebracht, kan dit niet alleen zorgen voor het verkeerd inschatten van de massas van de clusters die zijn gevonden, maar ook dat secundaire clustereigenschappen, zoals de temperatuur of massa in gas, systematische afwijkingen hebben.

Om deze systematische afwijkingen in kaart te brengen kijken we naar het verschil in de verdeling van massa's als we clusters selecteren met massa, X-rays, Compton-Y en sterrenstelsel abundantie. Hier vinden we verschillen tussen de massa verdelingen die je krijgt als je een andere selectiemethode gebruikt. Voor clusters op lage roodverschuiving, en met massa's rond de 10^{14} $M_{\odot}$, zorgen alle methoden voor vergelijkbare resultaten. Voor hogere roodverschuivingen en lage en hoge massa's zitten selecties met Compton-Y het dichtst bij een massa-selectie en hebben hierdoor het minst last van grote systematische afwijkingen.

Verder zijn de verschillen tussen selectiemethoden als het gaat over secundaire eigenschappen, zoals of het cluster een koude kern heeft, of de cluster dynamisch niet in evenwicht is, veel kleiner dan verwacht uit de observaties. Wel zijn de massafractie in gas, en de temperatuur lichtelijk afhankelijk van de selectiemethode. De minst zware clusters die worden geselecteerd kunnen hierdoor een waarde hebben boven het gemiddelde voor hun massa.

Hoofdstuk 4: Het FLAMINGO project: Beoordeling van de systematische fouten in het tellen van sterrenstelsels door modelaannames

Een ander aspect van het gebruiken van clusters van sterrenstelsels voor kosmologie, is dat er accurate voorspellingen gedaan moeten worden voor de hoeveelheid clusters die verwacht worden zichtbaar te zijn voor een specifieke observatie. Om deze voorspellingen te doen worden er een paar anames gedaan:

- We kunnen een accuraat model maken voor de hoeveelheid clusters in het Universum als een functie van massa.
- Voor elke massa kunnen we voorspellen hoe helder het cluster is voor het type observatie dat we gebruiken aan de hand van een simpele relatie, een machtswet met log-normaal verdeelde spreiding.

Aan de hand van de FLAMINGO simulaties verifiëren we of deze aanames de simulaties kunnen beschrijven. Hiervoor kijken we naar de impact op de resul-

taten van drie observatie campagnes: Planck, South pole telescope en Simon's observatory wat staat voor een oudere, huidige en opkomende quantiteit van data.

Qua modellen voor de hoeveelheid clusters, zijn alleen de meest recente modellen acuraat genoeg, zelfs voor vroegere en huidige observaties. De andere anames zijn goed genoeg om geen invloed te hebben op huidige en oudere observaties. Alleen in de toekomst zal er meer werk gedaan moeten worden om al onze aanames te verfijnen.

Hoofdstuk 5: Dynamiek van het kosmische web, krachten en getijdevormen

Zoals eerder genoemt, vormt het Universum op grote schalen een complexe, ruimtelijk verbonden structuur dat we het kosmische web noemen. Omdat het kosmische web de achtergrondstructuur is waar sterrenstelsels in ontstaan, is het interessant om te kijken naar de dynamische invloed van het kosmische web. Dit doen we door het kracht- en getijdeveld op te splitsen in het komponent van elk veld wat wordt veroorzaakt door filamenten, muren, nodes en leegtes.

Als je deze splitsing maakt ontstaat er een duidelijk beeld. Voor het kracht en getijdeveld spelen filamenten de belangrijkste rol. Voor het krachtveld domineren ze ver boven alle andere komponenten. Qua getijden is het iets meer gematigd, maar blijven ze bovenaan de hierarchie. Nodes zijn alleen dominant in hun omgeving, en dragen gemiddeld niet veel bij aan het kracht en getijde veld. Leegtes hebben een interessante rol. In het krachtveld zorgen ze voor grote uitdijende regios, waar ze aan de randen een aardig bijdrage leveren aan de total krachtsom. Vooral qua getijden hebben ze een grote invloed. Hier volgt het getijde veld veroorzaakt door de leegtes tot in detail de structuur van het kosmische web. Voor de grotere filamenten en muren hebben ze een normale bijdrage, maar als enige zorgen ze voor getijden op de plekken waar de kosmische structuur vrij lage dischtheden heeft. We concluderen dat filamenten de krachten in het Universum echt drijven, alleen leegtes zorgen voor het aanbrengen van structuur op kleine schaal.

Publications

First author

1. swift-emulator: A Python package for emulation of simulated scaling relations
   Roi Kugel and Josh Borrow Published in Journal of Open Source Software, vol. 7, issue 72, id. 4240

2. FLAMINGO: calibrating large cosmological hydrodynamical simulations with machine learning
   Roi Kugel, Joop Schaye, Matthieu Schaller, John C. Helly, Joey Braspenning, Willem Elbers, Carlos S. Frenk, Ian G. McCarthy, Juliana Kwan, Jaime Salcido, Marcel P. van Daalen, Bert Vandenbroucke, Yannick M. Bahé, Josh Borrow, Evgenii Chaikin, Filip Huko, Adrian Jenkins, Cedric G. Lacey, Folkert S. J. Nobels, and Ian Vernon Published in Monthly Notices of the Royal Astronomical Society, Volume 526, Issue 4, pp.6103-6127

3. The FLAMINGO Project: A comparison of galaxy cluster samples selected on mass, X-ray luminosity, Compton-Y parameter, or galaxy richness
   Roi Kugel, Joop Schaye, Matthieu Schaller, Ian G. McCarthy, Joey Braspenning, John C. Helly, Victor J. Forouhar Moreno, Robert J. McGibbon Accepted for publishing in MNRAS, ArXiv: eprint arXiv:2406.03180

4. Cosmic Web Dynamics: Forces and Strains
   Roi Kugel and Rien van de Weygaert Submitted to MNRAS, ArXiv: eprint arXiv:2407.16489

5. The FLAMINGO Project: An assessment of the systematic errors in the predictions of models for galaxy cluster counts used to infer cosmological parameters
   Roi Kugel, Joop Schaye, Matthieu Schaller, Victor J. Forouhar Moreno, Robert J. McGibbon Submitted to MNRAS, ArXiv: eprint arXiv:2408.17217

Co-author

1. The importance of black hole repositioning for galaxy formation simulations
   Yannick M. Bahé, Joop Schaye, Matthieu Schaller, Richard G. Bower, Josh Borrow, Evgenii Chaikin, **Roi Kugel**, Folkert Nobels and Sylvia Ploeckinger Published in Monthly Notices of the Royal Astronomical Society, Volume 516, Issue 1, pp.167-184

2. The FLAMINGO project: cosmological hydrodynamical simulations for large-scale structure and galaxy cluster surveys
   Joop Schaye, **Roi Kugel**, Matthieu Schaller, John C. Helly, Joey Braspenning, Willem Elbers, Ian G. McCarthy, Marcel P. van Daalen, Bert Vandenbroucke, Carlos S. Frenk, Juliana Kwan, Jaime Salcido, Yannick M. Bahé, Josh Borrow, Evgenii Chaikin, Oliver Hahn, Filip Huko, Adrian Jenkins, Cedric G. Lacey and Folkert S. J. Nobels Published in Monthly Notices of the Royal Astronomical Society, Volume 526, Issue 4, pp.4978-5020

3. The FLAMINGO project: revisiting the S_8 tension and the role of baryonic physics
   Ian G. McCarthy, Jaime Salcido, Joop Schaye, Juliana Kwan, Willem Elbers, **Roi Kugel**, Matthieu Schaller, John C. Helly, Joey Braspenning, Carlos S. Frenk, Marcel P. van Daalen, Bert Vandenbroucke, Jonah T. Conley, Andreea S. Font and Amol Upadhye Published in Monthly Notices of the Royal Astronomical Society, Volume 526, Issue 4, pp.5494-5519

4. The FLAMINGO project: the coupling between baryonic feedback and cosmology in light of the S_8 tension
   Willem Elbers, Carlos S. Frenk, Adrian Jenkins, Baojiu Li, John C. Helly, **Roi Kugel**, Matthieu Schaller, Joop Schaye, Joey Braspenning, Juliana Kwan, Ian G. McCarthy, Jaime Salcido, Marcel P. van Daalen, Bert Vandenbroucke and Silvia Pascoli Submitted, ArXiv, eprint arXiv:2403.12967

5. Non-linear CMB lensing with neutrinos and baryons: FLAMINGO simulations versus fast approximations
   Amol Upadhye, Juliana Kwan, Ian G. McCarthy, Jaime Salcido, John C. Helly, **Roi Kugel**, Matthieu Schaller, Joop Schaye, Joey Braspenning, Willem Elbers, Carlos S. Frenk, Marcel P. van Daalen, Bert Vandenbroucke and Jeger C. Broxterman Published in Monthly Notices of the Royal Astronomical Society, Volume 529, Issue 2, pp.1862-1876

6. Inferring the dark matter splashback radius from cluster gas and observable profiles in the FLAMINGO simulations
   Imogen Towler, Scott T. Kay, Joop Schaye, **Roi Kugel**, Matthieu Schaller, Joey Braspenning, Willem Elbers, Carlos S. Frenk, Juliana Kwan, Jaime Salcido, Marcel P. van Daalen, Bert Vandenbroucke and Edoardo Altamura

Published in Monthly Notices of the Royal Astronomical Society, Volume 529, Issue 3, pp.2017-2031

7. The FLAMINGO project: baryonic impact on weak gravitational lensing convergence peak counts
   Jeger C. Broxterman, Matthieu Schaller, Joop Schaye, Henk Hoekstra, Konrad Kuken, John C. Helly, **Roi Kugel**, Joey Braspenning, Willem Elbers, Carlos S. Frenk, Juliana Kwan, Ian G. McCarthy, Jaime Salcido, Marcel P. van Daalen and Bert Vandenbroucke Published in Monthly Notices of the Royal Astronomical Society, Volume 529, Issue 3, pp.2309-2326

8. Relativistic SZ temperatures and hydrostatic mass bias for massive clusters in the FLAMINGO simulations
   Scott T. Kay, Joey Braspenning, Jens Chluba, John C. Helly, **Roi Kugel**, Matthieu Schaller and Joop Schaye Submitted, ArXiv: eprint arXiv:2404.08539

9. Dark photon constraints from CMB temperature anisotropies
   Andres Aramburo-Garcia, Kyrylo Bondarenko, Alexey Boyarsky, Pavlo Kashko, Josef Pradler, Anastasia Sokolenko, **Roi Kugel**, Matthieu Schaller and Joop Schaye Submitted, ArXiv: eprint arXiv:2405.05104

10. The FLAMINGO project: Galaxy clusters in comparison to X-ray observations
    Joey Braspenning, Joop Schaye, Matthieu Schaller, Ian G. McCarthy, Scott T. Kay, John C. Helly, **Roi Kugel**, Willem Elbers, Carlos S. Frenk, Juliana Kwan, Jaime Salcido, Marcel P. van Daalen, Bert Vandenbroucke Published in Monthly Notices of the Royal Astronomical Society, Advance Access

11. Validating the clustering predictions of empirical models with the FLAMINGO simulations
    Sergio Contreras, Raul E. Angulo1, Jonás Chaves-Montero, **Roi Kugel**, Matthieu Schaller and Joop Schaye Submitted, ArXiv: eprint arXiv:2407.18912

12. The FLAMINGO simulation view of cluster progenitors observed in the epoch of reionization with JWST
    Seunghwan Lim, Sandro Tacchella, Joop Schaye, Matthieu Schaller, Jakob M. Helton, **Roi Kugel** and Roberto Maiolino Published in Monthly Notices of the Royal Astronomical Society, Volume 532, Issue 4, pp.4551-4569

Curriculum Vitae

I was born on the 21st of January 1998 in the city of Assen, Drenthe, The Netherlands. At the age of four we moved to Stadskanaal, Groningen. I have always been curious towards nearly anything. My eagerness to learn let me skip a year in elementary school, the CBS Oranje Nassau. I followed high school education at VWO level at the Ubbo Emmius school in Stadskanaal. While I always remained curious, this feeling was rarely sparked until towards the end of high school. At this point I wasn't quite sure what I wanted to study at University. After checking out multiple open days in software engineering and chemistry, it was the airing of the modern version of the show Cosmos, featuring Neil Degrasse Tyson, that first sparked my interest in astronomy as a potential study direction.

This spark quickly turned into a definite decision. The combination of physics and programming, both big interests, combined with the incredibly warm atmosphere of the Kapteyn institute, where I followed an open day, quickly confirmed that this would be the next big step in my life. I did both my Bachelor and Master at the University of Groningen and Kapteyn institute.

During this time I really started discovering the things I really enjoy. I started listening to a lot of music, started learning guitar and slowly started becoming more and more involved in the newly formed study association for astronomy: Sirius A. This involvement lead to to eventually become chair of the Kapteyn Learning Committee, which organises additional practise material for exams. During my masters I also became the chair of the program committee, who is in charge of gathering and giving feedback on university courses.

In terms of research, ever since the first mention of cosmology during an introductory course, this has been my main topic of interest. Therefore it was logical to pursue cosmology for both my BSc and MSc thesis, and eventually the PhD. During my Bachelor thesis I looked at using the phase information of Fourier transformed large-scale structure to learn more about structure formation. During my master thesis I investigated the large-scale forces, flows and tides caused by the cosmic web.

At the end of the MSc I was 22, and I was not ready yet to start working in industry, so I looked for a PhD position. As I ideally wanted to stay in The Netherlands, I applied to the one place in The Netherlands that does cosmological simulations that include hydrodynamics. I got accepted in the group as a PhD

student to work on large-scale structure simulations for cosmology.

During my PhD I developed the calibration method for the new FLAMINGO suite of simulations. The FLAMINGO simulations are designed to have all the required outputs to be informative about a wealth of cosmological observables. For this I made use of Gaussian Process machine learning. I also published an open-source python package to make the method available for other simulation projects as well.

After finishing the calibration and running the simulations, I started working on galaxy clusters, something that was timely both in terms of the simulation we created, but also because many cluster surveys started putting out their results. I investigated both cluster selection effects and models for cluster cosmology, and became particularly interested in clusters found using the Sunyaev-Zeldovich effect.

I also followed courses like: How to manage your professor, Communication in Science, time management, advanced project management for PhDs, scientific integrity, scientific writing, the Nova Fall school, the NOVA-SKIES program and bystander training. Furthermore I visited scientific conferences, workshops and a winter school across Europe where I presented my work.

During the PhD I further expanded my worldview and hobbies. I started singing, including taking lessons, I travelled a lot and became much more appreciative of the food at different places. I also became interested in Japan and started learning Japanese. I became chair of the social committee and organised institute-wide events. This includes the yearly barbecue, where I have joined playing with a band for three years in a row.

Now being at the end of the PhD, it is time to choose the next step. While I contemplated doing a postdoc and continuing an academic career, a strong desire to stay in The Netherlands, combined with a wavering motivation for astronomy and the desire to work on something a bit more tangible, lead me to make the decision to look for a job outside of astronomy. I will start work as a principle data engineer at the green energy provider GreenChoice.

Acknowledgements

Coming to the end of a four-year journey requires some level of self-reflection, and if one thing is definitely true, it is the fact that the completion of this thesis would have been impossible without the support of everyone around me. These last four years have had a huge impact on my interests, worldview and self-image, and I will hold all these experiences very dear for the rest of my life. Had I known four years ago what I was getting into, I probably would have accepted the PhD position immediately. I will try to thank as many of you as possible, and I apologize in advance if I accidentally left people out.

I want to start by thanking my supervisors. I want to thank Joop and Matthieu for their continued support, enthusiasm, and for always being there for any questions I had. The confidence I gained working on the big projects this year will stay with me for the rest of my life. I also want to thank Rien for further supporting me throughout the PhD.

I want to thank Joey and Evgenii for allowing me to be their unofficial office mate (until we became actual office mates). Our conversations gave a shine to every day. I want to thank Jeger for his countdowns and Will for his great song recommendations. Furthermore, I want to thank Bert, Elia, Folkert, Lydia, Nastasha, Orestis, Rob, Stijn, Victor, Yanling, Yannick, and Zorry for creating the amazing environment within our group. I also want to thank all the members of the FLAMINGO collaboration for always being very helpful and for providing such a welcome environment to talk about halo finders. I also want to give a special thanks to Ian and Josh, I always really enjoyed working together.

On a personal level, I want to thank Josh, Marta, Niccolò and Nicole. The last four years you have broadened my view on food, travel, Japan, One Piece, Dnd, board games, video games, etc. all while providing an amazing level of friendship. You have truly given me a new look on things, and I will carry that with me for the rest of my life.

The same goes for all the many others that made the Sterrewacht a home for me these last four years. I want to thank Alessia, Alfred, Amy, Aniruddh, Andres, Andrew, Anne, Anniek, Bas, Beth, Bianca, Chloe, Christian, Christiaan, Ciaran, Dario, Dilovan, Elena, Elina, Erik, Floor, Fraser, Frits, Gijs, Ivana, Jelle, Jurjen, Kevin, Kirsty, Lucie, Lucas, Mantas, Markus, Martje, Osmar, Paola, Pranjal, Rico, Rob, Roland, Silvia, Sill, Thijs, Thomas, Timo, Victorine and Willeke. Without

all of you, finishing the PhD would have been much harder. I also want to give a special thanks to both the gnarrly skiing committees.

I want to thank the people from the three BBQ bands, "No Shuffle", "PhDs and the moneymaker" and "Smoke cover", and Dario and the guinea pigs for joining me in one of my favourite hobbies. I want to thank my singing teacher Dany for unlocking a completely new instrument for me. I want to thank Big Dipper for many amazing Friday evenings and I want to thank everyone in the social committee for the great things we organised.

I want to thank all the people that I have met at conferences and meetings along the way, with a special thanks to all the SWIFT related meetings. These weeks have been some of the biggest highlights of the PhD.

I want to thank "the boys from Groningen", Arno, Bram, Frank, Gerben, Jasper, Jisk, Julien and Ramon for your ability to make it feel like we last saw each other yesterday, even if the actual time between is much longer. I also want to thank the people in my guild <It's not a cult> for carrying me through the long lockdown months. I also want to thank Annely, Diya, Gijsbert, Imge and Mahi for their friendship and support.

I want to thank my family for always supporting me and always showing interest in the work I'm doing. Though it was a bit more difficult from a distance, it was still much appreciated. Finally, I want to thank Veronica, besides being a great friend you were consistently there for me when I most needed it.

While I will hopefully keep up with all of you for the foreseeable future, the end of a PhD is still a goodbye. Therefore, I want to end this thesis with a piece of wisdom from a well renowned philosopher: Winnie the Pooh,

'How lucky I am to have something that makes saying goodbye so hard.'
-A.A Milne