Mapping Dark Matter in Galaxy Clusters with Gravitational Lensing

SUT-IENG TAM

How to cite:

TAM, SUT-IENG (2020) Mapping Dark Matter in Galaxy Clusters with Gravitational Lensing. Doctoral thesis, Durham University.

Mapping Dark Matter in Galaxy Clusters with Gravitational Lensing

Sut-Ieng Tam

A thesis submitted to the University of Durham

in accordance with the regulations for
admittance to the Degree of Doctor of Philosophy.

Durham
University

Institute for Computational Cosmology
Department of Physics
University of Durham
England

August 2020

Dedicated to

my parents

Mapping Dark Matter in Galaxy Clusters with Gravitational Lensing

Sut-Ieng Tam

Abstract

This thesis uses gravitational lensing to map the distribution of dark matter around galaxy clusters, and to infer their formation history. Galaxy clusters are the oldest and most massive gravitationally-bound objects in the Universe, exploited in the most discriminating tests of cosmology. It is therefore essential to understand the astrophysics of their formation. Indeed, clusters grow through filamentary connections with surrounding large-scale structures – and to chart their history is to trace the evolution and trajectory of the Universe itself.

Gravitational lensing is the apparent distortion in the shapes of distant galaxies due to foreground mass, such as a galaxy cluster. Many software algorithms have been developed to measure gravitational lensing and to reconstruct the distribution of foreground mass. In this thesis, we assess the performance of two mass-mapping techniques, using mock images of the BAHAMAS simulation, where the true distribution of mass is known. We find the methods suitable for different applications: MRLens suppresses noise without bias, while Lenstool suppresses noise further, but at a cost of over-estimating the mass in cluster outskirts ($R > 1\,\mathrm{Mpc}$) by up to a factor 2. We also develop a filter to search for large-scale filaments connected to galaxy clusters. We then use these calibrated techniques, and the largest ever mosaic of *Hubble Space Telescope* imaging, to study galaxy cluster MS 0451-03 ($z = 0.54$). We map the distribution of its dark matter, and discover six group-scale substructures, linked to the cluster halo by three possible filaments. By comparing lensing results with analyses of X-ray emission and optical spectroscopy, we conclude that the cluster collided with another 2–7 Gyr ago. Its star formation was quenched and its gas was heated; its gas has still not yet relaxed, and the dark matter halos are approaching second apocentre.

In the next decade, space-based telescopes will reveal this richness of detail about tens of thousands of galaxy clusters. If these observations are properly calibrated, via studies like this thesis, they will bring a new era of precision cosmology. As a final step towards this future, we present preliminary results from two ongoing projects: using deep learning to further suppress noise in lensing mass reconstruction, and the first successful measurement of gravitational lensing from a balloon-borne telescope at the edge of space.

August 16, 2020

Declaration

The work described in this thesis was undertaken between 2016 and 2020 while the author was a research student under the supervision of Professor Richard Massey and Dr Mathilde Jauzac in the Department of Physics at the University of Durham. This work has not been submitted for any other degree at the University of Durham or any other University.

Parts of Chapter 3 were submitted as part of the following paper:

- D. Harvey, S.-I. Tam, M. Jauzac, R. Massey, J. D. Rhodes 2019, ArXiv Astrophysics e-prints, arXiv: astro-ph/1911.06333

Chapter 4 of this thesis has been published in the form of a paper:

- S.-I. Tam, R. Massey, M. Jauzac, A. Robertson, 2020a, MNRAS, 496, 3973.

Chapter 5 of this thesis has been published in the form of a paper:

- S.-I. Tam, M. Jauzac, R. Massey, D. Harvey, D. Eckert, H. Ebeling, R. S. Ellis, V. Ghirardini, B. Klein, J-P Kneib, D. Lagattuta, P. Natarajan, J. Richard, G. Smith, 2020b, MNRAS, 496, 4032.

The work described in Chapters 3–5 was performed in collaboration with the co-authors of the above mentioned papers. The work described in Chapter 6 was performed in collaboration with R. Massey, M. Jauzac, K.-H. Leong, D. Lagattuta and the SuperBIT team. The author has carried out all of the analysis in this work with the following exceptions: (1) the mock data from the BAHAMAS simulation in Chapter 4, which were generated by A. Robertson and I. McCarthy ; (2) the photometry catalogue and photometric redshifts of galaxies in MS 0451-03 field, which were provided by H. Ebeling; (2) the catalogue of spectroscopic redshifts in

Chapter 5, which was provided by H. Ebeling; (3) the data reduction and analysis of MS 0451-03's XMM-Newton X-ray Observations presented in Chapter 5, which were conducted by D. Eckert; (4) the data reduction of SuperBIT's Abell 2218 images and the calibration of photometries in Chapter 6, which were carried out by D. Lagattuta.

August 16, 2020

Acknowledgements

First of all, I would like to thank my supervisors Professor Richard Massey and Dr Mathilde Jauzac, who have guided, taught and encouraged me. My work has only been made possible through their support. They have spent so much time on reading and polishing my manuscripts again and again. Their endless patience and enthusiasm keep me stay positive during the whole of my PhD time.

Thanks to all my colleagues, including David Lagattuta, David Harvey, Andrew Robertson, James Nightingale and the SuperBIT team, for their assistance and contribution to the work in this thesis. It has been a pleasure working with them all. Besides, It is also my pleasure to be part of the lensing group, whose meetings I enjoyed very much. I am grateful to the COSMA support team, Lydia Heck and Alastair Basden in particular, who offer their help with regard to computational issues.

I had the opportunity to share offices with many people who made this experience more memorable, including Stuart Stach, Fang-Xia An, Helen Johnson, Ugne Dudzeviciute, Steve Gillman, Will Collier, Lizelke Klindt, Jake Mitchell, Vicky Fawcett, and Sergio Ruiz-Bonilla. My very special thanks to all of my friends who give me support through this journey, especially to Lin Tsui, Lizelke Klindt, Cesar Hernandez-Aguayo, Ka-Hou Leong, and Qiuhan He.

Finally, I am deeply thankful to my family, my dearest Mum, Dad, and Brother for the support during all my life.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Human beings have always been fascinated and thus curious of our Universe. Who has never watched the night sky and not been attracted by the starlights surrounding this darkness? From this primitive curiosity, they gradually began to investigate the underlying mechanisms of the Universe, trying to understand its evolution and composition. With the advance of science, we now realise that this luminous matter, including dust clouds, stars and galaxies, constitutes merely 5% of the Universe. The other 95% are unfortunately not directly observable from telescopes, but nevertheless exist and are subsequently called "Dark Matter" and "Dark Energy". Since this discovery, cosmologists have paid increasing attention to the dark side of the Universe.

With the invention of telescopes of always higher resolutions, every decade has brought new findings. However, the nature of dark energy and dark matter remains one of the biggest mysteries in our Universe which scientists endeavour to unlock. Current cosmological probes suggest that our Universe is expanding at an accelerating rate, and dark energy being responsible for this expansion. The other dark component, dark matter, plays a crucial role in the origin and the evolution of structures. Their distinctive natures determine the ultimate fate of our Universe. If there is lots of mass, the expansion of the Universe will eventually reverse and recollapse, causing a "Big Crunch". On the other hand, if dark energy eventually overcomes gravity, it will tear everything apart, and the Universe will end in a "Big Rip".

In this thesis, we are particularly interested in the properties of dark matter. Studying the dark matter distribution of the whole Universe is difficult because the observable universe is limited. Nevertheless, galaxy clusters, the largest gravitationally-bound structures in the universe, can trace the “matter skeleton” of our Universe on the largest scales.

Based on the hierarchical structure formation scenario of the standard cosmological model, the ΛCDM (Λ for cosmological constant and CDM for cold dark matter) paradigm, small structures arose through the growth of density perturbations of dark matter which originated from the early universe. These structures then grew into more massive objects through matter accretion along their surrounding filamentary structures. In this scenario, galaxy clusters are the most massive objects in the observable Universe, consisting of 85% dark matter, 10% hot gas (called intra-cluster medium, ICM), and 5% stars. The vast amount of dark matter makes galaxy clusters fascinating cosmological laboratories for studying the “Dark Universe”. Furthermore, the shape of the distribution of mass inside a cluster reflects the nature of dark matter (whether it is collisionless; for example as in Robertson et al., 2019). On larger scales, a cluster’s orientation is governed by accretion of matter from the surroundings. Most substructures are accreted into clusters along filaments (Angulo et al., 2012; Aragón-Calvo et al., 2007; Bond et al., 1996; Yess & Shandarin, 1996). Consequently, clusters tend to align with these directions (e.g. Jing & Suto, 2002; Warren et al., 1992). In addition, since their growth spans the entire age of the Universe, and depends upon the density of infalling material and its gravity collapse, as opposed to its disruption by supernovae, active galactic nuclei, and dark energy, measurements of the precise number and properties of clusters represents highly sensitive tests of the standard cosmological model (e.g. Bahcall & Bode, 2003; Fluri et al., 2019; Ho et al., 2006; Jauzac et al., 2016; Mao et al., 2018; Rozo et al., 2010; Schwinn et al., 2017; Weinberg et al., 2015).

Accurate measurements of the mass and internal structure of clusters are key to unlock the mystery of dark matter. Although the majority of matter in clusters is not directly observable, the total mass along the line of sight can be mapped, regardless of its physical or dynamical state, via measurements of the gravitational lensing of the light emitted by background objects (behind clusters). The strong lensing effect causes a large distortion of background sources, producing multiple images or giant arcs. It probes the mass distribution in the inner core of clusters, while weak lensing provides constraints on the larger-scale environment. The phenomenon of gravitational lensing is predicted by General Relativity. The dense concentration of mass in a foreground galaxy cluster deflects light rays emitted by unrelated galaxies far in the background. Since adjacent light rays are almost coherently deflected, the shapes of those distant galaxies appear distorted, and typically stretched tangentially around the cluster. Crucially, the deflection of light rays depends only upon the total projected mass distribution. Measurements of gravitational lensing are therefore uniquely sensitive to the distribution of invisible-but-dominant dark matter, and unbiased by the nature and dynamical state of ordinary matter (e.g. Bartelmann & Maturi, 2017; Hoekstra, 2013; Kilbinger, 2015; Kneib & Natarajan, 2011b; Massey et al., 2010; Treu & Ellis, 2015).

This thesis is centered on the study of the weak gravitational lensing in galaxy clusters, aiming to constrain the physical properties of galaxy clusters by studying the total matter content of galaxy clusters and its surrounding environments. With the capabilities of the next-generation of telescopes, more clusters with wide-field and space-resolution observations will be made available. With regard to cluster mass distribution specifically, it is conducive to future research to look for an optimum method to convert weak lensing shear into precise mass distribution of lenses. Therefore, we first quantified two mass mapping techniques using data from the BAHAMAS simulation, and find the one suitable to the real galaxy clusters. Afterwards, we conducted a combined strong and weak lensing analysis of a massive galaxy cluster, MS 0451-03. Its wide-field mosaic of imaging data with the *HST* allowed us to detect possible large-scale filament directions extending from cluster centre, and funneling matter into its core. Significant properties of this cluster and its dynamical state were then inferred.

This thesis is structured as follows. In Chapter 2 I present a brief review of some aspects of cosmology which are relevant to our later discussions. Chapter 3 presents the theoretical basis of gravitational lensing and describes in detail the weak lensing shape measurement pipeline pyRRG used in this study. The details of the two weak lensing mass mapping techniques are also described. In Chapter 4, we quantify the performance of these two techniques using the BAHAMAS simulation data, and apply halo shape measurements and filament searches on the final mass reconstructions. Chapter 5 presents the combined strong and weak lensing analysis of the massive galaxy cluster, MS 0451-03. I also present the constrained X-ray and lensing analyses to infer its dynamic state. In Chapter 6, I present some preliminary results from a new mass reconstruction method with machine learning and the wide-field balloon-borne imaging telescope, SUPERBIT (Super-pressure Balloon-born Imaging Telescope). Finally, I summarize the discoveries and the results of this thesis in Chapter 7.

Chapter 2

Cosmological background

2.1 The Standard Model of Cosmology

2.1.1 ΛCDM model

Based on the current observational evidence from multiple cosmological probes, our Universe is well described by the ΛCDM model, where Λ stands for dark energy with constant energy density, and CDM for Cold Dark Matter. These two "dark" components determine the evolutionary fate of our Universe.

The ΛCDM Universe begins with a hot big bang and has been expanding since then, a prediction which was confirmed by the observations of distant type Ia supernovae (Perlmutter et al., 1998; Riess et al., 1998). The acceleration of the expansion rate nowadays is due to dark energy, which constitutes the majority ($\sim$ 70%) of the total energy density. This mysterious component causes a constant push of emptiness between masses, resulting in accelerating cosmic expansion over time.

The second important component is dark matter, which makes up $\sim$ 25% of the Universe, while only $\sim$ 5% of the energy density consists of normal baryonic matter. Dark matter seems to only interact via gravity, and has a very small electroweak interaction and self-interaction cross section, hence it does not emit light at any wavelength. This invisible matter plays an essential role in structure formation since its gravity holds matter together to collapse and form structure. Zwicky (1933) first proposed the existence of dark matter by studying the dynamics of galaxies in the

Coma cluster and other clusters. He found a high peculiar velocity which required a cluster to be about 100 times more massive than expected from its luminous contents to keep the galaxies bound. Another famous observational evidence of dark matter is given by the study of the merging Bullet cluster (Clowe et al., 2004), who discovered non-identical distributions of total mass and X-ray emitting gas. This finding supported the collisioness nature of dark matter: the two main cluster halos passed through each other and decelerated because of being affected by gravitational force only, whereas the collisional gas experienced hydrodynamical friction, causing large disturbances and shocks. Bullet clusters are useful to constrain the properties of dark matter. Robertson et al. (2017) performed numerical simulations of the merging galaxy cluster, including the effects of elastic dark matter scattering. They found that self-interacting dark matter transfers momentum between two merging halos, leading them, similar to the gas distribution, to lag behind the collisionless galaxies. Therefore, the offsets between the dark matter peaks and those of the baryonic matter (ICM and galaxies) offer a possibility to measure the self-interacting cross section of dark matter (e.g. Markevitch et al., 2004; Massey et al., 2015). In particular, Harvey et al. (2015) combined the measurements of 72 colliding systems and constrained the dark matter's momentum transfer cross-section to be $\sigma_{\rm DM}/m < 0.47{\rm cm}^2/{\rm g}$.

While many studies confirm the existence of missing mass in our Universe, the nature of dark matter remains unsettled. Traditionally, there are three kinds of dark matter: hot, warm and cold. Their temperature denotes how fast they can move in the very early Universe. Cold dark matter is non-relativistic and is favoured by modern observations. Hot dark matter particles travel with relativistic velocity, so that they can escape from small mass density fluctuations, resulting in a relatively slow structure formation. Warm dark matter has intermediate properties.

Instead of introducing an unknown matter, the missing mass problem can also be solved by a modification of the law of gravity at large distances, such as Modified Newtonian Dynamics (MOND; Milgrom, 1983). Nevertheless, ΛCDM is currently the standard model that successfully describes and recovers most of the current observations. Therefore, in this thesis, our cosmology assumptions are based on

ΛCDM.

2.1.2 Friedmann-Robertson-Walker matrix

In the standard cosmological model, our Universe is assumed to be homogenous and isotropic on large scales, meaning that at any time the Universe looks the same from any positions and along any directions in space. This Space-Time can be described by Friedmann-Robertson-Walker matrix which can be written as

$$ds^2 = -c^2dt^2 + a^2(t)[d\chi^2 + f_K^2\left(\chi\right)\left(d\theta^2 + sin^2\theta d\phi^2\right)], \tag{2.1.1}$$

$$d\chi = \frac{dR}{\sqrt{(1 - KR^2)}} \tag{2.1.2}$$

where c is the speed of light, and $a(t)$ is called the scale factor, which represents the expansion of the universe. This factor is defined as the change in physical distance between observers at the present and at time t; normally we choose $a(t_0) = a_0 = 1$ at the present epoch, t_0. (χ, θ, ϕ) are the spherical polar coordinates in comoving space where the observer with this coordinate moves along with the Hubble flow. $f_K\left(\chi\right)$ is called the comoving angular diameter distance which depends on χ and the curvature K. It is either a trigonometric, linear or hyperbolic function of χ for a closed Universe ($K > 0$), a flat Universe ($K = 0$) or an open Universe ($K < 0$), respectively.

$$f_K\left(\chi\right) = \begin{cases} \frac{1}{\sqrt{K}}\sin(\sqrt{K}\chi)) & \text{for } K > 0 \\ \chi & \text{for } K = 0 \\ \frac{1}{\sqrt{|K|}}\sinh(\sqrt{|K|}\chi)) & \text{for } K < 0 \end{cases} \tag{2.1.3}$$

2.1.3 Friedmann's Equations

To describe this Space-Time in a more comprehensive way, we need to know its dynamics, namely how the expansion rate changes with time and how the curvature, K, depends on matter. Einstein's field equation links the curvature distortion with the content of the Universe,

$$G_{\alpha\beta} = \frac{8\pi G}{c^2}T_{\alpha\beta} + \Lambda g_{\alpha\beta} \tag{2.1.4}$$

where $G_{\alpha\beta}$ is the metric tensor which describes the geometry of Space-Time, G is the gravitational constant, and $T_{\alpha\beta}$ is the stress-energy tensor of the matter. The second term of Λ proportional to a metric $g_{\alpha\beta}$ was introduced by Einstein to keep our Universe static. The field equation can be simplified to two cosmological field equations:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{Kc^2}{a^2} + \frac{\Lambda}{3} \tag{2.1.5}$$

and

$$\left(\frac{\ddot{a}}{a}\right) = -\frac{4}{3}\pi G\left(\rho + \frac{3p}{c^2}\right) + \frac{\Lambda}{3} \tag{2.1.6}$$

Dots here indicate time derivatives. Equation 2.1.5 is called Friedmann's equation. These two equations govern the expansion of the Universe, describing the time evolution of the scale factor $a(t)$. We introduce the *Hubble parameter*, $H \equiv \dot{a}/a$, as the relative expansion rate. At the present epoch t_0, $H(t_0) = H_0$ is the Hubble constant and its value is still uncertain. Recently, tension over the Hubble constant has increased. Studying the time delays of six distant quasars, Wong et al. (2019) found that $H_0 = 73.3^{+1.7}_{-1.8}, \mathrm{km\,s^{-1}\,Mpc^{-1}}$. Riess et al. (2019) obtained a similar result $H_0 = 74.03 \pm 1.42, \mathrm{km\,s^{-1}\,Mpc^{-1}}$ by measuring cosmic expansion via "cosmic distance ladders". In contrast, the"cosmic microwave background" measurement conducted by Planck Collaboration et al. (2018) predicted a much lower $H_0 = (67.4{\pm}0.5)\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. This inconsistency between "distant" and "local" measurements may be indication of new physics behind ΛCDM. For convenience, we define $h = H_0/100\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ to ignore the uncertainty in H_0. The Friedmann's equation can therefore be expressed in terms of the Hubble parameter,

$$H^2 + \frac{Kc^2}{a^2} = \frac{8\pi G}{3}(\rho) + \frac{\Lambda}{3} \tag{2.1.7}$$

2.1.4 Cosmic Components

During the evolution of Space-Time, our Universe is dominated by different cosmic components. In Eq 2.1.7, we decompose the energy density ρ into its matter content, ρ_m, and its radiation content, ρ_r, and therefore obtain

$$H^2 + \frac{Kc^2}{a^2} = \frac{8\pi G}{3}(\rho_m + \rho_r) + \frac{\Lambda}{3}. \tag{2.1.8}$$

We further define the relationship between pressure, p, and energy density, ρ, of the fluids using an *equation of state*,

$$p = \omega \rho c^2 \tag{2.1.9}$$

where ω is the equation of state parameter. Combing Equation 2.1.5 and Equation 2.1.6, we find $\rho \propto a^{-3(1+\omega)}$. For non-relativistic components, such as baryonic and dark matter, $\omega = 0$ and $\rho_m \propto a^{-3}$. For radiation, such as photons and neutrinos, their equation of state is $\omega = 1/3$, and therefore their density evolves as $\rho_r \propto a^{-4}$. For dark energy (or cosmological constant), $\omega = -1$, corresponding to the case of vacuum. This quantity, ρ_λ, remains a constant and does not change with time.

This time dependence of different density components enables us to draw a big picture of the history of our Universe. In the early Universe, when a was small, radiation dominated but quickly decayed due to the scaling factor a^{-4}. At present, it has a negligible contribution to the total energy density. After that, the Universe entered a matter-dominated era. When matter started to decay, with a scaling factor a^{-3}, dark energy began to dominate and caused an accelerated expansion.

It is useful to define the dimensionless density component at present time as

$$\Omega_m \equiv \frac{\rho_m}{\rho_{cr}}, \;\; \Omega_r \equiv \frac{\rho_r}{\rho_{cr}}, \;\; \Omega_\Lambda \equiv \frac{\Lambda}{3H_0^2}, \tag{2.1.10}$$

where ρ_{cr} is *critical density* of the Universe, defined as

$$\rho_{cr} \equiv \frac{3H_0^2}{8\pi G}. \tag{2.1.11}$$

Using Equation 2.1.8 and Equation 2.1.10, we can express the Hubble parameter in terms of the density parameters and scale factor:

$$H(a)^2 = H_0^2 \, (\Omega_\Lambda - Kc^2 a^{-2} + \Omega_m a^{-3} + \Omega_r a^{-4}). \tag{2.1.12}$$

This equation implies that the time evolution of the Hubble function depends on the fractional density of different components.

At present time, $H(t_0) = H_0$, and $\Omega_r \leqslant \Omega_m$, therefore

$$K = \left(\frac{H_0}{c}\right)^2 (\Omega_m + \Omega_\Lambda - 1). \tag{2.1.13}$$

The curvature K is determined by the energy density of matter and dark energy. If $(\Omega_m + \Omega_\Lambda) = 1$, the Universe is flat. For $(\Omega_m + \Omega_\Lambda) < 1$, the Universe is spatially open (hyperbolic). For $(\Omega_m + \Omega_\Lambda) > 1$, the Universe is spatially closed.

2.1.5 Redshift and Cosmological Distances

In this subsection, we introduce some concepts that will be used in our later discussions.

Redshift

Redshift is a phenomenon where the electromagnetic spectrum of a source is shifted toward longer wavelength due to the Doppler effect, or the expansion of the Universe. In an expanding Universe, a photon wavelength is stretched from $\lambda_{\rm emit}$ to $\lambda_{\rm obs}$ and the redshift, z, is defined as

$$1 + z \equiv \frac{\lambda_{\rm obs}}{\lambda_{\rm emit}}, \tag{2.1.14}$$

which is also related to the scale factor as: $a(t) = \frac{1}{1+z}$. Cosmologists often use redshift to parametrize the measure of the distance (or time) of an object. At the present time t_0, $z = 0$.

Cosmological Distance

Since our Universe could be a curved Space-Time, the definition of "distance" may not be unique, which is different from Euclidean space case. In the following, we define four different distances which are related to different observable properties.

Proper Distance The Proper Distance (also called Physical Distance), $D_{\rm prop}$, is measured by the travel time of a light ray emitted from a source at redshift z_e and received by an observer at redshift z_o. In Equation 2.1.1, Space-Time can be expressed as $ds^2 \equiv -c^2dt^2 + dD_{\rm prop}^2$. Considering the trajectory of light ray, $ds^2 = 0$, we obtain $dD_{\rm prop}^2 = c^2dt^2$.

Comoving Distance As mentioned in Section 2.1.2, the comoving distance $D_{\rm com}$ is measured by the observer moving along with the Hubble flow, which means that this distance is unaffected by the expansion of the Universe. In order to get rid of the space expansion, the proper distance is rescaled by a^{-1}, thus: $dD_{\rm com} = dD_{\rm prop}/a(t)$. For a light path, we have $cdt = a(t)dD_{com}$ from the metric, and the comoving distance can be calculated as

$$D_{com}(z_o, z_e) = \int_{a_(z_e)}^{a(z_o)} \frac{cdt}{a} = c\int_{a(z_e)}^{a(z_o)} \frac{da}{H(a)a^2} \tag{2.1.15}$$

where the Hubble parameter, $H(a)$, depends on the cosmology, and $a(z_e)$ and $a(z_o)$ are the scale factors at emission and observed times, respectively.

Angular-Diameter Distance The *angular-diameter distance* D_A is defined in terms of the object's actual size and its angular size viewed from an observer. Consider the physical cross section of an object ΔA at redshift z_e subtending to the solid angle $\Delta\theta$ for an observer at z_o. Hence,

$$D_A(z_o, z_e) \equiv \left(\frac{\Delta A}{\Delta\theta}\right)^{1/2} = a(z_e)f_K(\chi(z_o, z_e)) = a(z_e)f_K(D_{com}(z_o, z_e)) \tag{2.1.16}$$

where $f_K(\chi(z_o, z_e))$ is the comoving angular diameter defined in Equation 2.1.3. The scale factor, $a(z_e)$, in this equation is meant to rescale the comoving quantity back to the physical quantity at emission time.

Luminosity Distance The *luminosity distance* D_L is defined by the relation between an object's luminosity L at redshift z_e, and the flux received by an observer at redshift z_o. The observed flux can be written as

$$f = \frac{L}{4\pi D_A^2(z_o, z_e)}\left(\frac{a(z_e)}{a(z_o)}\right)^4 \tag{2.1.17}$$

The factor of $(a(z_e)/a(z_o))^2$ arises from the fact that photons are redshifted, their energy is thus decreased by a factor $a(z_e)/a(z_o)$, and their arrival time is also delayed by the same factor. We can then define the *luminosity distance* D_L as

$$D_L(z_o, z_e) = \left(\frac{a(z_o)}{a(z_e)}\right)^2 D_A(z_o, z_e), \tag{2.1.18}$$

and obtain

$$f = \frac{L}{4\pi D_L^2(z_o, z_e)}. \tag{2.1.19}$$

2.2 Structure formation

2.2.1 Density Fluctuations

In the previous section, we mentioned that the standard cosmological model is built upon the Cosmological Principle: *the Universe is homogeneous and isotropic*, which is valid on large scales. In fact, on smaller scales, there were inhomogeneities, i.e. small fluctuations at very early times. These density fluctuations originated from quantum fluctuations at the time of the Big Bang, and then by gravitational instability growing to form structures, such as clusters, galaxies, and stars we see today. These tiny density fluctuations can be described by the density contrast

$$\delta(\mathbf{x}) = \frac{\rho(\mathbf{x}) - \bar{\rho}}{\bar{\rho}}, \tag{2.2.20}$$

where $\rho(\mathbf{x})$ is the local density, and $\bar{\rho}$ is the mean density in the Universe. The amplitudes of these fluctuations will gradually grow due to their own gravity. When density contrasts, $\delta(\mathbf{x})$, are much smaller than unity at early time, they evolve linearly with time, and therefore can be described by *linear perturbation theory.* The evolution of density fluctuations in the early stage is given by (e.g. Peacock, 1999; Peebles, 1993)

$$\ddot{\delta} + 2\frac{\dot{a}}{a}\dot{\delta} = 4\pi G\rho\delta. \tag{2.2.21}$$

This shows that the time evolution of density fluctuations depends on the underlying cosmology, namely the background expansion and the fraction of different energy density components, both of which also change with cosmic time. It is convenient to decompose the density contrast, δ, into discrete Fourier modes with different wavelengths, λ, so that these individual modes evolve independently.

Before getting into the details of structure evolutionary history, we introduce a concept of "*horizon*". The horizon size is the maximum distance that a photon can travel in a time, t, since the Big Bang, $D_H = \frac{c}{aH(a)}$. As no information (or any physical interaction) can travel faster than a photon, particles are not able to communicate outside the horizon. However, the Universe is discovered to be very homogeneous on large scales (Smoot et al., 1992) where the widely separated patches of sky are causally disconnected regions. These patterns of structure similarity are known as the *horizon problem*, which can be explained by *cosmological inflation*.

After the big bang, our Universe experienced a period of accelerated expansion. This rapid expansion enlarged the connected regions out of the horizon regions, resulting in the homogeneous nature of our Universe. Since then, matter density starts to evolve. If density fluctuations of wavelength $\lambda < D_H(a)$ in the radiation domination era, the growth rate of these modes is $\propto a^2$. When the fluctuations of wavelength are smaller than the horizon during the radiation-dominated era, which means that these modes enter the horizon, the gravitational collapse of these density contrast is suppressed by the outward pressure of radiation. After entering the matter-dominated era, they continue to grow as $\propto a$. Finally, density fluctuations grow to the point where they cannot be linearly described ($\delta \sim 1$), with different modes interacting with each other. It is therefore difficult to describe this complex non-linear evolution, numerical calculation is thus needed.

2.2.2 The Comic Web

With the initial conditions of the Universe, large-scale N-body simulations (e.g. Springel et al., 2005) trace the movement of dark matter particles, helping us to model how dark matter is distributed and how structures have grown from early times to the present. Cosmologists find that our Universe evolves into an ensemble of large-scale structures (LSS) by self-gravitation. These highly non-linear structures form a complicated network of matter, called *"cosmic web"*. The network contains the under-dense regions called voids (Cautun et al., 2013), and the over-density regions centering in galaxy clusters and interconnected through filaments and sheets. Over-dense perturbations continue to grow by attracting nearby dark matter, and voids also grow in size with time to become emptier. These extended filamentary structures play an important role in driving the growth of structures since materials infall into dark matter halos along these preferred directions. Figure 2.1 shows a snapshot from the Millennium Simulation (Springel et al., 2005) where there is a massive galaxy cluster in the center with extended large-scale filaments connected to it.

When we zoom into the smaller scales of the dark matter halo, it is found that

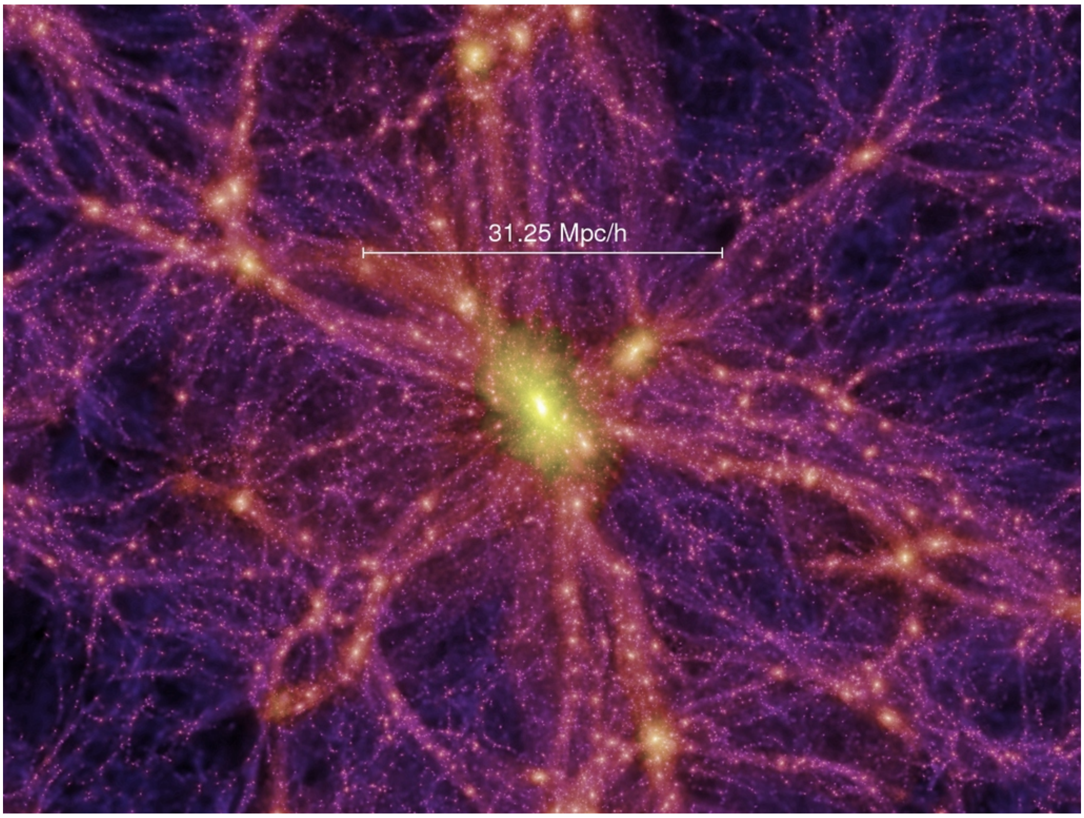


Figure 2.1: The snapshot of a simulated Universe in the N-body Millennium Simulations (Springel et al., 2005) at redshift $z = 0.0$. The white bar marks the length scale of 31.25 Mpc/h. The colour scale represents the mass density of dark matter, where the massive cluster halo in the center is shown in bright yellow, and the surrounding extended filaments are shown in purple. (Figure credit: Millennium Simulation Project [1], Springel et al., 2005)

---

[1] `https://wwwmpa.mpa-garching.mpg.de/galform/virgo/millennium/`

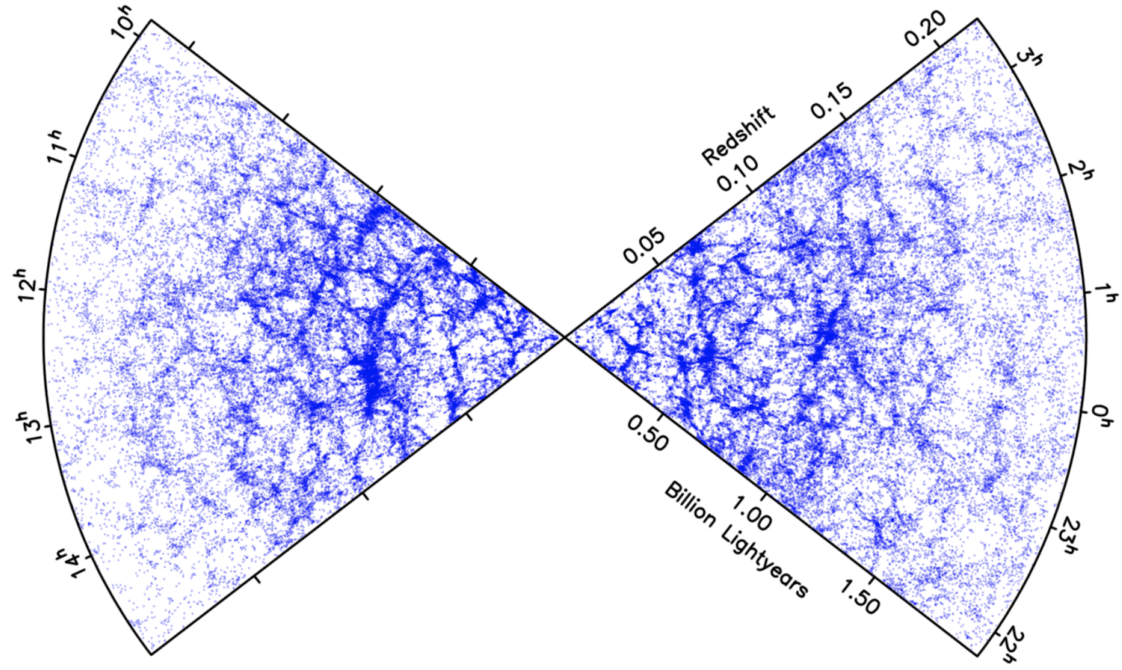


Figure 2.2: Galaxy distribution in the 2dF Galaxy Redshift Survey. Redshift is shown in the radial direction, and the polar angle is the right ascension. (Figure credit: M. Colless and the 2dF Galaxy Redshift Survey team).

smaller local peaks collapse and form subhalos, where infalling matter becomes hot enough for star formation to begin, eventually leading to galaxies. Therefore, galaxies are believed to exist at the centres of dark matter halos. Even though the majority of matter is invisible, the luminous part, the galaxies, can approximately trace the underlying mass distribution. Observational evidences of this web-like network were first provided by large-scale galaxy redshift surveys (e.g. Colless et al., 2001; Geller & Huchra, 1989; York et al., 2000); (figure 2.2) which map the spatial distributions of galaxies in redshift space with spectroscopic surveys. However, the spatial distribution of baryons may differ from that of underlying dark matter due to the varied physics of galaxy formation. (e.g. Norberg et al., 2002; Saunders et al., 1992). *Galaxy bias*, the ratio of mean overdensity of galaxies to mean overdensity of mass, $b = \delta_g/\delta$, provides a description of how well galaxies trace the total mass distribution (e.g. Kaiser, 1984).

Gravitational lensing provides a direct way to study the dark matter distribution, which contains rich information on structure formation history, and therefore can

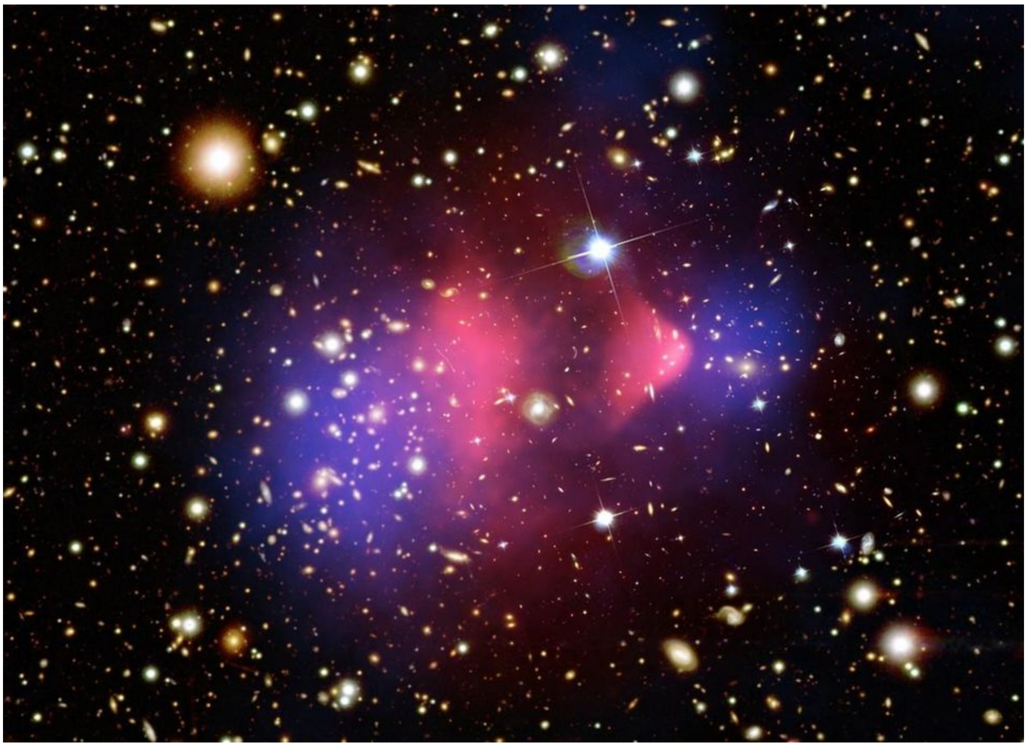


Figure 2.3: The “bullet cluster”, 1E0657-56, reconstructed from multi-wavelength data. The optical image shows the location of galaxies. The overlaid pink cloud shows the X-ray emission from the hot ICM. Both of these are associated with baryonic material. The overlaid blue color shows the total mass reconstruction obtained with gravitational lensing. (Figure credit: X-ray: NASA/CXC/CfA/ M.Markevitch et al.; Lensing Map: NASA/STScI; ESO WFI; Magellan/U.Arizona/ D.Clowe et al. Optical image: NASA/STScI; Magellan/U.Arizona/D.Clowe et al.)

be used to test different cosmological models. In particular, based on the ΛCDM model, LSS evolves through an upward hierarchical process of accretion and merging of small halos, eventually building the largest structures, clusters of galaxies. This is consistent with observations. On the other hand, hot dark matter model predicts the reverse process, in which the largest structures form first and then develop smaller systems. This model has been ruled out by observations.

2.2.3 Galaxy Clusters

According to hierarchical structure formation, galaxy clusters are the largest gravitationally bound objects, which formed via a series of mergers with smaller halos and continuous accretion of surrounding matter. They can contain several tens and up to thousand of galaxies, with total masses up to a few $10^{15}\mathrm{M}_{\odot}$. For example, the Virgo cluster with $M_{200} = (1.05 \pm 0.02) \times 10^{14}\mathrm{M}_{\odot}$ (Simionescu et al., 2017) has approximately 2,000 member galaxies and a size of ~ 3 Mpc in diameter. The Coma cluster with $M_{200} = (1.88^{+0.65}_{-0.56}) \times 10^{15}\mathrm{M}_{\odot}$ (Kubo et al., 2007) has more than 10,000 member galaxies and spans a physical scale of 6-8 Mpc. There are smaller objects, called *Galaxy Groups*, which contain fewer galaxies and have typical masses of about a few $10^{13}\mathrm{M}_{\odot}$. The composition of galaxy clusters is roughly 85% of dark matter, 10% of X-ray luminous and hot intracluster gas, and 5% of stars. This vast amount of dark matter makes galaxy clusters one of the most important cosmological probes in modern Astrophysics. They are the largest observable objects in the Universe, representing the high-mass end of collapsed structures originating from the highest peaks in the underlying density field. Their individual physical properties, mass functions and evolution histories are thus dependent on the initial density fluctuations and can be used to test the underlying cosmology.

Accurate measurement of their mass is one of the most important topics in galaxy cluster study. Traditionally we have three approaches to determine their mass. The first of them is the study of the dynamics of cluster members, using the virial theorem,

$$2\,T + U = 0, \tag{2.2.22}$$

where

$$\begin{aligned} T &= \frac{M_{\mathrm{tot}}\sigma_v^2}{2}, \\ U &= \frac{G\,M_{\mathrm{tot}}^2}{R_{\mathrm{tot}}}, \end{aligned} \tag{2.2.23}$$

and σ_v is velocity dispersion. It links the kinetic energy (T) of the total system with its gravitational potential energy (U) to yield the cluster mass. In order to use the virial theorem, we need to assume that the system is stable/in equilibrium. Second, since the hot gas in clusters emits the X-rays due to the *Bremsstrahlung* effect, the cluster mass can be estimated by using the X-ray temperature and flux, given the assumption of hydrostatic equilibrium (HSE) and the spherical symmetry of the gravitational potential. These two methods are based on assumptions regarding the dynamical state of the system. However, for an unrelaxed or merging cluster, the virial theorem and the HSE are not valid. Gravitational lensing though, enables us to directly probe the total cluster mass, including its dark matter content, independently of any equilibrium or symmetry assumptions. The theory of gravitational lensing is presented in detail in the next chapter.

2.2.4 Press-Schechter Mass Function

According to hierarchical structure formation, smaller objects formed first and then merged into bigger systems. Mergers between halos of similar masses are referred to major mergers. Those between halos at different mass ranges are called minor mergers. Supported by observations and N-body simulations, a parent halo contains smaller halos (called *subhalos*) orbiting around it, and the amount of subhalos depends on its merging history.

Since structure formation is a dynamic process, current observation cannot capture its whole evolution history. Nevertheless, its product, such as galaxy clusters and subhalos at a given redshift, can be seen. Therefore, the statistical distribution of halos as a function of mass is sensitive to cosmology, where the overall density controls the rate of structure growth. We introduce the concept of the *mass function*, $n(M)$, the number density of halos at a given redshift within a mass in the range M and $M + dM$. In principle, this quantity can be measured by counting structures of a given mass contained within a selected volume in space. Press & Schechter (1974) derived an analytical expression for the mass function and related it to the initial density field,

$$n(M) = \sqrt{\frac{2}{\pi}} \frac{d\sigma_M}{dM} \frac{\rho_0 \delta_c}{M\sigma_M^2} \exp\left(-\frac{\delta_\mathrm{c}^2}{2\sigma_\mathrm{M}^2}\right), \tag{2.2.24}$$

where ρ_0 is the background density, δ_c is the threshold of smoothed initial density field and it is common to use $\delta_c \approx 1.686$ as this is the linear density contrast at which a spherical top-hat perturbation would collapse and form virialized structures (Spherical Collapse Model; Gunn & Gott, 1972). σ_M is the variance of the smoothed density field,

$$\sigma_M^2 = \frac{1}{2\pi^2} \int dk k^2 P_m(k, z) \hat{W}_R^2(k), \tag{2.2.25}$$

where $P_m(k, z)$ is the matter power spectrum, and $\hat{W}_R^2(k)$ is the Fourier transform of the real-space spherical tophat window function which can filter the density field on a length scale R,

$$\hat{W}_R(k) = \frac{3}{(kR)^3}\left[\sin(kR) - (kR)\cos(kR)\right]. \tag{2.2.26}$$

Thus, the mass function depends on the primordial power spectrum (or the two-point correlation function of the initial density contrast), extrapolated to the present using linear theory.

2.2.5 Halo Density Profile

To investigate the dependence of the mass distribution with radial distance from the cluster centre, it is usual to measure the radial density profile of galaxy clusters. The density profile is often described by a parametric model with radial symmetry (e.g. elliptical model). Thus, for irregular mass distributions such as most unrelaxed clusters, the stacked density profile of a sample of clusters is needed to compare theoretical predictions with the real observed Universe. Moreover, there are some azimuthally averaged features of galaxy clusters which can be measured from radial density profiles. The splashback radius (Diemer & Kravtsov, 2014) is one of them, i.e. the physical boundary of dark matter halos that characterises the region of accreting materials at the outskirts.

There are several different halo models describing the radial mass dependence of a real galaxy cluster. In the following, we present the details of the mathematical definition and description of each halo model used in this thesis. Some of these models only have an analytical form for the 3D density; from this we can compute the projected surface density by integrating along the line of sight,

$$\Sigma(R) = 2\int_R^\infty \frac{r\rho(r)}{\sqrt{r^2 - R^2}}dr. \tag{2.2.27}$$

tPIEMD profile

Massive elliptical galaxies are empirically observed to have an approximately isothermal density distribution ($\rho \propto r^{-2}$), and a total mass proportional to the velocity distribution of their stars, σ. The truncated Pseudo-Isothermal Elliptical Mass Distribution (tPIEMD; Elíasdóttir et al., 2007; Kassiola & Kovner, 1993a; Limousin et al., 2005)

$$\rho_{\mathrm{tPIEMD}} = \frac{\rho_0}{(1 + r^2/r_{\mathrm{c}}^2)(1 + r^2/r_{\mathrm{t}}^2)} \tag{2.2.28}$$

removes the inconvenient mathematical singularity from the centre by softening the density inside a core radius r_{c}, and makes the integrated mass finite by truncating the profile at radius r_{t}. The central density is

$$\rho_0 = \frac{\sigma^2}{2\pi G} \frac{r_{\mathrm{c}} + r_{\mathrm{t}}}{r_{\mathrm{c}}^2 \, r_{\mathrm{t}}} \,. \tag{2.2.29}$$

The projected two-dimensional mass distribution is

$$\Sigma_{\mathrm{tPIEMD}}(R) = \frac{\sigma^2}{2G} \frac{r_{\mathrm{t}}}{r_{\mathrm{t}} - r_{\mathrm{c}}} \left(\frac{1}{\sqrt{R^2 + r_{\mathrm{c}}^2}} - \frac{1}{\sqrt{R^2 + r_{\mathrm{t}}^2}} \right) \,. \tag{2.2.30}$$

Within r_{c}, the surface density is approximately constant. For the region between r_{c} and r_{t}, its surface density is isothermal, whilst beyond r_{t} it sharply decreases as R^{-3}. It is a physically motivated model since it provides a finite total mass and central density to describe the flat region in the cluster centre.

NFW profile

The Navarro-Frenk-White (NFW) model is the universal density profile of dark matter halos suggested by numerical simulations (Navarro et al., 1996, 1997). The NFW 3D radial density profile has a two-parameter functional form of

$$\rho_{\mathrm{NFW}} = \frac{\rho_{\mathrm{s}}}{(r/r_{\mathrm{s}})(1 + (r/r_{\mathrm{s}}))^2} \tag{2.2.31}$$

where ρ_{s} and r_{s} are the characteristic density and radius respectively. At $r = r_s$, the logarithmic density slope equals to the isothermal value, $d\ln\rho(r)/d\ln r = 2$. For any given cosmology and cluster redshift, this model can also be parametrized in terms of the concentration parameter, $c_{200} = R_{200}/r_s$, where R_{200} is the radius at which the mean enclosed density is equal to 200 times the critical density, ρ_c, of the Universe, and halo mass $M_{200} = (4\pi/3)200\rho_c R_{200}^3$.

The projected two-dimensional mass distribution is given by (Bartelmann, 1996)

$$\Sigma_{\mathrm{NFW}}(R) = 2\Sigma_{\mathrm{s}} F(x), \tag{2.2.32}$$

where $\Sigma_\mathrm{s} = \rho_\mathrm{s} r_\mathrm{s}$, $x = R/r_\mathrm{s}$, and

$$F(x) = \begin{cases} \frac{1}{x^2-1}\left(1 - \frac{2}{\sqrt{x^2-1}}\arctan\sqrt{\frac{x-1}{x+1}}\right) & \text{if } x > 1 \\ \frac{1}{3} & \text{if } x = 1 \\ \frac{1}{x^2-1}\left(1 - \frac{2}{\sqrt{1-x^2}}\arctan\sqrt{\frac{1-x}{1+x}}\right) & \text{if } x < 1\ . \end{cases} \tag{2.2.33}$$

Generalized NFW (gNFW) profile

A generalized version of the NFW model has the form (Zhao, 1996)

$$\rho_\mathrm{gNFW} = \frac{\rho_s}{(r/r_s)^\alpha (1 + (r/r_s))^{(3-\alpha)}}\ . \tag{2.2.34}$$

It has a power-law shaped central cusp, $\rho \propto r^{-\alpha}$, which reduces to a NFW model when $\alpha = 1$. Similar to NFW, we describe the gNFW profile with a central slope α, a halo mass, M_{200}, and a concentration, $c_{200} = r_{200}/(2-\alpha)\, r_s$. The radial dependence of the gNFW lensing signal was calculated by Keeton (2001a).

Einasto profile

Several N-body simulations (e.g. Dutton & Macciò, 2014; Graham et al., 2006; Harvey et al., 2015; Klypin et al., 2016) have shown that CDM halos can best be described by the Einasto density profile (Einasto, 1965), which is written as

$$\rho_\mathrm{Einasto} = \rho_s \exp\left\{-\frac{2}{\alpha_E}\left[\left(\frac{r}{r_s}\right)^{\alpha_E} - 1\right]\right\}, \tag{2.2.35}$$

where α_E is the shape parameter describing the steepness of the logarithmic slope. Alternatively, we can express the density profile as

$$\frac{d\log\rho}{d\log r} = -2\left(\frac{r}{r_s}\right)^{\alpha_E}\ . \tag{2.2.36}$$

An Einasto profile with $\alpha_E \sim$0.18 has a similar shape as an NFW at a given concentration (Ludlow et al., 2013).

Burkert profile

In contrast to the other profiles presented, the Burkert (1995) dark matter halo has a core structure in the inner region, with a density profile described as

$$\rho_\mathrm{Burkert} = \frac{\rho_\mathrm{core}}{(1 + r/r_\mathrm{core})(1 + r^2/r^2_\mathrm{core})}, \tag{2.2.37}$$

where ρ_{core} and r_{core} parameterise the density and size of a (constant density) core.

Diemer & Kravtsov profile

The density profile proposed in Diemer & Kravtsov (2014) (hereafter referred to as DK14) is a more flexible function, which was calibrated with regard to a suite of ΛCDM simulations. This model is described by two components: (1) the collapsed matter which is modelled by a truncated Einasto profile (Einasto, 1965), and (2) the infalling material which is modelled by a power law function. The complete model is given by

$$\rho(r) = \rho_{\text{Einasto}}(r) \times f_{\text{trans}}(r) + \rho_{\text{infall}}(r) \tag{2.2.38}$$

$$f_{\text{trans}}(r) = \left[1 + \left(\frac{r}{r_t}\right)^{\beta}\right]^{-\gamma/\beta} \tag{2.2.39}$$

$$\rho_{\text{infall}}(r) = \frac{\rho_m b_e}{\Delta_{\text{max}}^{-1} + \left(\frac{r}{5r_{200m}}\right)^{s_e}} \tag{2.2.40}$$

where $\Delta_{\text{max}} = 10^3$, and the transition term, f_{trans}, captures the steepening of the profile around a truncation radius, r_t. The shape parameters, γ and β, define the steepness of the profile and how quickly the slope changes, respectively. For the infalling material, the power law profile which decreases with radius ($s_e > 0$) approaches the mean density of the Universe, ρ_m, at large radii. $\Delta_{\text{max}} = 10^3$ is introduced to avoid the spurious contribution toward the center of the cluster. We adapted the publicly available code COLOSSUS (Diemer, 2018) for the calculation of the DK14 density profile.

Chapter 3

Gravitational Lensing Formalism and Shear Measurement

3.1 Introduction

The phenomenon of gravitational lensing was initially predicted by General Relativity (Einstein, 1915). It states that a massive object locally distorts Space-Time. Photons traveling from a source to an observer along geodesics in a curved 3D space cause interesting features. The observed images of the background source can be distorted and magnified. Moreover, a compact and very massive object can sufficiently bend the light rays emitted from a single source such that multiple light rays eventually converge to the observer. This results in multiple images of the source at different angular positions. Since these geometrical effects are purely due to the over-density of mass, it is free of assumptions on the physical state of the lens (the massive object placed between the source(s) and the observer), which makes gravitational lensing an extremely powerful probe of underlying mass distribution in modern Astronomy and Cosmology.

Gravitational lensing was first confirmed in 1919 by Sir Frank Watson Dyson and Sir Arthur Eddington (Dyson et al., 1920). They measured the offset of positions of stars in the Hyades cluster when these stars passed along the line of sight to the Sun during the total solar eclipse in May 1919. The offset measured was consistent with the predictions from General Relativity, twice larger than the deflection predicted by

Newton's gravitation theory. Since then, the theory of General Relativity has been accepted and became the standard theory of gravity. In 1979, the first multiple-image system was observed (Walsh et al., 1979). Two images of the distant quasar Q0957+561 at redshift z = 1.4. These discoveries paved the way for further studies and observations of gravitational lensing. Nowadays, gravitational lensing is used in various areas, such as high-redshift galaxy studies, the detection of exoplanets, the investigation of dark matter's nature, and the constraint of Cosmology.

Depending on the configuration of the background sources, the lens object, the observer, and the mass of the lens, gravitational lensing can be classified into three regimes: strong lensing, weak lensing, and microlensing. Strong gravitational lensing refers to the lensing of a background galaxy which is strong enough to produce a noticeable distortion such as giant arcs, Einstein rings, or multiple-image systems. For weakly distorted images which correspond to weak gravitational lensing, the distortion in shape of each individual image is too weak to be directly observed. However, using statistical measurements, the net distortion averaged over a sample of background images can be calculated. Therefore, measurements of weak gravitational lensing are statistical in nature. The signal-to-noise ratio (S/N) of these measurements is limited by the density of galaxies *behind* the lens. To enhance statistical significance, in the near future, a large dataset will be available from *Euclid* (Laureijs et al., 2011), the Large Synoptic Survey Telescope *(LSST)* [1] and the Nancy Grace Roman Space Telescope (Spergel et al., 2013). Astrophysics and cosmology are entering a new era!

The last regime is microlensing, for which distortions are too small to be detectable, but changes of apparent magnitude can be observed. When a low mass lens object passes between a bright background source and an observer, the apparent brightness of the source is boosted and then diminished due to the changing configuration of the lens system. Through monitoring the source for a certain period of time, this transient effect can be characterised by the light curve. Microlensing has been used extensively for exoplanets searches/detections.

---

[1] `https://www.lsst.org/lsst/`

In this chapter, I present the gravitational lensing formalism which will be used in the following chapters. The presentation will start with the derivation of the basic lensing equations. Then I will introduce the lensing observables in Section 3.2. The procedure of weak lensing shear measurements of this study will be presented in Section 3.3. Finally, Section 3.4 describes the weak lensing mass mapping methods that are used in this work. For more in-depth discussion on gravitational lensing, I refer the reader to reviews by Bartelmann & Schneider (2001), Massey et al. (2010) and Kneib & Natarajan (2011b).

3.2 Gravitational Lensing Theory

3.2.1 Lens Equation

First of all, we introduce the assumption used in most of the lensing studies, the *thin lens approximation*: all deflection of the light rays occur in the lens plane. Neglecting the contributions from the cosmic large-scale structures, this assumption is valid for most of the lens systems since the scale of the lens object is usually much smaller than the angular diameter distances involved. As shown in figure 3.1, we consider a typical gravitational lensing system. A light ray propagates from the source with redshift z_S at the position $\boldsymbol{\eta}$ on the source plane, to the observer at redshift z_L, passing the position $\boldsymbol{\xi}$ on the lens plane where it is deflected by an angle $\widehat{\boldsymbol{\alpha}}$. The angle between the optical axis and the source position is $\boldsymbol{\beta}$, and the angle between the optical axis and the image is $\boldsymbol{\theta}$. The angular diameter distances between the source and lens, between the lens and observer, and between the source and observer, are D_{LS}, D_L and D_S, respectively. We can easily find the geometrical relation in this system:

$$\boldsymbol{\eta} = (D_S/D_L)\boldsymbol{\xi} - D_{LS}\widehat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) \tag{3.2.1}$$

Introducing angular coordinates $\boldsymbol{\beta} = \boldsymbol{\eta}/D_s$ and $\boldsymbol{\theta} = \boldsymbol{\xi}/D_L$, we can transform the geometric relation into

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \frac{D_{LS}}{D_S}\widehat{\boldsymbol{\alpha}}(D_d\boldsymbol{\theta}) \equiv \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta}), \tag{3.2.2}$$

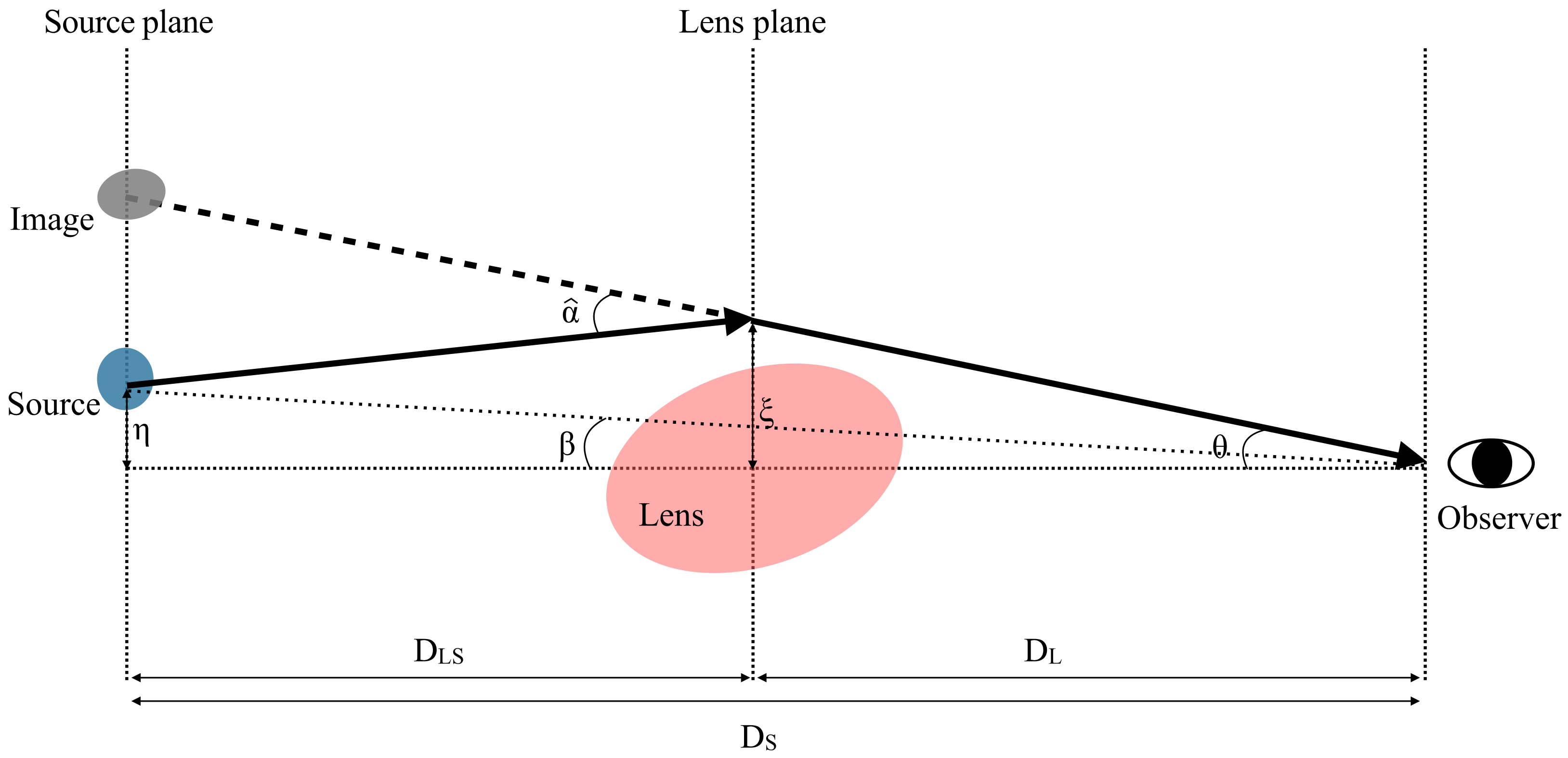


Figure 3.1: Sketch of a typical gravitational lens system: A light ray propagates from the source at the position $\boldsymbol{\eta}$ on the source plane to the observer, passing the position $\boldsymbol{\xi}$ on the lens plane where it is deflected by an angle $\hat{\boldsymbol{\alpha}}$. The angle between the optical axis and the source position is $\boldsymbol{\beta}$, and the angle between the optical axis and the image is $\boldsymbol{\theta}$. The angular diameter distances between the source and lens, between the lens and observer, and between the source and observer are D_{LS}, D_L and D_S, respectively.

where we define the reduced deflection angle $\boldsymbol{\alpha}(\boldsymbol{\theta}) \equiv \frac{D_{LS}}{D_S}\widehat{\boldsymbol{\alpha}}$. Equation 3.2.2 is called the *lens equation*. In general, the lens equation is non-linear with respect to $\boldsymbol{\theta}$. It may have more than one solution for a fixed $\boldsymbol{\beta}$. In this case, a source at $\boldsymbol{\beta}$ has images at several positions on the sky, producing multiple images of a single source. Consider a lens with point mass M, General Relativity predicts the deflection angle

$$\widehat{\boldsymbol{\alpha}} = \frac{4GM}{c^2\boldsymbol{\xi}}. \tag{3.2.3}$$

For a source aligned with the lens center, $\boldsymbol{\beta}$=0, one defines the Einstein radius:

$$\boldsymbol{\theta_E} = \sqrt{\frac{4GM}{c^2}\frac{D_{LS}}{D_S D_L}}. \tag{3.2.4}$$

which is directly related to the lens mass. Hence, the Einstein radius is usually used to describe the strength of a lens.

For an extended lens, its mass distribution can be seen as an ensemble of point deflectors. Furthermore, since the cosmological distance to a typical background galaxy is vastly larger than the size of an intervening mass, the 3D distribution of that mass can be approximated as a projected surface density $\Sigma(\boldsymbol{\xi})$,

$$\Sigma(\boldsymbol{\xi}) \equiv \int \rho(\boldsymbol{\xi}, z)dz, \tag{3.2.5}$$

where ρ is the 3D mass density (Equation 2.2.20), and z is the distance along the line of sight. Then, using $\boldsymbol{\xi} = D_L\boldsymbol{\theta}$, the deflection angle can be written as

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{4G}{c^2}\frac{D_{LS}D_L}{D_S}\int \Sigma(\boldsymbol{\theta'})\frac{\boldsymbol{\theta} - \boldsymbol{\theta'}}{|\boldsymbol{\theta} - \boldsymbol{\theta'}|^2}d^2\boldsymbol{\theta'}. \tag{3.2.6}$$

Convergence

Here, we introduce a useful quantity in lensing, the dimensionless critical surface density, commonly called *convergence*,

$$\kappa = \frac{\Sigma(D_L\boldsymbol{\theta})}{\Sigma_{\rm crit}}, \tag{3.2.7}$$

where $\Sigma_{\rm crit}$ is the critical surface mass density which depends on the redshifts (z_L, z_s) and the geometry of the Universe $(H_0, \Omega_m, \Omega_\lambda)$:

$$\Sigma_{\rm crit} \equiv \frac{c^2}{4\pi G}\frac{D_S}{D_L D_{LS}} = \frac{c^2}{4\pi G D_L}\beta^{-1}(z_L, z_s). \tag{3.2.8}$$

The lensing sensitivity function $\beta(z_L, z_S) = D_{LS}/D_S$ describes the lensing strength as a function of lens and source redshifts (z_L, z_S). For a foreground galaxy with $z_S < z_L$, $\beta(z_L, z_S) = 0$. Convergence can be used to determine the lensing regime. When $\kappa > 1$, multiple images are produced and we refer to it as strong lensing. For weak lensing, $\kappa < 1$ and galaxies are weakly distorted.

Lensing potential

Combinning equations 3.2.6 and 3.2.7, the deflection angle can be written as

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{1}{\pi}\int \kappa(\boldsymbol{\theta'})\frac{\boldsymbol{\theta}-\boldsymbol{\theta'}}{|\boldsymbol{\theta}-\boldsymbol{\theta'}|^2}d^2\boldsymbol{\theta'}. \tag{3.2.9}$$

We define the two-dimensional projected gravitational potential, called deflection potential

$$\psi(\boldsymbol{\theta}) = \frac{1}{\pi}\int \kappa(\boldsymbol{\theta'})\mathrm{ln}|\boldsymbol{\theta}-\boldsymbol{\theta'}|\mathrm{d}^2\boldsymbol{\theta'}, \tag{3.2.10}$$

which satisfies the Poisson equation

$$\nabla^2\psi = 2\kappa. \tag{3.2.11}$$

The deflection angle can therefore be written as a gradient of the deflection potential: $\boldsymbol{\alpha} = \boldsymbol{\nabla}_\theta\psi$, and the lens equation can be expressed in terms of $\psi(\boldsymbol{\theta})$ as

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \boldsymbol{\nabla}_\theta\psi(\boldsymbol{\theta}). \tag{3.2.12}$$

3.2.2 Weak Lensing Distortion

The lens equation gives the relation of source plane points $\boldsymbol{\beta}$, and image plane points θ. The local properties of the lens equation, the mapping between small elements in the source plane, $d\boldsymbol{\beta}$, and those in the image plane, $d\boldsymbol{\theta}$, are described by the lensing Jacobian matrix.

$$A(\boldsymbol{\theta}) = \left(\frac{\partial\boldsymbol{\beta}}{\partial\boldsymbol{\theta}}\right) = \left(\delta_{ij} - \frac{\partial^2\psi(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\right) = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix} \tag{3.2.13}$$

where i and j denotes the axes of the angular coordinate on the sky plane $\boldsymbol{\theta} = (\theta_1, \theta_2)$.

Lensing Covergence and Shear

In the Jacobian matrix, the quantities of shape distortion are introduced, γ_1 and γ_2. They are the components of the complex shear field, $\gamma = \gamma_1 + i\gamma_2 = |\gamma|e^{i2\varphi}$, where φ is the phase angle. The factor of 2φ shows that the shear is a spin-2 vector which maps to itself under a rotation of 180°. They can also be written as functions of the deflection potential,

$$\gamma_1 = \frac{1}{2}(\psi_{,11} - \psi_{,22}), \quad \gamma_2 = \frac{1}{2}(\psi_{,12} + \psi_{,21}) = \gamma_{,12}, \tag{3.2.14}$$

where $\psi_{,ij} = \partial^2\psi/\partial\theta_i\partial\theta_j$ (i, j = 1, 2). κ is related to ψ through Poisson's equation. The Jacobian matrix can be decomposed into two terms:

$$A(\boldsymbol{\theta}) = \begin{pmatrix} 1-\kappa & 0 \\ 0 & 1-\kappa \end{pmatrix} + \begin{pmatrix} -\gamma_1 & -\gamma_2 \\ -\gamma_2 & \gamma_1 \end{pmatrix}, \tag{3.2.15}$$

where the first term describes isotropic size magnification caused by the convergence κ, and the second term is a trace-free matrix dependent on the shear. This term corresponds to the tidal gravitational field contributed by the shape distortion, deforms the circular source to an ellipse, and rotates based on the phase angle φ. Figure 3.2 gives examples of the deformation of a circular source under different types of distortions. Since the tidal gravitational field induces the coherent tangential pattern of the shear, we decompose the shear into two terms, the tangential shear γ_+, and the cross shear $\gamma_\times$,

$$\begin{aligned} \gamma_1 &= -\gamma_+ \cos(2\phi) + \gamma_\times \sin(2\phi) \\ \gamma_2 &= -\gamma_+ \sin(2\phi) - \gamma_\times \cos(2\phi) \end{aligned} \tag{3.2.16}$$

where ϕ is the angle of the position vector of the source pointed from the lens. The tangential component, γ_+, measures the tangential coherence of the shape distortions due to weak lensing. The cross term, $\gamma_\times$, corresponds to 45 degree rotated distortion patterns of background images which is not the contribution from weak lensing. Therefore, $\gamma_\times$ is usually used to check systematics in weak lensing analyses.

Finally we introduce a quantity, the *reduced shear* $\boldsymbol{g}$, defined as

$$\boldsymbol{g}(\boldsymbol{\theta}) \equiv \frac{\boldsymbol{\gamma}(\boldsymbol{\theta})}{1 - \kappa(\boldsymbol{\theta})}. \tag{3.2.17}$$

As we shall see in Section 3.3.1, in general, the observable quantity for weak shear lensing is not the shear $\boldsymbol{\gamma}$, but the reduced shear, $\boldsymbol{g}$.

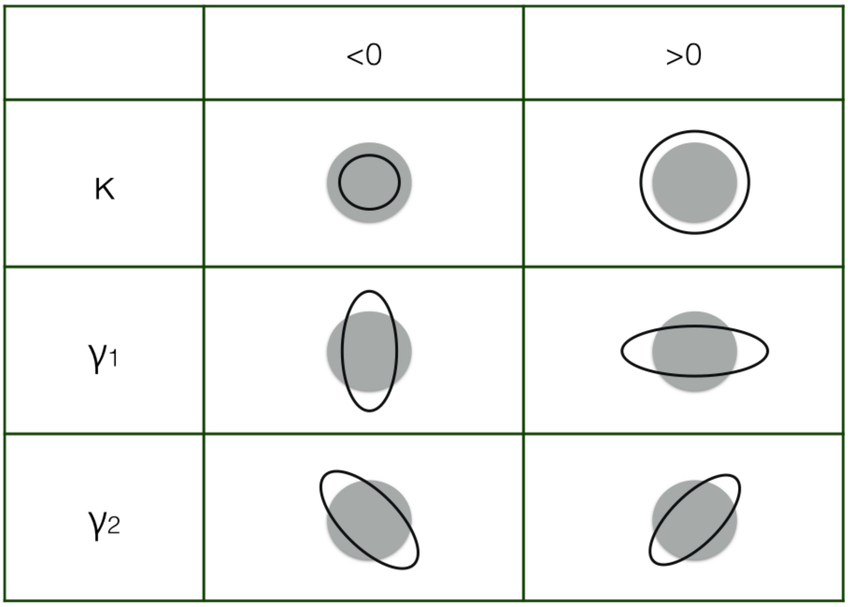


Figure 3.2: Illustration of different types of lensing distortion. The black circle represents a circular source whose shape and size deform into the grey shapes under lensing effect.

Magnification

Since the gravitational light deflection does not involve emission or absorption of photons, the surface brightness density is conserved with gravitational lensing. Hence the flux magnification is caused by the area distortion: $\delta\Omega^I = \mu\delta\Omega^S$ only. We can define the magnification as the ratio between the lensed and the unlensed image solid angles:

$$\mu = \frac{1}{detA} = \frac{1}{(1-\kappa)^2 - |\gamma|^2} \tag{3.2.18}$$

3.2.3 Strong Lensing

Critical and Caustic Lines

The magnification, μ, is a function of covergence, κ, and shear, γ, which can also be expressed as $\mu^{-1} = (1-\kappa)^2 - \gamma^2$. The infinite magnification will happen when $\mu^{-1} = 0$. The corresponding locus in the image plane of infinite magnification

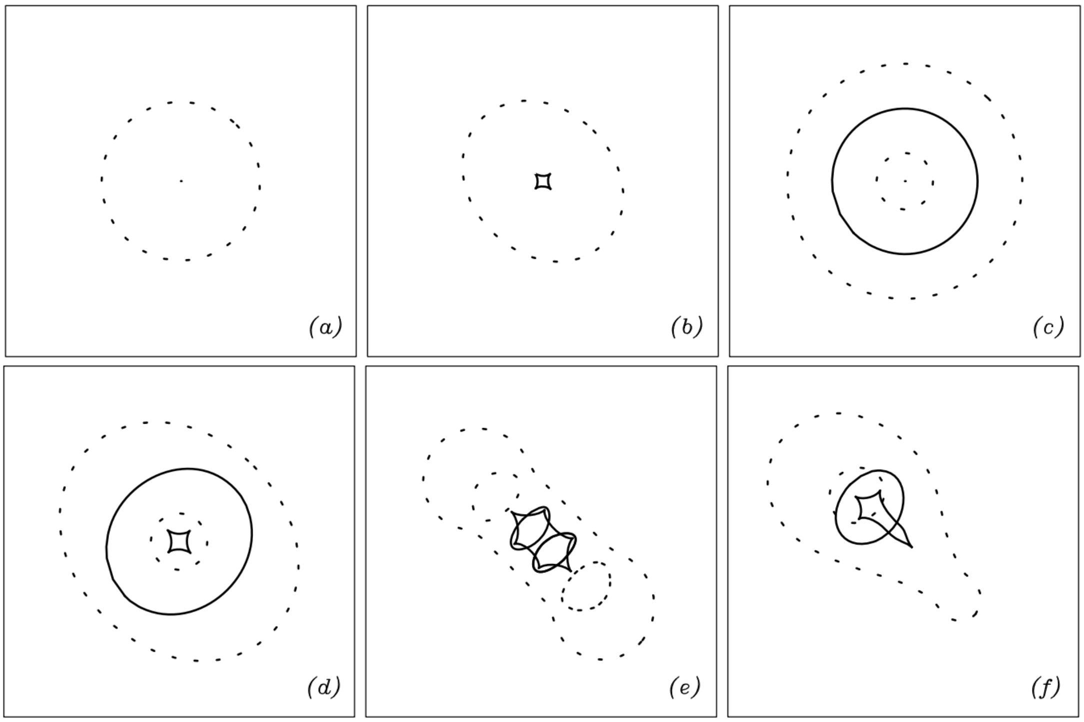


Figure 3.3: Critical lines (dashed) and caustics (solid) for different mass models: (a) for a singular isothermal circular mass distribution, the radial critical line is the central point, and the corresponding caustic is at infinity; (b) for a singular isothermal elliptical mass distribution, the tangential caustic line is an astroid; (c) for a circular mass distribution with an inner slope shallower than isothermal mass distribution, a radial critical curve appears, and both caustics are circles; (d) is the same as (c), but for an elliptical mass distribution, the relative size of both caustic lines will depend on the mass profile and the ellipticity of the mass disribution; (e) for a bimodal mass distribution with two clumps of equal mass, similar to (d); and (f) for a bimodal distribution with different unequal masses. (Figure credit: Kneib & Natarajan, 2011a)

defines two closed lines, called *critical lines.* The critical lines define the limits that can produce multiple images in the image plane, where outside the closed lines multiple images merge and disappear. If an image lies exactly on the critical line, then the flux of the source is infinitely amplified. The location of critical lines can be mapped back to the source plane, and the corresponding limits, called *caustic lines*, are defined. We defined the Jacobian matrix, A, in the Cartesian frame in Equation 3.2.13, which can be rewritten in polar coordinates (r, θ) as (Kneib & Natarajan, 2011a)

$$A = \begin{pmatrix} 1 - \frac{\partial^2\psi}{\partial r\partial r} & -\frac{1}{\partial r}\left(\frac{\partial\psi}{r\partial\theta}\right) \\ -\frac{1}{\partial r}\left(\frac{\partial\psi}{r\partial\theta}\right) & 1 - \frac{\partial\psi}{r\partial r} - \frac{1}{r^2}\frac{\partial\psi}{r\partial\theta}. \end{pmatrix} \tag{3.2.19}$$

For a circularly symmetric mass distribution, the Jacobian Matrix can be simplified to:

$$A = \begin{pmatrix} 1 - \frac{\partial^2\psi}{\partial r\partial r} & 0 \\ 0 & 1 - \frac{\partial\psi}{r\partial r}. \end{pmatrix} \tag{3.2.20}$$

Therefore, the two critical lines can also be defined as $(1 - \frac{\partial^2\psi}{\partial r\partial r}) = 0$ and $(1 - \frac{\partial\psi}{r\partial r}) = 0$, called the *radial* and *tangential critical curves* respectively. They refer to the deformations in radial and tangential directions. For a mass distribution with circular symmetry, critical and caustic lines are circles and the tangential caustic line always reduces to a single point. Some significant properties can be derived by studying the critical curve. First, a projected mass with circular symmetry enclosed within the radius r is

$$M(r) = \frac{c^2}{4G}\frac{D_L D_s}{D_{LS}} r \frac{\partial\psi(r)}{\partial r}. \tag{3.2.21}$$

Using the definition of the tangential critical curve, $r = \partial\psi(r)/\partial r$, the mass enclosed within the tangential critical radius (also called Einstein radius θ_E) can thus be written as

$$M(\theta_E) = \pi\Sigma_{\rm crit}\theta_E^2, \tag{3.2.22}$$

which is the same as Equation 3.2.4. Therefore, given the known redshifts of the lens and the source and the underlying cosmology, the position of the tangential critical curve can be used to precisely determine the enclosed mass within a circular aperture. Secondly, according to the definition of the radial critical curve,

$$\frac{\partial^2\psi(r)}{\partial r\partial r} = \frac{\partial}{\partial r}\left(\frac{M(r)}{\pi\Sigma_{\rm crit} r}\right) = 1, \tag{3.2.23}$$

its location depends on the gradient of the mass profile, meaning that given the location of the radial critical curve and the information in Σ_{crit}, the slope of the mass profile close to the cluster center can be constrained. Information about the critical lines are therefore valuable in probing the lens objects, galaxy clusters, in this thesis. In the case of general non-circular mass distributions, critical lines cannot be expressed analytically except for certain simple elliptical mass profiles. Figure 3.3 shows critical and caustic curves for different mass distributions. For a singular isothermal circular mass distribution, the tangential critical line is a circle while the radial one is a central point. In the case of a singular isothermal elliptical mass distribution, the tangential critical line is an ellipse and its caustic line an astroid. If a mass distribution has an inner slope shallower than the isothermal mass profile, the radial critical curve will appear. For a more complex mass distribution, such as bimodal distribution, two sets of tangential and radial critical lines are produced.

Multiple images

As we have shown, critical curves are useful for the determination of the mass of the lens. While they cannot be directly mapped, multiple images can be used to constrain the location of the critical curves. For example, images distorted in the tangential (radial) direction are located close to the tangential (radial) critical lines. The number of multiple images equals the number of solutions to the lens equation (Equation 3.2.2), which depends on the complexity of the mass distribution of the lens. It is predicted by theory that an odd number of sources is produced for one source. However, some images can be less magnified, or de-magnified, so that they cannot be observed. For a cluster dominated by a single halo, *fold*, *cusp* and *radial arcs* can be observed depending on the configuration of the lens system, as shown in figure 3.4. Outside the critical curves, only one image can be observed. When the source moves across the radial caustic line, two additional lensed images are produced which are stretched radially and appear near the radial critical curve. The positions of the radial multiple images can be used to probe the shape of the density profile in the central regions. When the source moves close to the astroid spike of the tangential caustic line, cusp arcs are formed with three images located

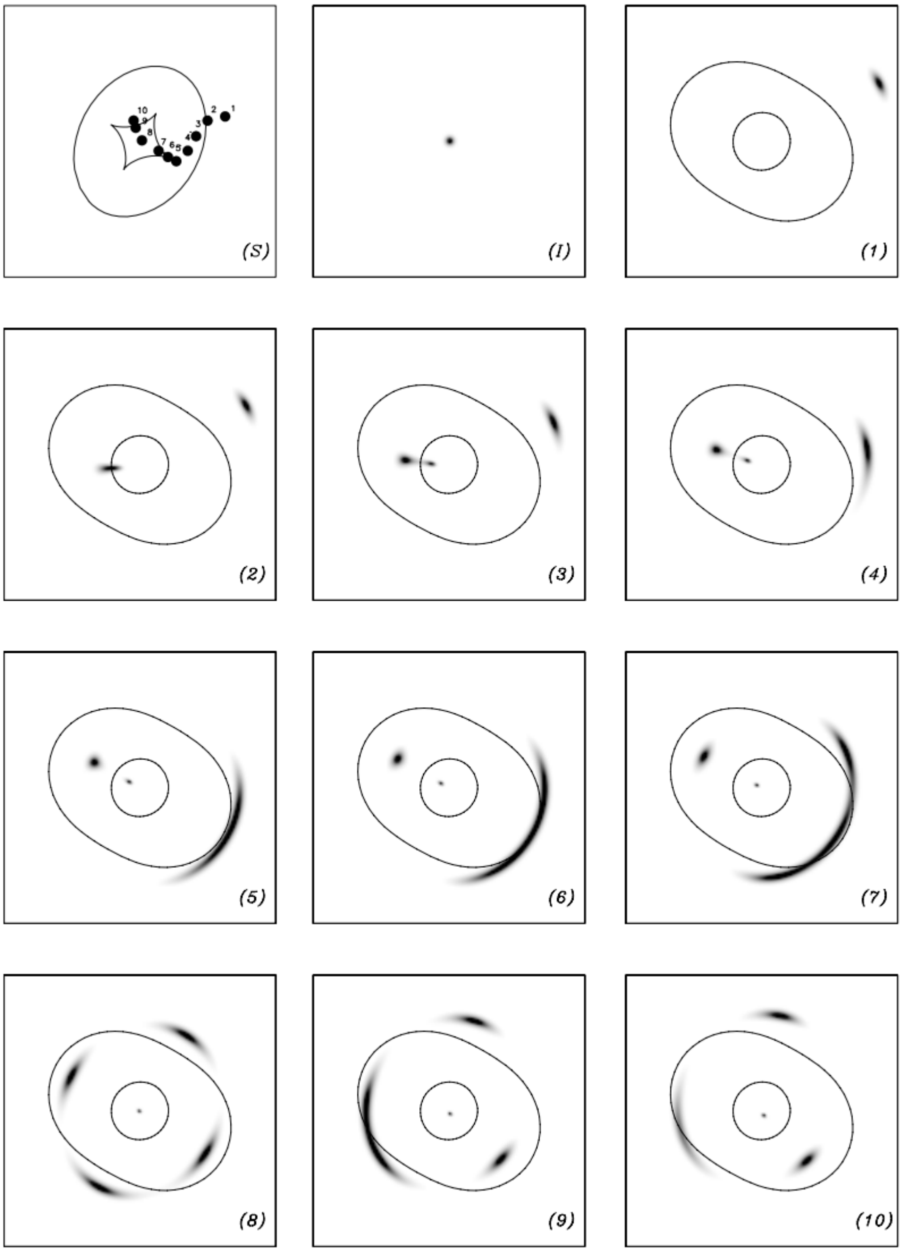


Figure 3.4: Multiple-image configurations produced by a simple elliptical mass distribution. The panel (S) shows the caustic lines in the source plane and the source positions numbered 1 to 10. The panel (I) is the image of the source without any lensing effect. The panels (1) to (10) show the lensed images for the various source positions in panel (S). Certain configurations are very typical and are denoted as follows: (3) radial arc, (6) cusp arc, (8) Einstein cross, (10) fold arc. (Figure credit: Kneib & Natarajan, 2011a)

at the epicentres of the tangential critical line. For a fold configuration, the source is close to the astroid side of the tangential caustic line, with the two images produced on both sides of the tangential critical line, and the third image produced in the opposite side. The well-predicted configuration of multiple images are valid only for simple elliptical mass model. A bimodal cluster or other complex structures can produce more complicated configurations. For more details in multiple images and the study of strong lensing, we refer the reader to the review of Kneib & Natarajan (2011a). In the following chapters of this thesis, we will mainly focus on the weak lensing analysis.

3.3 Weak Gravitational Lensing in Galaxy Clusters

The study of gravitational lensing in galaxy clusters can be seperated into two aspects: strong and weak lensing. In the central region of clusters where the mass density is high with $\kappa \sim 1$, multiple images can be clearly observed and used to constrain mass distribution of the cluster core. Outside the strong lensing region, $\kappa < 1$, lensing distortions get smaller. The shapes of background galaxies are dominated by their intrinsic shape, which is assumed to have random orientations. Therefore, weak lensing needs to be measured statistically, such as the mean distortion of a sample of background sources. After that, the underlying mass distribution of the lens cluster on larger scales can be reconstructed from the weak lensing shear signal. Weak lensing is normally observed in the optical and near-infrared bands. The images are then processed to detect individual galaxies and to extract the lensing signal from the background galaxy shapes. In this section, we will first present the galaxy shape measurement and the weak lensing shear estimation for *Hubble Space Telescope* (*HST*) imaging data. The second part is devoted to the reconstruction of the surface mass density distribution, κ. There are two mass reconstruction methods used to analyze the cluster targets in this thesis. The details of them are presented in Section 3.4.

3.3.1 Galaxy Shape Measurements for *HST* Data

Measuring galaxy shapes to a very high precision is essential in weak lensing because galaxy ellipticities are typically distorted by only a few per cent. It is challenging to detect this weak signal because the image shapes are also affected by convolution with the point spread function (PSF) of the telescope, detector and atmosphere, resulting in a blurred image in observations. These instrumental effects must be modelled and corrected before estimating the lensing signal. In this thesis, we focus on the weak lensing shear calibration on the *HST* Advanced Camera for Survey (ACS) images, using the publicly available *HST* weak lensing shape measurement code PYRRG (Harvey et al., 2019). This PYTHON3.7 code is based on Rhodes et al. (2000) method (hereafter called RRG), designed for correcting the small, diffraction-limited PSF obtained from space, and has been calibrated on simulated data containing a known shear (Leauthaud et al., 2007). The algorithm of PYRRG is graphically illustrated in Fig 3.5. It consists of six key steps:

1) Source finding

Galaxies are identified in the *HST*/ACS image using the SEXTRACTOR photometry package (Bertin & Arnouts, 1996). To maximise sensitivity to distant (small and faint) galaxies that contain most of the lensing signal, PYRRG adopts the 'Hot–Cold' technique originally developed by Rix et al. (2004). This technique first runs with a 'hot' scan, using a smaller minimum number of pixels to count as a source to the faint objects which contain most of the lensing signal. However, these low detection thresholds may deblend the outer features of bright galaxies. To avoid spurious deblending of the largest galaxies, we then run a 'cold' scan optimised to only detect the brightest objects. The final source catalogue consists of all cold and hot detections not overlapping with a cold object. We record the detection signal-to-noise ratio of each object from SEXTRACTOR parameters, S/N ≡ FLUX_AUTO/FLUXERR_AUTO.

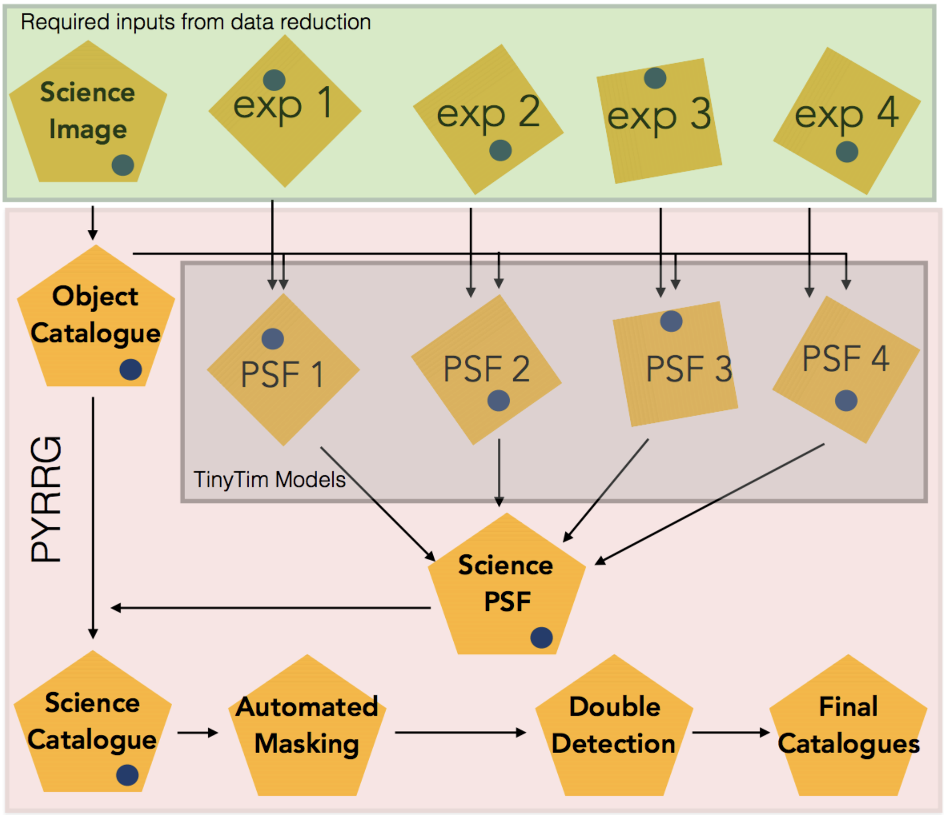


Figure 3.5: The flowchart of the pyRRG algorithm. Science image and all the associated exposures are input to the pipeline. Galaxy catalogues are generated. pyRRG measures the PSF from the TinyTim models for each exposure and combines them to produce a stacked PSF. The shape of galaxies are corrected according to the PSF model, and the lensing signal is estimated from that. Finally, several cleaning procedures are carried out to produce a clean shear catalogue. (Figure credit: Harvey et al., 2019)

2) Moment measurement

Following the source detection, PYRRG measures the weighted multipole moments of each object in order to characterise its shape. We define the second and fourth order normalized moments of surface brightness distribution of each galaxy,

$$\begin{aligned} I_{ij} &= \frac{\sum \omega(\theta)\theta_i\theta_j I(\boldsymbol{\theta})}{\sum \omega(\theta) I(\boldsymbol{\theta})}, \\ I_{ijkl} &= \frac{\sum \omega(\theta)\theta_i\theta_j\theta_k\theta_l I(\boldsymbol{\theta})}{\sum \omega(\theta) I(\boldsymbol{\theta})}, \end{aligned} \tag{3.3.24}$$

where I is the pixel intensity, ω is a Gaussian weight function included to suppress noise, and the sum is over all pixels. There are multiple ways to quantify the shape of an image, and in PYRRG we follow Rhodes et al. (2000) to define the size as the combination of the quadrupole moments,

$$d = \sqrt{\frac{(I_{xx} + I_{yy})}{2}} \tag{3.3.25}$$

and ellipticity $\epsilon \equiv \sqrt{\epsilon_1^2 + \epsilon_2^2}$ with orientation $\varphi = \frac{1}{2}\text{arctan}\,(\epsilon_1/\epsilon_2)$, where

$$\begin{aligned} \epsilon_1 &= \frac{I_{xx} - I_{yy}}{I_{xx} + I_{yy}}, \\ \epsilon_2 &= \frac{2I_{xy}}{I_{xx} + I_{yy}} \ . \end{aligned} \tag{3.3.26}$$

The relation between the ellipticity defined here, and the semi-major axis a and the semi-minor axis b of an ellipse, is given by

$$\epsilon = \frac{a^2 - b^2}{a^2 + b^2} \ . \tag{3.3.27}$$

3) Star - Galaxy Classification

Following the measurements of the normalised image moments, we perform the star-galaxy classification by examining the distribution of objects in the overall brightness (MAG_AUTO) and peak surface brightness (MU_MAX) plane. This diagram allows us to separate three distinct categories: galaxies, stars, and noises (i.e., artifacts and residual cosmic rays). In this thesis, we classify different objects manually. But PYRRG recently implements a new automatic classification using Random Forest approach.

4) PSF Measurement

The shape of galaxies is set by lensing but blurred by the PSF. It is necessary to model the PSF and correct the observed galaxy shapes since lensing measurements are particularly sensitive to it. Compared with all other telescopes, the *HST* PSF is very stable. However, according to Rhodes et al. (2007), the ACS PSF varies over time, as thermal expansion and contraction change the distance between the primary and secondary mirrors due to the heating of the Sun. Even a few microns away from nominal focus, the PSF becomes larger and more elliptical, and there is no atmosphere to average away variations. The PSF can be measured from non-saturated stars within an image, and can be split into two components. The isotropic part of the PSF circularizes an image, enlarging shapes and reducing their ellipticity. The anisotropic component elongates all objects within the image, increasing their ellipticity. Both components affect the observed shear and need to be corrected. First we measure the high order moments of the stars in each of the individual exposures, and then we compare ellipticities of observed stars with models created by the TINYTIM ray-tracing software (TINYTIM; Rhodes et al., 2007) which creates simulated PSF images at varying focus offsets. This comparison allows us to determine the effective focus of the images. Once their effective focus positions have been determined, the shape moments of the net PSF are interpolated by a polynomial fit to the known positions of the galaxies. Finally, the PSFs from each individual exposure at the position of the galaxy are stacked and combined to mimic the drizzling of multiple exposures to acquire the final PSF. The moments of galaxies measured in 4) are then corrected from the PSFs linearly, size and ellipticity are re-calculated. For a full description of the PSF correction, please see Rhodes et al. (2000).

5) Shear estimation

Having corrected galaxy shapes from the impact of instrumentation, we then calculate the gravitational shear. In theory, the transformation between the image complex ellipticity, $\boldsymbol{\epsilon} = \epsilon_1 + i\epsilon_2$, defined in Eq. 3.3.26 and the intrinsic ellipticity of source, $\boldsymbol{\epsilon}^{(s)}$, is given by (Schneider & Seitz, 1995)

$$\boldsymbol{\epsilon}^{(s)} = \frac{\boldsymbol{\epsilon} - 2\boldsymbol{g} + \boldsymbol{g}^2\epsilon^*}{1 + |\boldsymbol{g}|^2 - 2\,\mathbb{R}\mathrm{e}(\boldsymbol{g}\boldsymbol{\epsilon}^*)}, \tag{3.3.28}$$

where the asterisk denotes complex conjugation, $\mathbb{R}\mathrm{e}(x)$ defines the real part of the complex number x, and $\boldsymbol{g}$ is the reduced shear (Equation 3.2.17). Equation 3.3.28 shows that the transformation of ellipticity between the source and the image depends only on the reduced shear. Thus, the reduced shear is the only observable quantity which can be calibrated from the measurements of image ellipticities. In the weak lensing limit ($|\kappa|, |\gamma| \ll 1$), the transformation between the image ellipticity and its intrinsic source ellipticity can be simplified as $\boldsymbol{\varepsilon} \simeq \boldsymbol{\varepsilon}^{(s)} + 2\boldsymbol{\gamma}$. This relation requires information about the intrinsic shape of each galaxy. However, galaxies display a very wide range of shapes which makes it difficult to measure the low shear signal for an individual object. Consequently, weak lensing shear can only be measured statistically. Assuming the random orientation of sources, the ensemble average of intrinsic ellipticity of background sources vanishes: $< \epsilon^{(\mathrm{s})} >= 0$, thus the local distortion can be inferred from a local ensemble of image ellipticities

$$< \boldsymbol{\epsilon} > \simeq 2 < \gamma > . \tag{3.3.29}$$

In the case of RRG, the shear estimator is more sophisticated and has been successfully calibrated and tested in Leauthaud et al. (2007),

$$\boldsymbol{\gamma} = C\frac{\boldsymbol{\epsilon}}{G} \tag{3.3.30}$$

It incorporates the calibration factor, $C = (0.86^{+0.07}_{-0.05})^{-1}$, which is empirically measured from mock *HST* images in the same band and to the same depth (Leauthaud et al., 2007) and the "shear susceptibility factor" G is measured from the global distribution of ϵ and fourth-order moments (Rhodes et al., 2000),

$$G = 2 - < \boldsymbol{\epsilon}^2 > -\frac{1}{2} < \lambda > -\frac{1}{2} < \boldsymbol{\epsilon} \cdot \boldsymbol{\mu} >, \tag{3.3.31}$$

where

$$\begin{aligned} \lambda &= (I_{1111} + 2I_{1122} + I_{2222})/(2d^2\omega^2), \\ \mu_1 &= (-I_{1111} + I_{2222})/(2d^2\omega^2), \\ \mu_2 &= -2(I_{1112} + I_{1222})/(2d^2\omega^2), \end{aligned} \tag{3.3.32}$$

and ω is the size of the Gaussian weight function in Eq. 3.3.24. From these, we have the final estimator of the shear γ.

6) Catalogue cleaning and masking

Finally we apply cleaning operations and lensing cuts to the weak lensing shear catalogue to remove spurious and duplicated detections. We first generate polygons using the known positions of stars and saturated stars, and mask any object that lies within these polygons. Secondly, for double detections, we remove objects that lie within the isophote of a larger object. In the last step in constructing a "clean" shear catalogue, we eliminate sources with uncertain shapes. Shapes of very small or faint galaxies are difficult to measure and may be biased, since they are dominated by that of the PSF. We therefore exclude galaxies with size $d < 0.11''$, detection S/N< 4.5, or unphysical values of $\epsilon > 1$ (which can arise after PSF correction in the presence of noise; for a discussion of this effect, see Jauzac et al., 2012).

3.4 Weak Lensing Mass Mapping Techniques

Traditionally, the modelling of the cluster mass distribution in lensing can be done by using "parametric models" or "non-parametric approach". The major distinction between these two mehtods is whether the calculation is "model-based" (parametric) or "model-free" (non-parametric). In the parametric method, the mass distribution is described by a finite number of physical halo models. In practice, this can be done by fitting the observed data, such as shear or magnification, to different mass models (see Section 2.2.5) with relatively few defined parameters. A catalogue of mass models for lensing is summarized in Keeton (2001b). Although a parametric model provides a simple description of the cluster mass, the cluster's spatial distribution may not be well-described by any of the common parametric models, especially for merging clusters. Moreover, the flexibility of a parametric model is limited to only a few free parameters and the restricted shapes. On the other hand, the free-form or non-parametric models, provide a more flexible and accurate description of the mass distribution. For most of the "non-parametric" methods, the mass distribution is typically pixelated into a regular grid of small mass elements. However, these large number of parameters can lead to over-fitting issues. Further detail of non-parametric methods are presented later on.

In this section, I will focus on two frequently used methods to reconstruct the mass: the one frequentist, the other Bayesian. These two methods will be used in the following chapters.

3.4.1 Direct inversion with KS93+MRlens

The first method we used to reconstruct the mass is the simple inversion method. According to Eqs. 3.2.11 and 3.2.14, the convergence and shear can be written as

$$\kappa(\boldsymbol{\theta}) = \frac{1}{2}\partial^*\partial\psi(\boldsymbol{\theta}), \tag{3.4.33}$$

$$\gamma(\boldsymbol{\theta}) = \frac{1}{2}\partial\partial\psi(\boldsymbol{\theta}) \equiv \hat{D}_\theta\psi(\boldsymbol{\theta}), \tag{3.4.34}$$

where $\partial \equiv \partial_1 + i\partial_2$ is a complex gradient operator, and $\hat{D}_\theta = \partial\partial/2 = (\partial_1^2 - \partial_2^2)/2 + i\partial_1\partial_2$ is a spin-2 operator. Since both κ and γ fields are linear combinations of the second derivatives of $\psi(\boldsymbol{\theta})$, they are related to each other (Kaiser, 1995) and this non-local relation between κ and γ is given as (Kaiser & Squires, 1993),

$$\kappa(\boldsymbol{\theta}) = \frac{1}{\pi}\int d^2\theta' D^*(\boldsymbol{\theta} - \boldsymbol{\theta'})\gamma(\boldsymbol{\theta'}) \tag{3.4.35}$$

$$\gamma(\boldsymbol{\theta}) = \frac{1}{\pi}\int d^2\theta' D(\boldsymbol{\theta} - \boldsymbol{\theta'})\kappa(\boldsymbol{\theta'}) \tag{3.4.36}$$

where $D(\boldsymbol{\theta})$ is the complex kernel defined as

$$D(\boldsymbol{\theta}) \equiv 2\pi\hat{D}_\theta \triangle^{-1} (\boldsymbol{\theta}) = \frac{\theta_2^2 - \theta_1^2 - 2i\theta_1\theta_2}{|\theta|^4}, \tag{3.4.37}$$

where the $*$ denotes the complex conjugate and $\triangle^{-1}(\boldsymbol{\theta}) = \ln|\boldsymbol{\theta}|/2\pi$ is the Green's function. Equation 3.4.35 shows that the lens' convergence field is a convolution of the shear γ with the complex kernel, D^*, meaning that κ can be directly calculated if the shear field, γ, induced by the deflector can be measured locally as a function of the angular position, $\boldsymbol{\theta}$. In practical application, these relations are usually inverted in Fourier space to obtain linear functions between the Fourier transformation of κ and γ (denoted as $\hat{\kappa}$ and $\hat{\gamma}$),

$$\hat{\kappa}(\boldsymbol{k}) = \frac{k_1^2 - k_2^2 - 2ik_1k_2}{k_1^2 + k_2^2}\hat{\gamma}(\boldsymbol{k}), \tag{3.4.38}$$

$$\begin{aligned}\hat{\gamma}_1(\boldsymbol{k}) &= \frac{k_1^2 - k_2^2}{k_1^2 + k_2^2}\,\hat{\kappa}(\boldsymbol{k})\,,\\ \hat{\gamma}_2(\boldsymbol{k}) &= \frac{2k_1k_2}{k_1^2 + k_2^2}\,\hat{\kappa}(\boldsymbol{k})\,,\end{aligned} \tag{3.4.39}$$

where $\boldsymbol{k} = (k_1, k_2)$ is the wave vector conjugate to $\boldsymbol{\theta}$. These transformations (called "KS93" hereafter) will be used to reconstruct the surface mass density from the measured shear. However, in observational analyses, the intrinsic shape noise of background galaxies induces lots of artifacts in the reconstructed mass maps which can bias the inferred mass distributions of the clusters. In the original incarnation of KS93, it was usual to suppress noise by convolving the reconstructed mass distribution with a smoothing kernel whilst in Fourier space. In this study, we omit this step in order to initially preserve as much spatial resolution as possible. To suppress the high frequency noise, we then filter the noisy convergence map by the Multi-Resolution method for gravitational lensing (MRLens[2]; Starck et al. 2006). Note that a 3D extension of the method, GLIMPSE, has also been developed (Leonard et al., 2015).

MRLens is a non-linear filtering method that decomposes an image into wavelets and applies non-linear regularisation on each wavelet scale. It aims to retain statistically significant signals but minimise noise through an approach that, under the assumption of a multiscale entropy prior, optimises the False Discovery Ratio (FDR) of false detections to true detections. The multiscale entropy method considers the entropy of an input image as the sum of the information at each scale of its wavelet transformation which is related to the probability of being caused by noise. For the non-significant wavelet coefficients selected by FDR, a regularization (i.e. filtering) is applied on its entropy to suppress noise. Starck et al. (2006) describe the method fully, and demonstrate that MRLens outperforms Gaussian or Wiener filtering in this context. In particular, Pires et al. (2010) show that MRLens improves the reconstruction of non-Gaussian structures, like the distribution of mass in galaxy clusters.

---

[2] We implement MRLens using the June 26, 2017 version of software available from `https://www.cosmostat.org/software/mrlens`.

3.4.2 Multi-scale grid technique

The second method is based on the adaptive grid technique developed by Jullo & Kneib (2009), using the Lenstool[3] software. This non-parametric technique was originally designed for strong lensing mass modelling, and then adapted for weak lensing. In the following, I will give a brief introduction to this technique. For a full description, we refer the reader to Jauzac et al. (2012).

The fundamental idea is to follow the physical assumption of light-traces-mass. We first create a multi-scale grid of Radial Basis Functions (RBFs) based on the light distribution of the cluster. We initialise a grid of points by drawing a large hexagon over the field of interest, split into six equilateral triangles (see Figure. 1 in Jullo & Kneib, 2009). It is then recursively refined by applying a splitting criterion which is based on the surface density of the light map. More precisely, if a single pixel inside any of these triangles exceeds a predefined light-surface-density threshold, we split that triangle into four smaller triangles. This refinement continues for several levels of recursion, until the brightest parts of the cluster are covered by the highest resolution grid. The resulting grid of RBFs of different sizes provides a higher resolution grid with more flexibility to describe the higher density region, making this method ideal for describing the irregular mass distribution. Following the creation of multi-scale grid, at each grid node we place a physically motivated halo model, circular ($q = 1$) tPIEMD (Eq. 2.2.30), with core radius $r_{\rm c}$ set to the side length of the triangle, the truncation radius $r_{\rm t}$ is set to three times the core radius, and the velocity dispersion is free to vary.

We optimise free parameters in this model using the MassInf Markov Chain Monte Carlo algorithm implemented in Lenstool. The parameter space is highly dimensional. So to optimise the multiscale grid, we use the Bayesian MassInf algorithm and the Gibbs approach to maximise the likelihood

$$\mathcal{L}_\gamma = \frac{1}{Z_L} \exp\left(-\frac{\chi^2}{2}\right), \tag{3.4.40}$$

---

[3] We implement Lenstool using version 7.1 of the software available from `https://projets.lam.fr/projects/Lenstool/wiki`.

where χ^2 is the goodness-of-fit statistic

$$\chi^2 = \sum_{i=1}^{M} \sum_{j=1}^{2} \frac{(\gamma_{j,i} - 2\gamma_{j,i}^{\text{model}}(\boldsymbol{R}_i))^2}{\sigma_\gamma^2} \tag{3.4.41}$$

(following Schneider et al. 2000)[4], and M is the number of background sources. These expressions are normalised by the statistical uncertainty in the shear measured from each galaxy

$$\sigma_\gamma^2 = \sigma_{\gamma,\text{intrinsic}}^2 + \sigma_{\gamma,\text{measurement}}^2 \ , \tag{3.4.42}$$

which is a combination of intrinsic shape noise and shape measurement error. The normalization factor is given as

$$Z_L = \prod_{i=1}^{M} \sqrt{2\pi}\sigma_{\gamma i}. \tag{3.4.43}$$

At each step of the iteration, the 2% most discrepant masses are adjusted. One should note that Lenstool assumes that mass densities are positive by definition. This is not necessarily true since we are really fitting departures from the mean density of the Universe. For example, the convergence of the large-scale structures is consistent with fluctuations around zero. However, it is frequently used, and reasonable near the extreme mass of a galaxy cluster. From the resulting MCMC sample of the mass map, we finally compute the marginalised mean convergence and its 68% confidence limits.

---

[4] Note that we include a factor of 2 because Lenstool takes inputs in the form of ellipticity $e = (a^2 - b^2)/(a^2 + b^2)$ instead of shear (Jullo et al., 2014a).

Chapter 4

Mapping dark matter and finding filaments: calibration of lensing analysis techniques on simulated data

4.1 Abstract

We quantify the performance of mass mapping techniques on mock imaging and gravitational lensing data of galaxy clusters. The optimum method depends upon the scientific goal. We assess measurements of clusters' radial density profiles, departures from sphericity, and their filamentary attachment to the cosmic web. We find that mass maps produced by direct inversion (KS93) of shear measurements are unbiased, and that their noise can be suppressed via filtering with MRLens. Forward-fitting techniques, such as Lenstool, suppress noise further, but at a cost of biased ellipticity in the cluster core and over-estimation of mass at large radii. Interestingly, current searches for filaments are noise-limited by the intrinsic shapes of weakly lensed galaxies, rather than by the projection of line-of-sight structures. Therefore, space-based or balloon-based imaging surveys that resolve a high density of lensed galaxies, could soon detect one or two filaments around most clusters.

The following has been published in MNRAS as Tam et al. (2020a)

4.2 Introduction

The CDM standard model of cosmology suggests that structures in the Universe formed hierarchically, via mergers of small over-densities in the early Universe into larger and larger objects (Springel et al., 2005; White & Rees, 1978). Thirteen billion years after the Big Bang, the largest objects are currently clusters of hundreds or thousands of galaxies. Their properties depend on the growth of structure, which emerges from primordial density fluctuations, collapses under gravity, and is funnelled into the cluster along filaments. They can therefore be used to test cosmological models (e.g. Bahcall & Cen, 1993; Jauzac et al., 2016; Meneghetti et al., 2005; Rozo et al., 2010; Schwinn et al., 2017; de Haan et al., 2016).

Gravitational lensing is particularly efficient at investigating clusters. When a light ray emitted by a distant source passes though a massive structure, the path of this light ray is deflected by the gravitational field, resulting in distorted images of the distant source. To study this image distortion, we can probe the distribution of total mass content, including dark matter, in our Universe.

Ground-based observations of gravitational lensing by galaxy clusters have been successfully used to measure clusters' average or bulk properties, such as mass (e.g. Herbonnet et al., 2019; McClintock et al., 2019; Medezinski et al., 2017; Miyatake et al., 2019; Okabe & Smith, 2016; Rehmann et al., 2019; Schrabback et al., 2018; Sereno et al., 2017; Umetsu et al., 2014b; Umetsu et al., 2019; von der Linden et al., 2014), and ellipticity (e.g. Chiu et al., 2018; Clampitt & Jain, 2016; Evans & Bridle, 2009; Oguri et al., 2010; Shin et al., 2018; Umetsu et al., 2018; van Uitert et al., 2017). The CLASH survey (Cluster Lensing and Supernova Survey with Hubble; Postman et al., 2012) measured the mass and concentration of 25 clusters, by combining wide-field Subaru imaging with *Hubble Space Telescope* (*HST*) imaging of the cluster cores (Merten et al., 2015). However, ground-based observations have yielded only marginally significant detections of filaments (e.g. Clowe et al., 2006; Dietrich et al., 2012; Gavazzi et al., 2004; Gray et al., 2002; Kaiser et al., 1998; Martinet et al., 2016), whose dark matter density is too low (and the filaments too narrow to resolve).

Space-based imaging reveals the shapes of more background galaxies, and increases the S/N of lensing measurements in multiple resolution elements across an individual cluster. Thus the shape and morphology of individual mass distributions can be precisely *mapped*, without the need to average out features over a population of clusters. Space-based lensing reconstructions have resolved substructure near cluster cores (e.g. Merten et al., 2011; Natarajan et al., 2017); bimodality even in relatively distant clusters like the 'Bullet Cluster' (Bradac et al., 2006) or 'El Gordo' (Jee et al., 2014); and filaments in Abell 901/902 (Heymans et al., 2008) and MACSJ 0717+3745 (Jauzac et al., 2012). Nonetheless, these analyses remain rare because the $\sim 3' \times 3'$ field of view of *HST*'s Advanced Camera for Surveys (ACS) is smaller than a typical cluster's angular size. Furthermore, both of *HST*'s contiguous surveys (GOODS and COSMOS) unluckily sampled regions of the Universe that are underdense at the $z = 0.2$–0.4 redshifts where lensing is most sensitive (Heymans et al., 2005; Krolewski et al., 2018; Massey et al., 2007a), so happen to contain few lensing clusters (Guzzo et al., 2007). Until recently, only around one cluster, MS 0451-03, had a dedicated wide-field mosaic of contiguous *HST* imaging been obtained (Moran et al., 2007b).

With the capabilities of the next-generation of telescopes, there will soon be more wide-field, space-resolution imagings taken through the *HST*/BUFFALO survey (Steinhardt et al., 2020), *SuperBIT* (Redmond et al., 2018; Romualdez et al., 2016), *Euclid* (Laureijs et al., 2011) and *Nancy Grace Roman Space Telescope* (Spergel et al., 2013).

The intent of this work is to prepare for future observations, much as Van Waerbeke et al. (2013) calibrated mass mapping methods for the current generation of wide-field ground-based lensing surveys. We use mock space-based weak-lensing data to develop and quantify the performance of two different methods to map dark matter around galaxy clusters, to measure deviations from sphericity, and to search for filaments connecting it with the cosmic web. Where we must make decisions about general properties (e.g. distance, mass) of clusters that we simulate, we shall use MS 0451-03 as a template, so our predictions can be immediately tested on real observations (see Chapter 5).

This chapter is organised as follows. We introduce the simulated data in Section

4.3. We describe weak-lensing mass mapping and show the results of mass reconstruction in Section 4.4. The halo shape measurements are presented in Section 4.5. We finally search for filamentary structures from the reconstructed mass maps in Section 4.6 and conclude in Section 4.7. Throughout the chapter, we define angular diameter distances assuming a background cosmology with $\Omega_m = 0.287$, $\Omega_\Lambda = 0.713$, and $h = H_0/100\mathrm{kms}^{-1}\mathrm{Mpc}^{-1} = 0.693$ (WMAP 9-year cosmology; Hinshaw et al., 2013). All magnitudes are quoted in the AB system.

4.3 Simulation Data

We use N-body particle data from the BAHAMAS suite of cosmological simulations (McCarthy et al., 2017, 2018). These were run with different background cosmologies and implementations of sub-grid galaxy formation physics, designed to test the impact of baryonic physics on large-scale structure tests of cosmology. For this chapter, we use the version with a WMAP 9-year (Hinshaw et al., 2013) cosmology, and sub-grid feedback model that is calibrated to produce a good match to the observed stellar mass function, X-ray luminosities and gas fractions of galaxy clusters. This simulation occupies a periodic cubic volume, $400\,h^{-1}$ Mpc on a side, with dark matter and (initial) baryon particle masses of $5.5 \times 10^9\,\mathrm{M}_\odot$ and $1.1 \times 10^9\,\mathrm{M}_\odot$, respectively.

4.3.1 Distribution of mass in clusters

We extract the ten most massive clusters from the $z = 0.5$ simulation snapshot. We first use the friends-of-friends algorithm (FOF; More et al. 2011) to identify all matter overdensities. For each FOF group, we calculate r_{200} and M_{200}, the total mass enclosed within this sphere. For the ten most massive clusters, which have $4 \times 10^{14}\mathrm{M}_\odot < M_{200} < 2 \times 10^{15}\mathrm{M}_\odot$, we store the 3D distribution of dark matter, stars and gas.

To generate a 2D, pixellated convergence map, we follow the method of Robertson et al. (2019). In summary, we project the location of all simulation particles within $5\,r_{200}$ of the centre of a cluster along a line of sight (here, the simulation z-

	$M_{\mathrm{FOF}}(10^{14}\mathrm{M}_\odot)$	$M_{200}(10^{14}\mathrm{M}_\odot)$
Group 1	27.7	17.3
Group 2	17.9	15.0
Group 3	17.8	17.7
Group 4	16.6	14.6
Group 5	14.3	9.7
Group 6	13.3	11.0
Group 7	12.9	8.9
Group 8	11.1	4.0
Group 9	9.4	8.2
Group 10	9.3	5.7

Table 4.1: Masses of the 10 most massive BAHAMAS clusters used in this analysis. We here give their FOF mass, M_{FOF}, and M_{200}.

axis). In a 25×25 Mpc (2048×2048 pixel) map centred on the most bound particle, we use an adaptive triangular shaped cloud scheme to smooth each particle's mass over a kernel whose size depends on the 3D distance to that particle's 32nd nearest neighbour. Resulting convergence maps are shown in figure 4.1, adopting the lens redshift $z_l = 0.55$ of galaxy cluster MS0451-03 as a concrete example, and source redshift $z_s = 0.97$ typical of *HST* observations to single-orbit depth (Leauthaud et al., 2007). The masses of the clusters are listed in Table 4.1.

Before proceeding further, we identify 40 filaments in the ten projected mass maps, defined as radially extended regions with convergence $0.005 < \kappa < 0.01$, which is equivalent to a surface density of $1.7 \times 10^7 < \Sigma\,(\mathrm{M}_\odot/\mathrm{kpc}^2) < 3.4 \times 10^7$. These are indicated by white dashed lines in the bottom panel of figure 4.1.

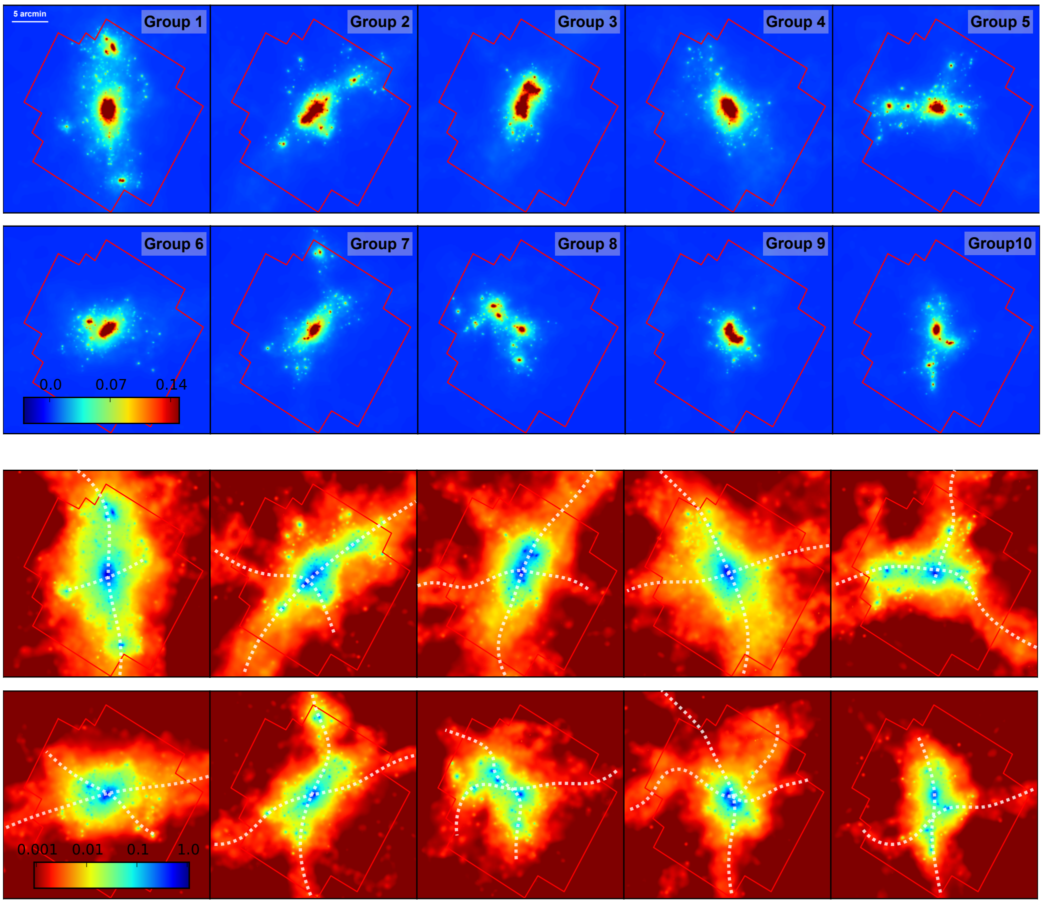


Figure 4.1: Noise-free maps of the total mass distribution in the ten most massive clusters of the BAHAMAS simulations, projected along a randomly-oriented line of sight. Clusters have masses M_{200} from $2 \times 10^{15}\mathrm{M}_{\odot}$ (cluster 1) to $4 \times 10^{14}\mathrm{M}_{\odot}$ (cluster 10), and are sorted in descending order of M_{FOF}, as in Table 4.1. Colours show the lensing convergence κ (Top panel: linear scale; Bottom panel: logarithmic scale). Dotted white lines show filaments identified from the noise-free, projected mass distribution, above density thresholds defined in section 4.3.1. For reference, red lines indicate the field of view in which *HST* observations exist for real cluster MS 0451-03.

4.3.2 Distribution of all other mass along a line of sight

In addition to the mass of the galaxy cluster itself, we also account for large-scale structure (LSS) projected by chance along the same line of sight. This is a source of noise in the projected mass of the cluster.

To quantify the expected level of noise, we generate realisations of LSS along 1000 random lines of sight through the BAHAMAS simulation box. We then integrate the 3D mass along the line of sight, weighted by the lensing sensitivity function $\beta(z)$ with $\langle z_s \rangle = 0.97$, interpreting it as a mass distribution in a single lens plane at $z_l = 0.55$. For each realisation of LSS, we calculate an effective radial density profile, $\kappa(R)$. The mean of these realisations is (unsurprisingly) consistent with zero; we also calculate the rms scatter $\sigma_{\rm LSS}$. In concentric annuli of width $\Delta R = 25''$, these are well-fit by

$$\sigma_{\rm LSS} = \frac{A}{\sqrt{R(\text{arcsec})} + B}, \tag{4.3.1}$$

with best-fit values for free parameters

$$A = 0.197 \pm 0.008, \quad B = 6.441 \pm 0.502\,. \tag{4.3.2}$$

We add this in quadrature to the statistical uncertainty on the reconstructed density profiles in Sect. 4.4.3. Note that it would also be possible to compute the full covariance matrix between LSS at different radii or in adjacent pixels of a mass map. Here we use only the diagonal elements, but in Chapter 5, we fit to real observations using the full covariance matrix.

4.3.3 Mock near-IR imaging

To generate a mock catalogue of the cluster galaxies' K-band magnitudes, we run SUBFIND algorithm (Springel et al., 2001) on the particle distribution from the simulations, to identify individual galaxies. We sum their stellar masses, and convert these to K-band luminosity based on the relation presented by Arnouts et al. (2007) for the evolution of stellar mass to light ratio, (M/L_K), with redshift for a sample of quiescent galaxies, and based on the Salpeter (1955) initial mass function. The power-law fitting function is defined as

$$\log_{10}(M/L_K) = a\,z + b, \tag{4.3.3}$$

where the mass M and luminosity L_K are in units of M and L, respectively. The best-fit value for parameters a and b from Arnouts et al. (2007) are

$$a = -0.18 \pm 0.04, \quad b = +0.07 \pm 0.04. \tag{4.3.4}$$

4.3.4 Mock weak-lensing shears

To generate mock weak-lensing observations, we convert the pixellated mass distributions into pixellated shear fields. Since both $\kappa(\boldsymbol{R})$ and $\boldsymbol{\gamma}(\boldsymbol{R})$ fields are linear combinations of second derivatives of $\varphi(\boldsymbol{R})$, it is possible to directly convert between their Fourier transforms $\hat{\kappa}(\boldsymbol{k})$ and $\hat{\boldsymbol{\gamma}}(\boldsymbol{k})$

$$\hat{\gamma_1}(\boldsymbol{k}) = \frac{k_1^2 - k_2^2}{k_1^2 + k_2^2}\,\hat{\kappa}(\boldsymbol{k}) \tag{4.3.5}$$

$$\hat{\gamma_2}(\boldsymbol{k}) = \frac{2k_1 k_2}{k_1^2 + k_2^2}\,\hat{\kappa}(\boldsymbol{k})\,, \tag{4.3.6}$$

where $\boldsymbol{k} = (k_1, k_2)$ is the wave vector conjugate to $\boldsymbol{R}$ (Kaiser & Squires, 1993, hereafter KS93). To implement this in practice, we pixellate the fields within a $34' \times 34'$ (2048×2048 pixel) grid, add zero padding [1] to twice that linear size, then use discrete Fourier transforms. When adding projected LSS to the cluster mass maps (see Sect. 4.3.2), we first sum the two convergence fields then generate $\boldsymbol{\gamma}(\boldsymbol{R})$. Finally, we use Eq. 3.2.17 to calculate $\boldsymbol{g}(\boldsymbol{R})$.

We generate a mock shear catalogue by randomly placing source galaxies throughout the high-resolution pixellated shear field. Mimicking typical single-orbit depth *HST* observations, we sample 50 source galaxies arcmin^{-2}. Note that we achieve a uniform density of background galaxies; in real observations, the number density of background galaxies is both clustered, and dips near the centre of a cluster because of obscuration by, and confusion with, its member galaxies. To each shear value, we

---

[1] In this work, we follow the suggestions in literature (e.g. Merten et al., 2009; Umetsu et al., 2015) and use the zero-padding technique to mitigate boundary effects. However, this technique is not sufficient to perfectly reduce the boundary effects due to discontinuity. A more ideal way would be to adapt the 'inpainting' (Pires et al., 2009) method which extrapolates the input data to the boundary regions.

add Gaussian random noise with width $\sigma_\gamma = 0.36$, representing each galaxy's unknown intrinsic shape, plus uncertainty in shape measurement. This value matches that measured in *HST* measurements near MS 0451-03 (see Chapter 5), and is consistent with that measured for faint galaxies in the *HST* COSMOS field (see figure 17 in Leauthaud et al., 2007).

4.4 Weak Lensing Mass Reconstruction

In this section, we first describe the methods that are used (or suggested) to analyse the distribution of mass in clusters in this work. A common theme will be the suppression of noise — the two main sources of which are projected LSS, and galaxies' intrinsic shapes. In particular, sophisticated nonlinear noise-suppression techniques have been developed to map the 2D distribution of mass. Even for measurements that could be obtained directly from the shear field, it may therefore be efficient to first infer (and suppress noise in) a mass map, then to measure equivalent quantities from that. We will then show the reconstructed mass maps and the recovered density profiles.

4.4.1 Methods

Direct inversion with KS93+MRlens

As mentioned in Section 3.4.1, convergence κ can be directly converted from weak lensing shear γ, using equation 3.4.38. This is a non-local mapping. In observations of the real Universe, any missing shear values (e.g. outside the survey boundary or behind bright stars) must be replaced via 'inpainting' (Pires et al., 2009; Raghunathan et al., 2019) to avoid suppressing the convergence signal inferred nearby. We avoid this effect by using mock shear catalogue that is contiguous and covers a larger area ($34' \times 34'$) than the mosaicked *HST* imaging of MS 0451-03. We bin the shear field $\boldsymbol{\gamma}(\boldsymbol{R})$ into $0.4'$ pixels, add zero padding out to $105' \times 105'$, and implement equation (3.4.38) using discrete Fourier transforms.

Noise was suppressed in early incarnations of KS93 by convolving the mass distribution with a larger smoothing kernel, such as Gaussian filter. We instead filter the final convergence map using the software, MRLens (see Section 3.4.1). The software implementation has various free parameters: we use ten iterations during the filtering process, and decompose the noisy 2D convergence map into six wavelet scales, starting at $j = 3$ (highest resolution). The starlet wavelets (eq (11) of Leonard et al., 2012) have size $\vartheta = 2^j$ pixels. For comparison to older analyses, we also repeat the analysis after smoothing and rebinning the shear field into larger, $1'$ pixels.

Forward fitting with Lenstool

We also use Lenstool[2] (Jullo & Kneib, 2009) to fit the reduced shear catalogues $\boldsymbol{g}(\boldsymbol{R})$ with a sum of analytic mass distributions. The field of view considered is the same size as the mosaicked *HST* imaging around MS 0451-03. Jullo & Kneib (2009) advocate a mass model built of three components.

- **Cluster-scale halo:** For clusters that produce strong gravitational lensing, the observed positions of multiple images are typically used to pre-fit the smooth, large-scale distribution of mass (Jauzac et al., 2015b; Kneib et al., 1996; Richard et al., 2011; Smith et al., 2005). Like many clusters, our mock data do not include strong-lensing, so we omit this component. Note that our performance forecasts will therefore be conservative, because this information efficiently captures the broad features of a mass distribution in only a few parameters, and removes degeneracies between the remaining parameters that we shall fit (Jauzac et al., 2015a).

- **Cluster member galaxies:** We model the total mass of each galaxy in the cluster as a tPIEMD (Eq. 2.2.30). Following Jauzac et al. (2012), their core radii, truncation radii and velocity dispersions are scaled using empirical relations

$$r_{\mathrm{c}} = r_{\mathrm{c}}^{*}\left(\frac{L}{L^{*}}\right)^{\frac{1}{2}}, \quad r_{\mathrm{t}} = r_{\mathrm{t}}^{*}\left(\frac{L}{L^{*}}\right)^{\frac{1}{2}}, \sigma = \sigma^{*}\left(\frac{L}{L^{*}}\right)^{\frac{1}{4}}, \tag{4.4.7}$$

---

[2] We implement Lenstool using version 7.1 of the software available from `https://projets.lam.fr/projects/Lenstool/wiki`.

where $r_{\rm c} = 0.15$kpc, $r_{\rm t} = 58$kpc and $\sigma^* = 163.10$kms^{-1} for a typical galaxy with K-band magnitude $m^* = 18.699$ at $z = 0.55$. These scaling relations describe early-type cluster galaxies (Wuyts et al., 2004), and assume a constant mass-to-light ratio for all cluster members.

- **Multi-scale, free-form grid:** We add a free-form (pixellated) mass distribution with spatially-varying resolution that is adapted to the cluster's light distribution. This method is described in Section 3.4.2 in details. After six levels of recursion, the brightest parts of the cluster are covered by the highest resolution grid with $r_c = 18''$. We extend this grid into the cluster centre, which is inevitably modelled at the highest resolution. At the centre of every triangle, we place a circular ($q = 1$) tPIEMD (Eq. 2.2.30) to represent its potentials. This process represents a prior of light-traces-mass. Therefore, multi-scale grid is well suited to describe irregular mass distributions, such as large-scale structures.

4.4.2 Results of 2D mass maps

We quantify the precision and accuracy of mass maps produced by KS93+MRLens (figure 4.2) and Lenstool (figure 4.3) by comparing them to the noise-free distributions of mass, $\kappa_{\rm true}$ (which includes only the mass of the cluster, not projected LSS). We first measure deviations from this truth, $\kappa_{\rm res} \equiv \kappa - \kappa_{\rm true}$, to obtain the residual maps. For each map, we compute the noise level σ_κ, defined as the root mean square (rms) deviation from the mean of $\kappa_{\rm res}$, over all pixels in a field of view equivalent in size to the *HST* observations of MS 0451-03. We then average the performance of each method over all 10 clusters (table 4.2).

In observations of the real Universe, σ_κ cannot be calculated because there is no privileged knowledge of $\kappa_{\rm true}$. For comparison with observations, we therefore also measure $\sigma_\kappa^{\rm obs}$, the rms deviation from the mean of κ. We find values of $\sigma_\kappa^{\rm obs}$ roughly consistent with σ_κ being added in quadrature to an irreducible component that is the rms deviation from the mean of $\kappa_{\rm true}$, 0.022 ± 0.0007 on average (0.027 for the five highest mass clusters, or 0.017 for the five lowest).

Direct inversion mass reconstruction

MRLens suppresses galaxy shape noise by a factor 3.8 (a factor 1.5 better than smoothing with $1'$ pixels, and retaining higher spatial resolution). However, galaxy shapes still contribute more noise to the mass maps than (physically real) LSS noise. Spurious noise peaks are found in all regions of the field of view. Massive substructures with $\kappa > 0.096$ can be detected with $\mathrm{S/N} > 3$.

Mass reconstructions using KS93+MRLens are statistically consistent with being unbiased. Both positive and negative noise fluctuations are produced, at all radii. The mean residual of maps with both sources of noise is $\langle \kappa - \kappa_{\mathrm{true}} \rangle = -0.0005 \pm 0.0018$, where the averaging is over 10 clusters, and the uncertainty is the standard deviation between them. The marginally negative mean may be because density is underestimated in a small region near cluster cores (see Sect. 4.4.3).

Forward-fitting mass reconstruction

Lenstool suppresses noise even further. Galaxy shape noise is an additional factor 2 lower than KS93+MRLens (averaged across the field of view) — and LSS noise becomes the dominant component.

The spatial distribution of noise is nonuniform. A Lenstool reconstruction has more freedom in regions with a high resolution free-form grid (section 3.4.2), such as the cluster core and associated substructures. Spurious κ peaks appear preferentially in those regions, even when we replace the shear catalogue with one that contains only (spatially uniform) galaxy shape noise. To further investigate this effect, we split the ten clusters into two subsamples: higher mass (clusters 1 to 5), and lower mass (clusters 6 to 10). Multi-scale grids of the high mass sample have larger high-resolution regions, resulting in noisier maps on average. Assessing the S/N of any identified peak must therefore involve bootstrap analysis at the specific region of interest. This confirms Jullo et al. (2014b)'s similar assessment of the performance of Lenstool. For many scientific purposes, spatially varying noise is a useful feature: the lower resolution and positive definite constraints help to suppress positive LSS noise and remove negative noise at large radii. Even filaments contain a statistically significant overdensity of galaxies (Galárraga-Espinosa et al.,

	σ_κ			$\sigma_\kappa^{\mathrm{obs}}$		
	Full mock	**Shape noise only**	**LSS noise only**	**Full mock**	**Shape noise only**	**LSS noise only**
KS93 (pixel scale $0.4'$)	0.088 ± 0.001	0.091 ± 0.001	0.017 ± 0.002	0.090 ± 0.002	0.092 ± 0.001	0.027 ± 0.006
KS93 (pixel scale $1'$)	0.037 ± 0.001	0.037 ± 0.001	0.013 ± 0.002	0.042 ± 0.003	0.039 ± 0.002	0.024 ± 0.006
KS93+MRLens	0.026 ± 0.001	0.024 ± 0.001	0.014 ± 0.002	0.032 ± 0.004	0.028 ± 0.002	0.024 ± 0.006
High mass clusters	0.026 ± 0.001	0.024 ± 0.001	0.016 ± 0.001	0.035 ± 0.004	0.030 ± 0.002	0.029 ± 0.005
Low mass clusters	0.026 ± 0.002	0.024 ± 0.001	0.012 ± 0.001	0.029 ± 0.003	0.026 ± 0.001	0.019 ± 0.003
Lenstool	0.015 ± 0.004	0.012 ± 0.003	0.013 ± 0.004	0.023 ± 0.007	0.022 ± 0.007	0.024 ± 0.008
High mass clusters	0.018 ± 0.002	0.014 ± 0.002	0.016 ± 0.003	0.031 ± 0.005	0.030 ± 0.005	0.030 ± 0.008
Low mass clusters	0.012 ± 0.002	0.010 ± 0.002	0.010 ± 0.001	0.018 ± 0.003	0.018 ± 0.003	0.019 ± 0.004

Table 4.2: Noise level in mass maps created using different methods, measured as the standard deviation of all pixels inside a field of view equivalent to *HST* observations of MS 0451-03. Central values and uncertainties show the mean and standard deviation between clusters. The first three columns show deviations from the true, noise-free mass map; the second three columns show deviations from zero — which can be compared to observations of the real Universe. The 2^{nd}, 3^{rd}, 5^{th} and 6^{th} columns refer to analyses in which the shear catalogues contain only certain sources of noise, so their relative effect can be assessed. The first two rows quantify the performance of KS93 direct inversion, with noise suppressed only via convolution with a top hat window function of radius R_f. The middle rows use $R_f = 0.2'$, and suppress noise using MRLens. The bottom rows use Lenstool.

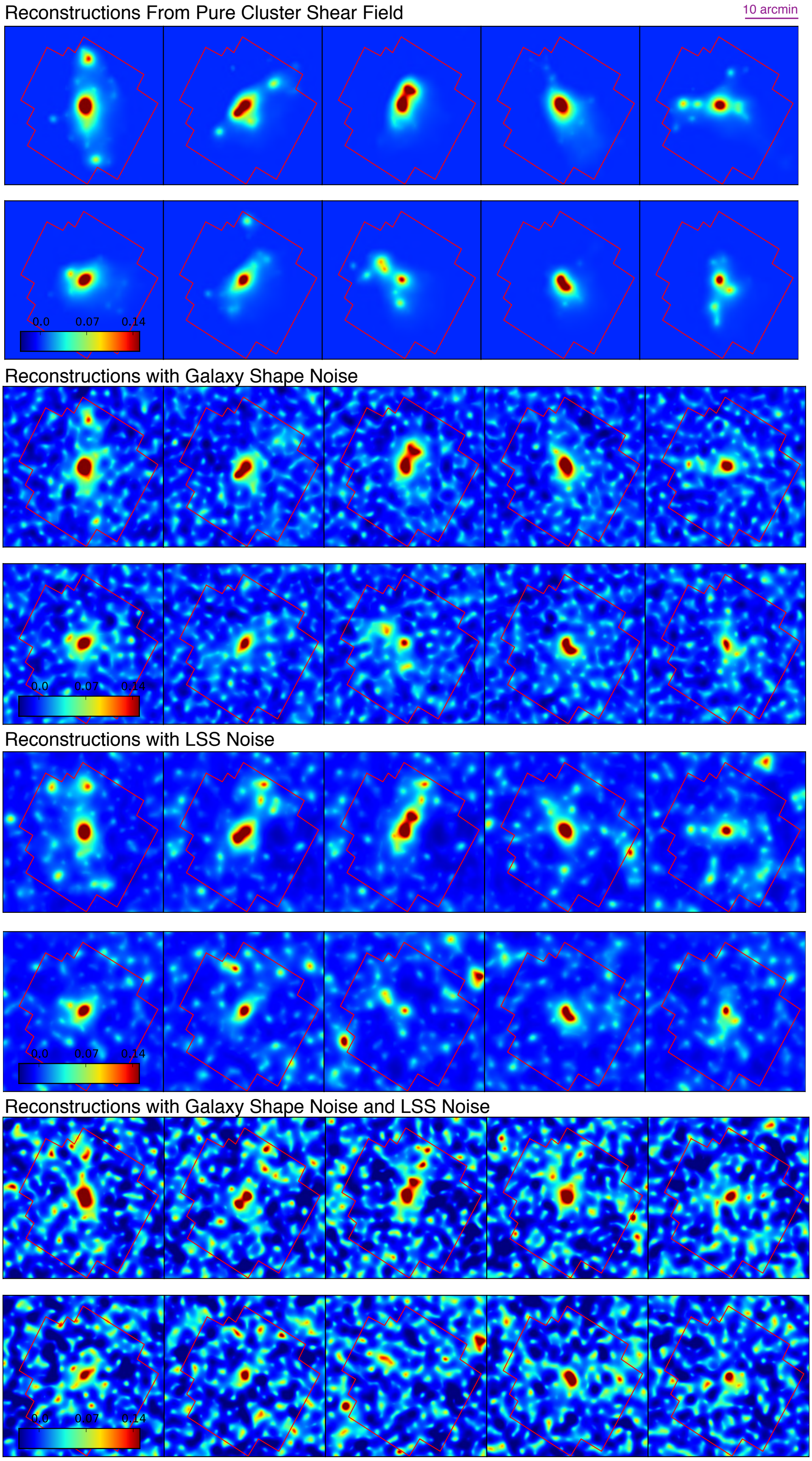


Figure 4.2: *(see next page)*

Figure 4.2: Projected mass maps of the 10 simulated clusters reconstructed using the KS93+MRLens direct inversion method, including different components of noise. *Top panels:* reconstruction with no noise. *Second panels:* including only shape noise from 50 background galaxies per square arcminute. *Third panels:* including only projected large-scale structure. *Bottom panels:* including both sources of noise simultaneously. Colour scales are identical for all panels. For reference, red lines indicate the field of view of the largest *HST* mosaic obtained around a massive galaxy cluster, MS 0451-03.

2020), so the reconstruction can be given sufficient flexibility to include (rather than suppress) them.

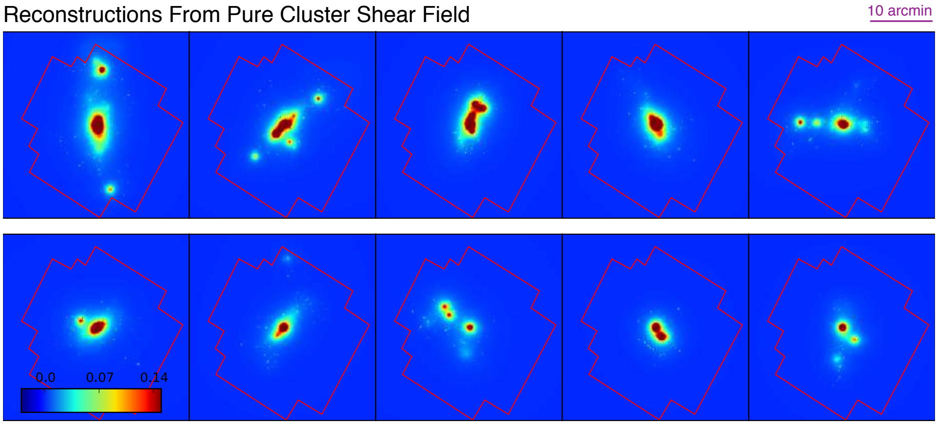
Mass reconstructions using Lenstool slightly overestimate the total mass, because of its positive-definite constraint. Averaged over the field of view, the mean residual of maps with both sources of noise is $\langle\kappa_{\rm res}\rangle = 0.0088 \pm 0.0064$ (we quote the mean and standard deviation of $\kappa_{\rm res}$ for 10 clusters).

4.4.3 Radial density profiles

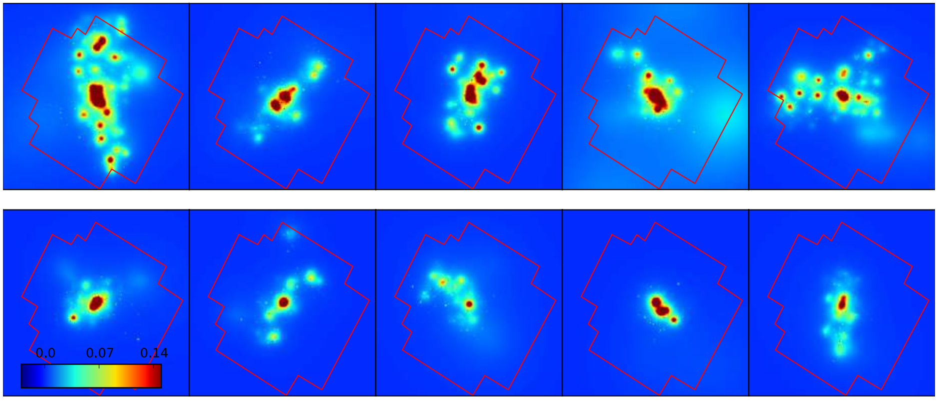
Most analyses of galaxy clusters involve fitting models of an azimuthally-averaged density profile. Measuring density profiles is a key test of e.g. cosmological structure formation (the 'splashback' feature reveals a characteristic build-up of accreted mass, pausing at first apocentre after first core passage Diemer & Kravtsov, 2014) and the nature of dark matter (Newman et al., 2013, 2015; Robertson et al., 2019). Because almost all clusters have irregular features, and approximately half are significantly unrelaxed (Smith et al., 2010), it is necessary to statistically combine the profiles of many clusters. This can be achieved by rescaling and averaging their density profiles in radial bins, or by fitting parametric models with radial (or elliptical) symmetry, then averaging the best-fit parameters.

We calculate the radial density profiles of each simulated cluster by azimuthally averaging the reconstructed density maps within linearly spaced annuli of fixed width $\Delta R = 25''$. For Lenstool reconstructions, we quote the statistical uncertainty in

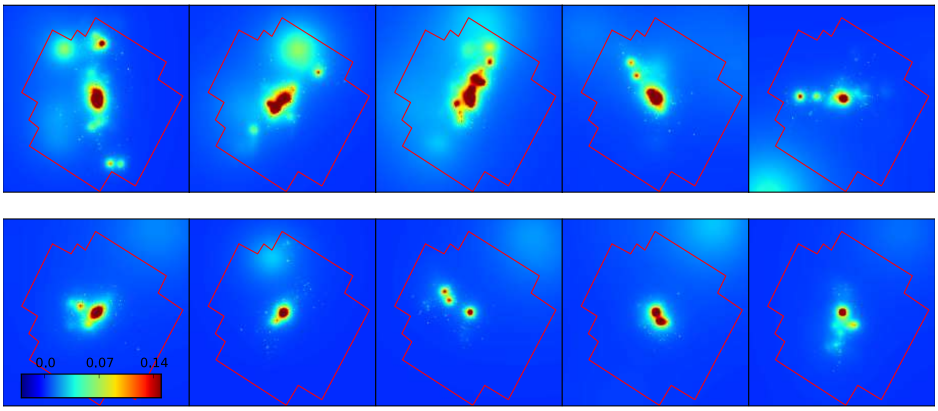
Reconstructions From Pure Cluster Shear Field



Reconstructions with Galaxy Shape Noise



Reconstructions with LSS Noise



Reconstructions with Galaxy Shape Noise and LSS Noise

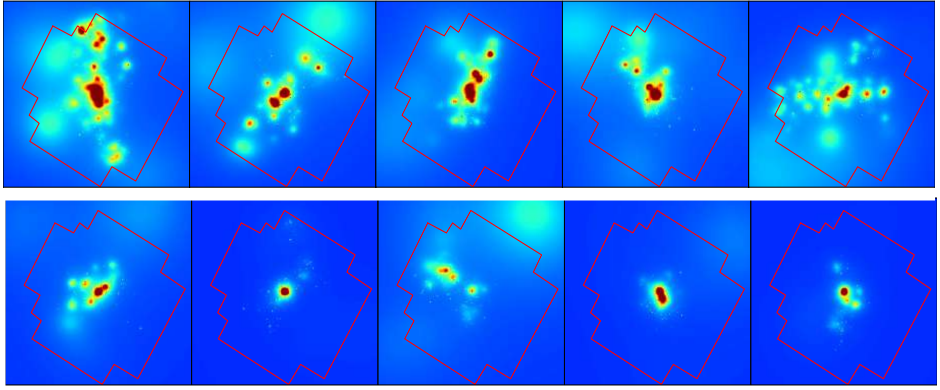


Figure 4.3: Same as figure 4.2, but reconstructed using Lenstool.

each annulus, σ_{stat}, determined during the MCMC sampling. When the signal from projected LSS is included, we add σ_{LSS}, as detailed in Sect. 4.3.2, such that the total uncertainty error on the density profile, $\sigma^2_{\text{tot}} = \sigma^2_{\text{stat}} + \sigma^2_{\text{LSS}}$. Figure 4.4 shows the clusters' density profiles recovered by azimuthally averaging the convergence maps. The smoothing inherent to KS93+MRLens results in an underestimation of density in the cluster core, and an overestimate just outside. This biases the inner profile slope that is often used to distinguish between cusps and cores. Lenstool is accurate in the cluster core, because its basis functions have a density profile that matches those of the simulated clusters. This is not affected by Lenstool's positive-definite constraint, because the true mass distribution is very positive near the core. In the cluster outskirts, Lenstool strongly suppresses galaxy shape noise, and the reconstruction is dominated by LSS noise. Because of the positive-definite constraint, this is also potentially biased. The amplitude of LSS noise varies a great deal depending on environments along the line-of-sight LSS, but we typically find artificial boosts in inferred density of up to $\sigma_{\text{LSS}} = 4 \times 10^7\,\text{M}_\odot\,/\,\text{kpc}^2$, at large projected radii, $R > 1000\,\text{kpc}$. This effect must be taken into account when measuring properties at large radius (e.g. M_{200}, c_{200}, splashback radius). To militate against this, measurements of galaxy redshifts will be invaluable to disentangle structures connected to the cluster from those lying in the foreground or background.

4.5 Halo Shape Measurement

On large scales, the accretion of matter from the surrounding large-scale environment plays a key role in determining the shape and orientation of cluster dark matter halos (Shaw et al., 2006). Halos are not necessarily self-similar (concentric ellipsoids with the same orientation and ellipticity; Suto et al., 2016), but align with the infall direction of subhalos and surrounding filaments at large radii. Thus, the shape of galaxy clusters is a fundamental probe of the history of its mass accretion. Numerical simulations with collisionless dark matter predict cluster halos to be triaxial (Jing & Suto, 2002; Warren et al., 1992). Allowing DM particles to self-interact isotropizes the orbits of dark matter particles, and makes the inner mass distribution more

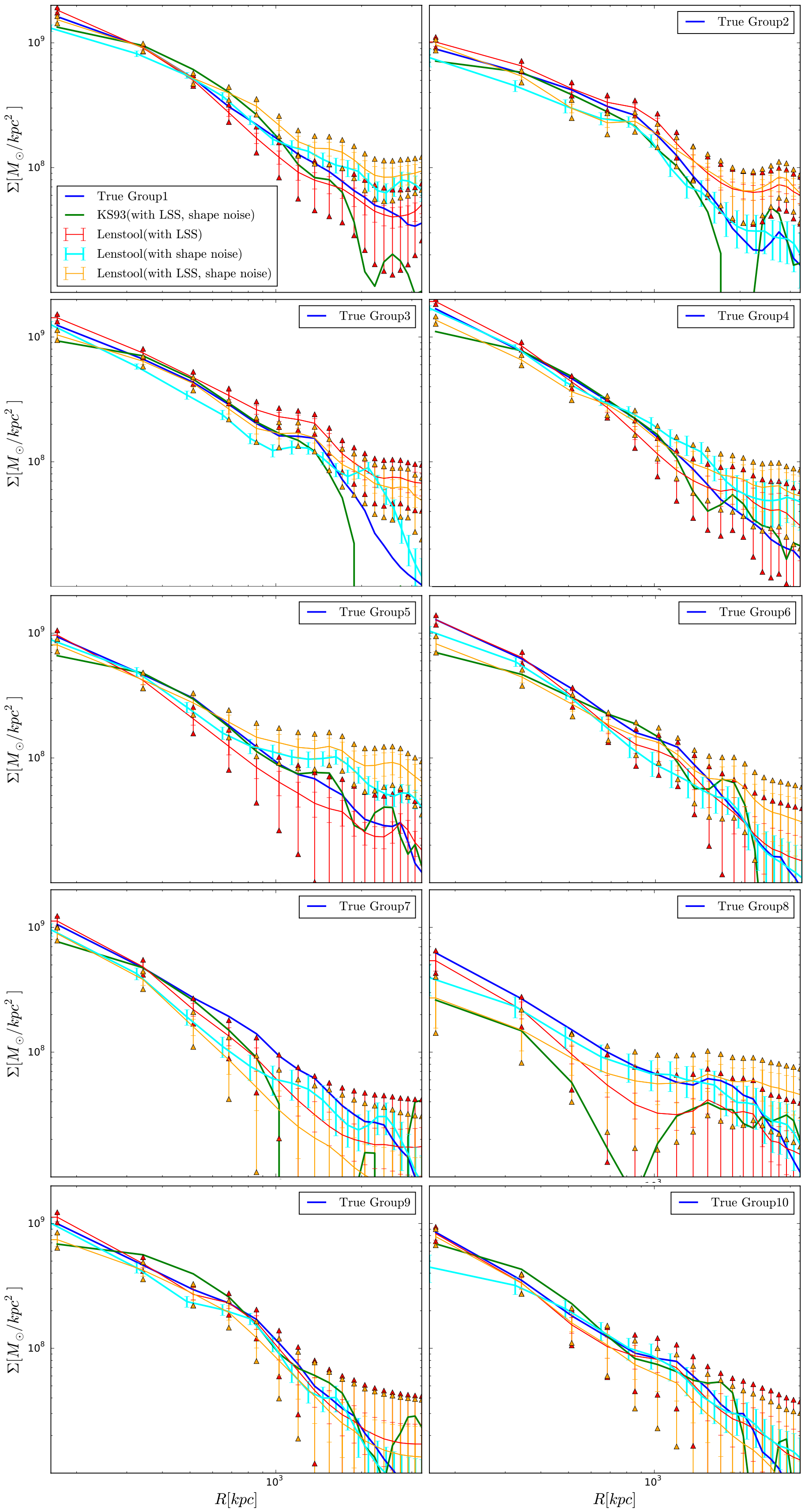


Figure 4.4: *(see next page)*

Figure 4.4: Surface mass density profiles for 10 simulated clusters. Blue solid lines show the the density profile calculated from the true mass distribution in Fig 4.1. Green solid lines are the density profiles of KS93+MRLens reconstructed maps after adding shapes noise and LSS. Cyan, orange, and red lines show the results recovered by Lenstool including shape noise, projected LSS, and both shape noise and LSS, respectively. Error bars with line caps are statistical errors from the MCMC sample. Error bars with triangle caps are total errors which is the combination of statistical errors with the estimated noise from the projected LSS.

spherical. For a cross-section of $1\,\mathrm{cm}^2/\mathrm{g}$, the median minor-to-major axis ratio 100 kpc from the halo centre is ∼0.8, compared with ∼0.5 with CDM (Robertson et al., 2019).

4.5.1 Ellipcial NFW mass modelling

We first modify the spherical NFW model to be elliptical by a coordinate transformation

$$|\boldsymbol{R}'|^2 = q(x^2\cos^2\phi + y^2\sin^2\phi) + (y^2\cos^2\phi - x^2\sin^2\phi)/q\ , \tag{4.5.8}$$

(Kassiola & Kovner, 1993a; Oguri et al., 2010) that maps a circle to an ellipse with axis ratio $0 < q \leq 1$ and orientation ϕ. We apply this transformation to the projected mass distribution. Applying it instead to the gravitational potential yields different results, and no simple mapping exists between them.

We then fit the elliptical NFW mass distribution to the 2D convergence maps reconstructed from shear catalogues with no noise, with shape noise, with LSS noise or both. The fit[3] minimizes the sum of the squared difference between the reconstructed surface mass density of each BAHAMAS simulated cluster and an elliptical NFW model, within a circular aperture of radius R_{ap}. We then vary R_{ap}, to investigate changes between the cluster's inner and outer halos. During the fits, we

---

[3] We use the scipy.minimize implementation of the L-BFGS-B algorithm (Byrd et al., 1995), available from `https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.minimize.html`.

fix the centre of the NFW (to the location of the most bound particle) because it is degenerate with axis ratio. We adopt flat priors on other free parameters: $0.1 \leq M_{200}$ $(10^{15}\mathrm{M}_{\odot}) \leq 5$, $0.1 \leq c_{200} \leq 8$, $0 \leq \phi \leq 180$ and $0.1 \leq q \leq 0.9$, and neglect covariance between adjacent pixels. So that the uncertainties in this test match those in observational data, we add only one, fixed realisation of LSS along the line-of-sight associated with each cluster.

4.5.2 Results

Both mass reconstruction methods produce distributions that are rounder than the truth (figure 4.5). eNFW models fitted to the reconstructed mass maps (figure 4.2) have a higher mean axis ratio $\langle q \rangle$ than models fitted to the true mass maps (figure 4.1). However, they successfully capture the decrease in $\langle q \rangle(R)$ at large radii that is seen in the true mass maps (reflecting a transition from dominant baryonic effects to the infall of structures along filaments; Suto et al., 2017). The orientation of most inner ($R = 650\,\mathrm{kpc}$) and outer ($R = 3\,\mathrm{Mpc}$) halos also remain aligned within $\Delta\phi \leq 10°$, matching the true distributions (and also the simulations by Despali et al., 2017). Two exceptions to this are clusters 5 and 9, which have complex cores and $\Delta\phi = 17°$ and $\Delta\phi = 15°$. This likely indicates a transitory state during a major merger.

Using KS93+MRLens leads to inferred values of $\langle q \rangle$ that are too high by about 6%. The level of bias is not significantly influenced by either source of noise in the shear catalogue (although adding noise increases scatter in individual measurements of q as expected). It is likely due to the isotropic blurring associated with pixellisation and MRLens filtering.

Using Lenstool leads to inferred values of $\langle q \rangle$ that are too high by 10% in the cluster core and 15% in the outskirts. The bias appears to be caused by two effects:

- The mass distribution is built from components that are all individually spherical. If the dominant halo in the cluster core is anomalously spherical (see clusters 4, 5, 8 or 10 in figure 4.6), it can bias the apparent axis ratio of the mass inside an aperture by up to 10%, almost regardless of the size R_{out} of that aperture.

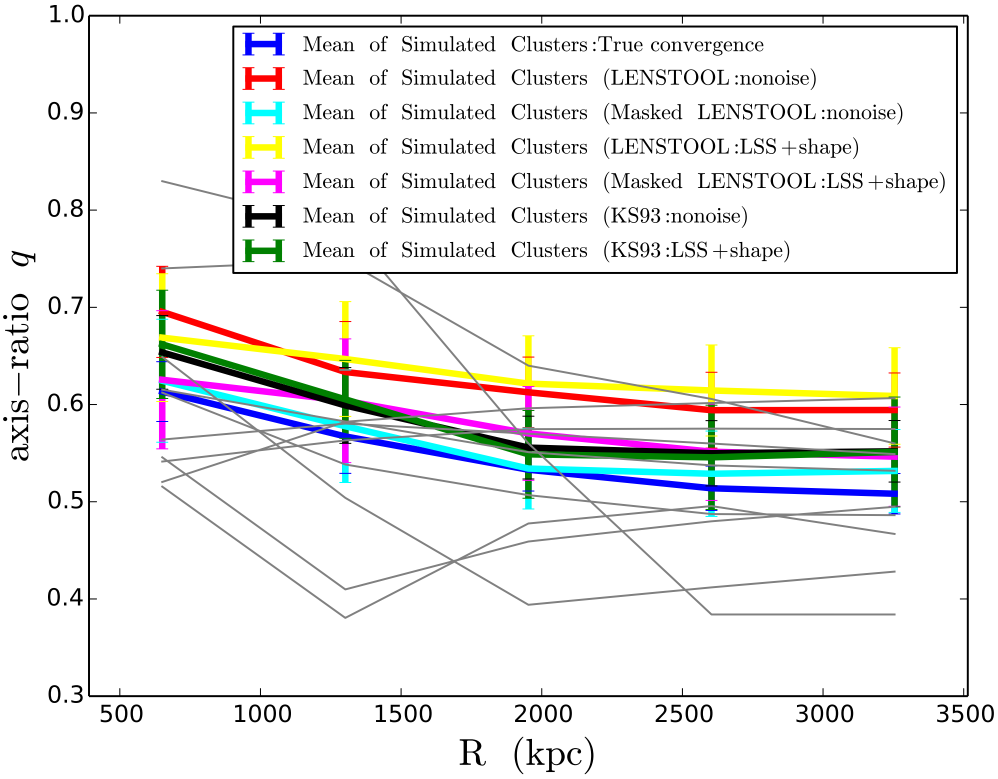


Figure 4.5: Best-fit axis ratios of the mass distribution in galaxy clusters, as a function of projected, clustercentric radius R. Grey lines show the BAHAMAS simulated clusters, whose axis ratio profiles are measured from the true mass distribution. Blue lines show the mean and standard deviations from this set of clusters. Black (green) lines show the mean axis-ratio and its scatter measured from noise-free KS93+MRLens reconstruction (with LSS and shape noise). Red (yellow) lines show the mean results measured from noise-free Lenstool reconstruction (with LSS and shape noise). Cyan (magenta) lines show the axis-ratio measured from the masked $R < 35''$ (228 kpc) Lenstool reconstruction (with LSS and shape noise).

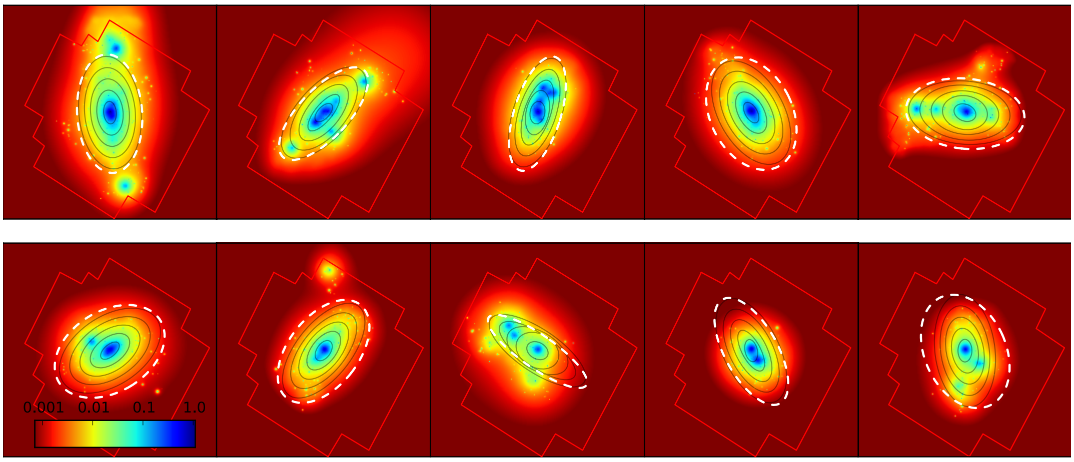


Figure 4.6: Elliptical NFW models fitted to the LENSTOOL mass maps are $\sim 6\%$ too round, on average (see figure 4.5). Black ellipses have the same axis ratio of the true mass distribution (see figure 4.1) inside annulus $R < R_{\rm ap}$, where different values of $R_{\rm ap}$ are indicated by the length of the major axis. White dashed ellipses show the axis ratio measured from masked Lenstool reconstructions, inside the largest $35'' < R < R_{\rm ap}$. The background image shows the mass distribution reconstructed by LENSTOOL, as in figure 4.3 but with a logarithmic scale to highlight one problem with the LENSTOOL method: overly circular central halos.

- The mass distribution is constrained to be positive definite. In the absence of noise, this has no effect. If we add galaxy shape noise, it is also relevant that the reconstructed mass distribution is higher resolution (has more freedom) along its major axis. The positive-definite bias in noise artefacts then exaggerates the major axis, reducing $\langle q \rangle$ by $\sim$5%. If we add LSS noise, $\langle q \rangle$ increases by 8% because there is a larger area at close to zero convergence along the minor axis.

It is possible to mitigate the first effect by masking the cluster core. We successfully recover the true axis ratio when fitting an eNFW using to *noise-free* data inside an annulus $35'' < R < R_{\rm ap}$ (instead of a circle of radius $R_{\rm ap}$). Fitting inside annuli also decorrelates measurements of $\langle q \rangle$ at different radii, and steepens the apparent gradient in $\langle q \rangle(R)$. Note that the second effect still increases $\langle q \rangle$ by $\sim$6% in the presence of both sources of noise.

A different strategy to mitigate sphericity bias could be to pre-fit the axis ratio of central halos, then hold them fixed while the rest of the grid is constrained. A similar two-step process happens naturally in most combined analyses of strong plus weak lensing, where strong lensing information constrains a cluster core. This bias should therefore not affect LENSTOOL strong lensing analyses. However, it would be difficult to characterise statistical uncertainty in such analysis, because shear data would be used twice.

4.5.3 Comparison with previous studies

Previous work by simulators to measure the shape of cluster-scale halos split into two distinct conclusions. Hopkins et al. (2005) found that 2D cluster ellipticity increases with clustercentric radius, in agreement with our results. However, they also found that the ellipticity is $\epsilon \approx 0.05z + 0.33$ for the redshift range $0 < z < 3$, which implies $q = 0.64$ at the $z = 0.55$ redshift of our simulated clusters. Similarly, Ho et al. (2006) found $q \sim 0.616$ for halos with masses $M > 10^{14}{\rm M}_\odot$ at $z = 0.55$ assuming $\Omega_m = 0.3$, and $\sigma_8 = 0.7$, and little dependence upon cosmological model. Both of these results are slightly rounder than our measurement of $\langle q \rangle_{\rm true} \sim 0.55 \pm 0.03$.

More recently, Despali et al. (2017) found that $M \sim 10^{15}\mathrm{M}_\odot/h$ halos in the SBARBINE N-body simulations had more elliptical shapes, with $q \sim 0.55$. Suto et al. (2016) studied the probability distribution function (PDF) of q from projected density distributions without assumptions of self-similarity. Using their PDF fit formula for M_{vir} at $z = 0.4$, we obtain $q = 0.57 \pm 0.17$. These results match ours closely, and more recent independent analyses appear to be converging. Note that the other simulations were DM-only, but Suto et al. (2017) found that non-sphericity is unaffected by baryonic physics beyond half of the virial radius, so it is reasonable to compare to our measurements.

Several observational studies of weak-lensing have attempted to measure cluster halo ellipticity. In the Sloan Digital Sky Survey (SDSS), Evans & Bridle (2009) found a mean projected axis ratio $\langle q \rangle = 0.48^{+0.14}_{-0.09}$ in the redshift range $0.1 < z < 0.3$. By directly fitting 2D shear-maps with eNFW models, Oguri et al. (2010) measured a mean projected axis ratio $\langle q \rangle = 0.54 \pm 0.04$ for a sample of 18 X-ray luminous clusters in the redshift range $0.15 < z < 0.3$. Shin et al. (2018) measured $\langle q \rangle = 0.56 \pm 0.09$ for 10,428 SDSS clusters. These results are consistent with our measurement. Intriguingly, Umetsu et al. (2018) measured the median projected axis-ratio of 20 high-mass galaxy clusters in the HST-CLASH survey to be $\langle q \rangle = 0.67 \pm 0.07$, within a scale of $2\,\mathrm{Mpc}h^{-1}$. However, their measurement from the CLASH high-magnification subsample was $\langle q \rangle = 0.55 \pm 0.11$, consistent with our results. This suggests a lensing selection bias towards halos that are more elliptical (in the plane of the sky as well as along a line of sight). In contrast, X-ray selected clusters tend to be relaxed clusters with rounder dark matter halo shapes. For clusters selected by the red sequence technique, it is more likely that they are elongated along the line of sight, causing an over-density of red galaxies in the projected sky-plane. Since our simulated cluster sample is selected by their high mass, with each halo projected along a random line-of-sight, we can only give the mass-selected mean halo shape. For direct comparison with observational data, future theoretical predictions will need to take the selection function of the observed sample into effect.

Other shape measurement techniques are possible. Studies using quadrupole estimators to quantify halo shape include Adhikari et al. (2015); Clampitt & Jain (2016); Shin et al. (2018); van Uitert et al. (2017). In particular, Clampitt & Jain (2016) developed a new estimator to measure the quadrupole weak-lensing signal from 70,000 SDSS Luminous Red Galaxies halos, and found a best-fit axis-ratio $\langle q \rangle \sim 0.78$. Their analysis assumes that dark matter perfectly aligns with light, so one potential systematic in their study is the possibility of light and dark matter misalignment. The determination of the orientation of each lens-source pair could become inaccurate due to this misalignment, and result in the dilution of the final stacked signal of the halo ellipticity. Indeed, applying the misalignment distribution of Okumura et al. (2009) to their measurement, they obtain $q \sim 0.6$, consistent with our results.

4.6 Searches for filaments

Dark matter and gas are accreted onto a cluster mainly through filaments that connect it to the 'cosmic web'. Filaments are key transition regions in the evolution of galaxy morphology (Einasto et al., 2007; Kuutma et al., 2017; Liu et al., 2019; Martizzi et al., 2019; Nuza et al., 2014; Pandey & Bharadwaj, 2006) and star formation (Alpaslan et al., 2015, 2016; Crain et al., 2009; White et al., 2010; Yuan et al., 2019).

Filaments are much lower density environments than a cluster, so appear in gravitational lensing observations with correspondingly lower signal-to-noise. While it is possible to search for filaments directly in shear data (Dietrich et al., 2005; Dietrich et al., 2012; Jauzac et al., 2012), we explore whether it is efficient to leverage the de-noising techniques developed for mass mapping, then to analyse the inferred convergence field. In this section, we present a new method that is optimal for filament search in the reconstructed field.

4.6.1 Removing the smooth mass component

First, we subtract the smooth distribution of mass in the clusters, which would otherwise dominate the lower density contrast in the filaments.

We fit mock reduced shear data (with or without LSS and galaxy shape noise), using an elliptical NFW potential. This model has 6 free parameters: the coordinates of the centre of mass, $(x_\mathrm{c},\, y_\mathrm{c})$, the ellipticity, $e = (1-q^2)/(1+q^2)$ where q is the axis ratio, the position angle, ϕ, the scale radius, r_s, and the concentration, c. We set flat priors on x_c and y_c within a $15'' \times 15''$ box centred on the most bound particle, and flat priors on $e \in [0.05, 0.7]$, $\phi \in [0, 180]$, $r_s \in [50, 1000]\,\mathrm{kpc}$, and $c \in [0.5, 10]$. Note that we introduce ellipticity to this model via a coordinate transformation to the gravitational potential (rather than the mass, as in Sect. 4.5.1) because code to achieve this already exists within LENSTOOL[4]. The smooth distribution of mass in most simulated clusters is well approximated by a single potential. However, we use two to fit bimodal clusters 1, 2 and 9, and three for cluster 3.

We then subtract the best-fit smooth halos from the convergence maps. Since the mass distribution of simulated clusters cannot be perfectly described by elliptical NFW potentials, small residuals are left near the cluster centre. Such residuals do not impact searches for filaments at much larger radii.

4.6.2 Aperture multipole moments

Schneider & Bartelmann (1997) first suggested looking for substructures or filaments using multipole moments of a convergence field within circular apertures. These are

$$Q_n(\boldsymbol{R}) = \int_0^\infty |\boldsymbol{R}' - \boldsymbol{R}|^n \, \mathrm{e}^{ni\phi} \; U_n(|\boldsymbol{R}' - \boldsymbol{R}|) \; \kappa(\boldsymbol{R}') \; \mathrm{d}^2\boldsymbol{R}' \,, \tag{4.6.9}$$

where n is the order of the multipole, $(R,\, \phi)$ are polar coordinates, and $U_n(R)$ is a radially symmetric weight function, for which Dietrich et al. (2005) suggested

$$U_n(R) = \begin{cases} 1 - \left(\frac{R}{R_{\mathrm{max},n}}\right)^2 & \text{for } R \leqslant R_{\mathrm{max},n} \\ 0 & \text{otherwise.} \end{cases} \tag{4.6.10}$$

Eq. (4.6.9) can also be expressed in terms of shear measurements, which Dietrich et al. (2005) used to detect filament candidates in close pairs of clusters. Since

---

[4] An elliptical gravitational potential produces a ‘boxy’ mass distribution if $e > 0.6$. However, for the low values of ellipticity that we obtain, the maximum distance δR between a projected density contour and a true ellipse is $\delta R/R < 10\%$ (see figure 6 in Golse & Kneib, 2002).

modern mass reconstruction methods successfully suppress noise, we attempt instead to measure multiple moments directly from the pixellated convergence field

$$Q_n(\boldsymbol{R}) = A_{\rm pix} \sum_{i=1}^{N_{\rm pix}} R_i^n \; \mathrm{e}^{ni\phi_i} \; U_n(R_i) \; \kappa(\boldsymbol{R_i}) \,, \tag{4.6.11}$$

where $N_{\rm pix}$ is the total number of pixels inside the aperture and $A_{\rm pix}$ is an area per pixel. For $n > 0$, Q_n is complex; we shall generally take its modulus, $|Q_n|$.

Multipoles of different orders highlight different features in a mass distribution (see figure 4.7). Monopole moments ($n = 0$) are the aperture mass or normalisation. Dipole moments ($n = 1$) are the local gradient of a convergence field. They form ring-like structures around mass clumps. Quadrupole moments ($n = 2$) are the locally-weighted curvature or Hessian of the convergence field. As Dietrich et al. (2005) explain using a toy model, linear overdensities with a lower mass on either side (i.e. filaments) have large quadrupole moments. However, regions *between* two substructures also have large quadrupole moments. To identify the former and suppress the latter, Mead et al. (2010) suggested combining multipole moments

$$Q \equiv \alpha_0 \, Q_0 + \alpha_1 \, Q_+ 1 + \alpha_2 \, Q_2 + ... \tag{4.6.12}$$

where the constants, α_i, can be adjusted to boost a signal of interest. We have tried different combinations and aperture sizes, and find that a choice of

$$\alpha_0 = -\alpha_1 = 0.7 \quad \text{and} \quad \alpha_2 = 1 \,, \tag{4.6.13}$$

$$R_{\rm max,0} = 1' \quad \text{and} \quad R_{\rm max,1} = R_{\rm max,2} = 2' \,. \tag{4.6.14}$$

typically highlights narrow filaments (see figure 4.8). The quadrupole term is sensitive to linearly extended mass distributions, and the rings that it adds around substructures are removed by the negative dipole term. The monopole term fills in the subtracted mass, and suppresses regions between two substructures but without mass.

4.6.3 Filament identification

To identify individual filaments, we search for spatially extended regions with Q above a threshold $Q_{\rm threshold}$. The normalisation of coefficients in eq. (4.6.13) conveniently ensures that regions inside a contour $Q_{\rm threshold}$ have mean convergence

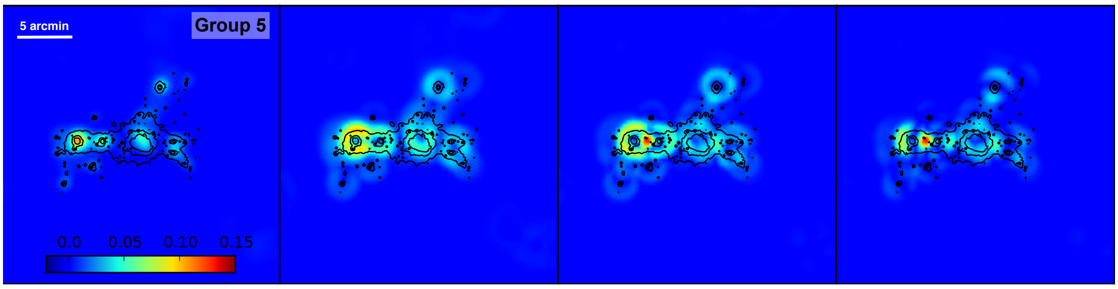


Figure 4.7: An example of aperture multipole moments of various orders, which pick out different features of the noise-free mass distribution of one simulated cluster (Cluster 5, which happens to have several features in the plane of the sky). Moments are calculated after subtracting the large-scale smooth mass distribution. From left to right, panels show: (a) monopole, (b) dipole, (c) quadrupole moments and (d) the radial component of the quadrupole moment. For reference, black contours show the true mass distribution.

$\langle\kappa\rangle \approx Q_{\mathrm{threshold}}$ (figure 4.9). We identify as possible filaments any region with $Q > Q_{\mathrm{threshold}}$ in a contiguous area or multiple peaks with total area $> 1.13\,\mathrm{arcmin}^2$, that is aligned within $\sim 45^\circ$ of the radial direction to the cluster centre. Applied to noise-free data and using $Q_{\mathrm{threshold}} = 0.005$, this recipe identifies 22 of the 40 filaments, all of which are real, i.e. 55% *completeness* (the number identified divided by the true number) and 100% *purity* (the number identified that are true divided by the number identified). The identified filaments are highlighted in magenta in figure 4.8.

4.6.4 Additional noise suppression strategies

Measurements of multipole moments will be more difficult in noisy data — especially for high n moments, where the diverging $|\boldsymbol{R}' - \boldsymbol{R}|^n$ term is particularly sensitive to noise in κ near the aperture boundary. We shall explore three strategies to reduce noise. First, noise can be averaged away by enlarging the aperture. However, signal is also averaged away for a filter that is not matched to the size of the feature – and filaments are relatively narrow, even around clusters at low redshift. Second, negative noise peaks can be eliminated by forcing $\kappa = \max\{\kappa, 0\}$. Negative con-

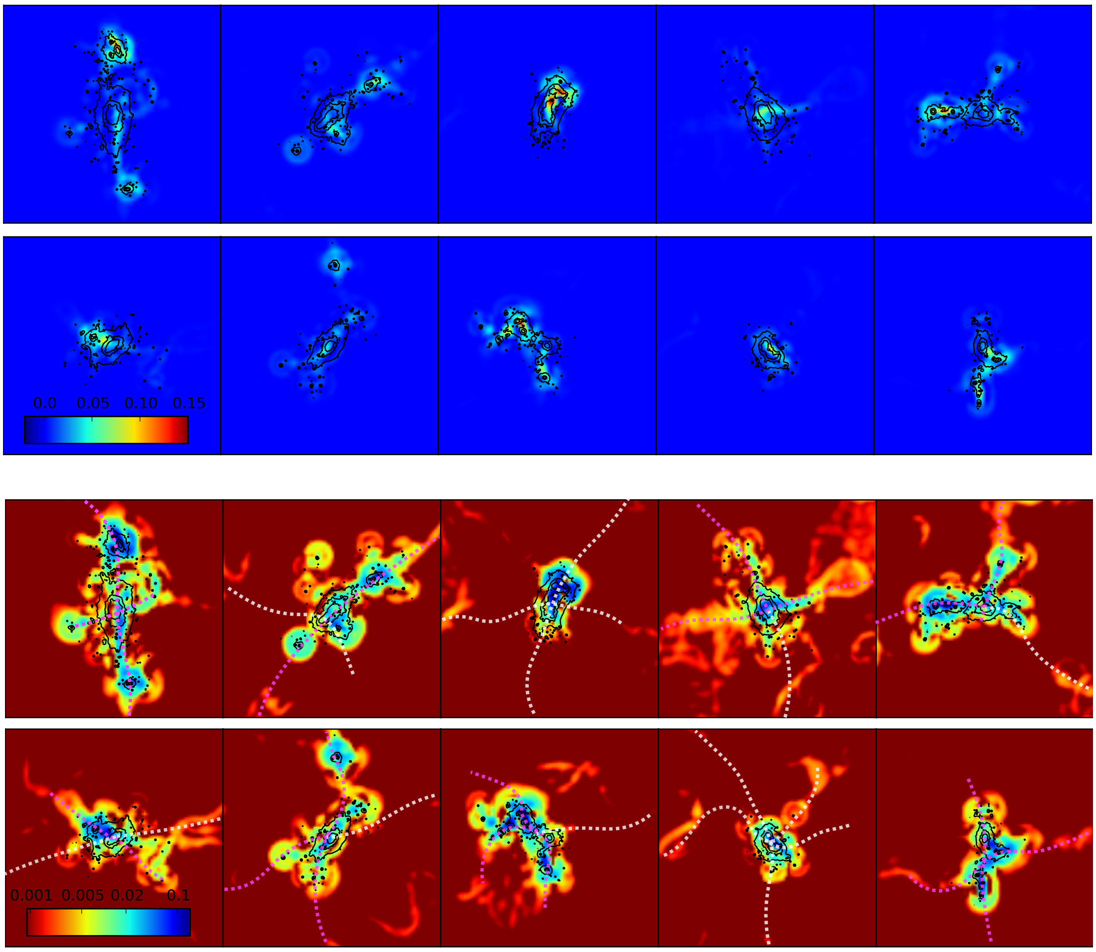


Figure 4.8: A combination of aperture multipole moments, Q (equations 4.6.12–4.6.14), can be used to identify filamentary features in a mass map. Colours (Top panel: linear scale, Bottom panel: logarithmic scale) show Q calculated from the true convergence map (without shape noise or LSS noise; black contours), after subtracting its best-fitting smooth component. Dotted lines reproduce the 40 filaments from figure 4.1. The 22 filaments successfully identified using Q and the procedure described in section 4.6.3 are highlighted in magenta.

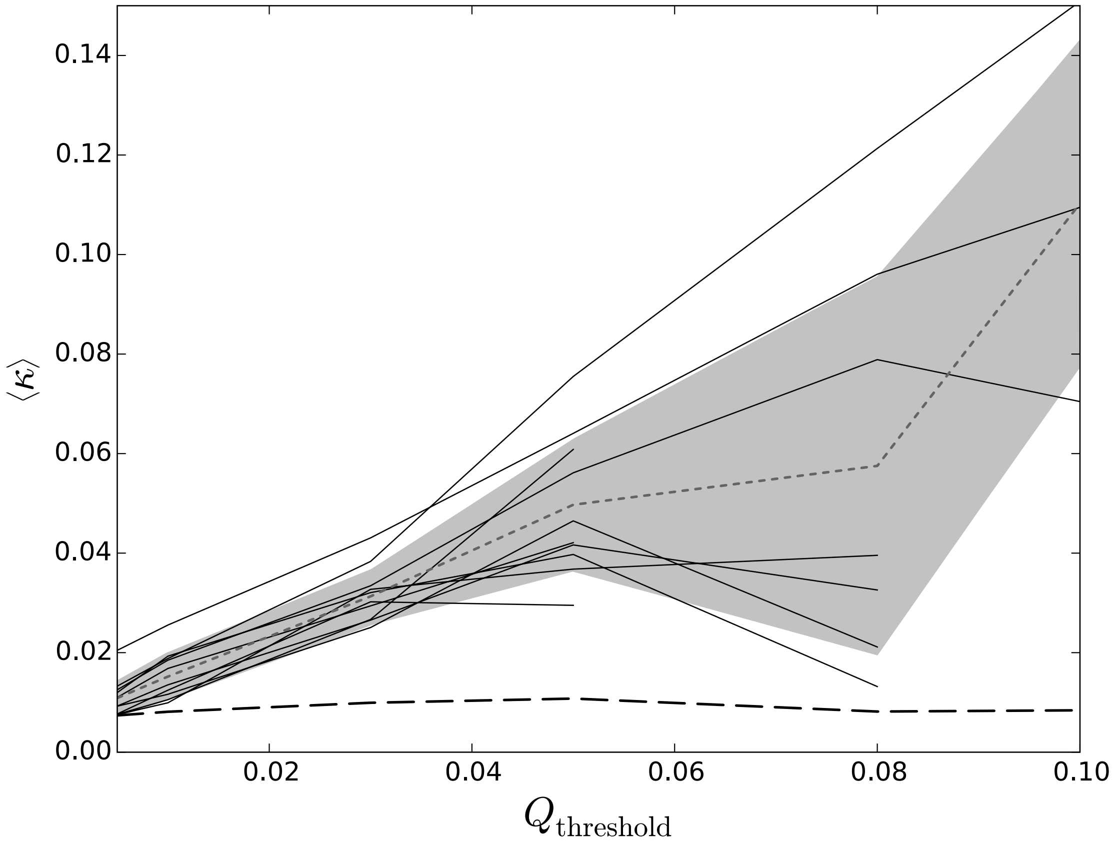


Figure 4.9: A combination of aperture multipole moments, Q (equations 4.6.12–4.6.14), can be used to identify features in a mass distribution with filamentary topology (see figure 4.8) and higher density than the background. Solid lines show the mean projected density $\langle\kappa\rangle$ inside a contour defined by $Q_{\mathrm{threshold}}$, for all 10 simulated clusters. The dotted line and shaded region show their mean and standard deviation. The normalisation of coefficients (4.6.13) is chosen so that $\langle\kappa\rangle = Q_{\mathrm{threshold}}$. The lower dashed line shows the mean convergence, weighted by the number of pixels that contain $Q > Q_{\mathrm{threshold}}$.

vergence is physically possible, because convergence represents deviation from the mean cosmic density; but it is unlikely along the line-of-sight to even a low density structure, and probably noise rather than signal. Third, we could assume that all filaments extend radially away from the cluster, while noise is isotropic, and suppress quadrupole and dipole moments whose phases are tangential. We calculate

$$Q_{n,\mathrm{projected}} = |Q_n| \, \cos{(\phi - \theta)}, \quad \text{with } \; n = 1,2 \tag{4.6.15}$$

where θ is an phase angle of Q_n. Figure 4.7(d) shows the projected quadrupole moments in the noise-free case, as an example.

4.6.5 Results

In the presence of galaxy shape noise and LSS noise, maps of our combination of aperture multipole moments Q have lower signal-to-noise than maps of convergence κ (figure 4.10; given the noise level, we show them only in linear scale, not logarithmic). We quantify the noise level by defining σ_Q as the standard deviation of all pixels in the final Q map. Despite our attempt to eliminate isolated substructures from the Q maps by combining different multipole moments, clusters 1, 2 and 5 contain sufficiently massive substructures to induce higher Q than lower-density filaments. Following the methodology in section 4.6.3, we then search for filaments as extended regions with $Q > 3\,\sigma_Q$ (illustrated in figure 4.10) or $Q > 4\sigma_Q$. Results for both are listed in table 4.3.

In the default Lenstool mass reconstructions, we find $\langle\sigma_Q\rangle = 0.011$ and, with $Q_{\mathrm{threshold}} = 3\sigma_Q$ we identify 17 of the 40 filaments (42.5% completeness), plus 5 false positive detections (77.3% purity). Increasing the detection threshold to $4\,\sigma_Q$ removes all but one false detection, but finds only 12 real filaments.

Identifying filaments in the noisier KS93+MRLens mass reconstructions is much more difficult. To obtain useful results, we need to apply all three denoising strategies presented in Sect. 4.6.4.

We enlarge the apertures to $R_{\mathrm{max},0}$=$2'$, $R_{\mathrm{max},1}$=$R_{\mathrm{max},2}$=$2.5'$; we replace negative convergence by zeros; and we project all quadrupole and dipole moments in the radial direction. In combination, these strategies reduce $\langle\sigma_Q\rangle$ from 0.11 to 0.06.

	Galaxy number density[arcmin^{-2}]	**Purity**		**Completeness**	
		$3\,\sigma_Q$	$4\,\sigma_Q$	$3\,\sigma_Q$	$4\,\sigma_Q$
KS93+MRLens	20	35.0%	40.0%	50.0%	35.0%
	50	41.7%	44.4%	37.5%	30.0%
	100	50.0%	57.9%	42.5%	27.5%
Lenstool	20	76.0%	78.0%	40.0%	27.5%
	50	77.3%	92.3%	42.5%	30.0%
	100	81.8%	93.3%	45.0%	35.0%

Table 4.3: Filament identification efficiency at 3σ or 4σ detection significance, from multipole aperture moments in mass maps created by KS93+MRLens or Lenstool, assuming different densities of weakly lensed galaxies. Completeness indicates the fraction of the 40 real filaments (see section 4.3.1) that are successfully identified. Purity indicates the fraction of the identified filaments that are real.

Filament identification statistics after this noise suppression are listed in table 4.3. At $3\,\sigma_Q$ detection threshold, we identify 15 of the 40 filaments (37.5% completeness), but also 21 false positive detections (41.7% purity).

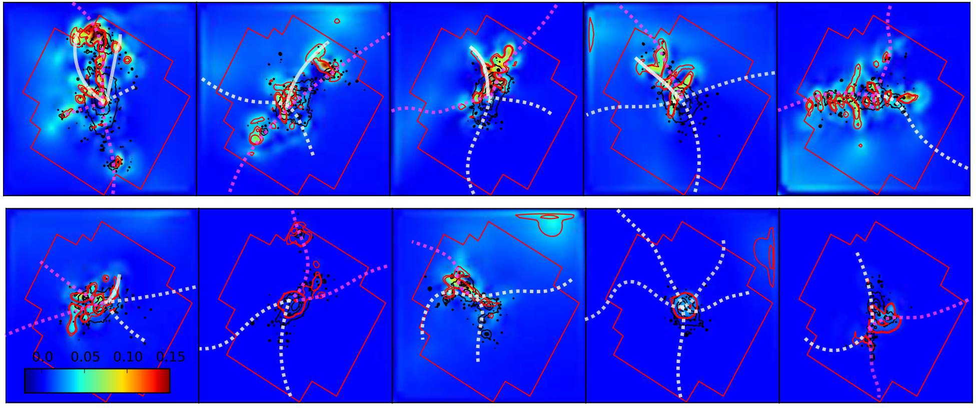
Most of the false-positive filament detections are caused by galaxy shape noise. Repeating the KS93+MRLens analysis with only shape noise yields a Q map with $\sigma_Q = 0.058$; with only LSS noise, it is $\sigma_Q = 0.033$. Because shape noise is apparently so dominant, we also investigate the effect of different survey strategies on the success of filament identification. We simulate ground-based observations, which typically resolve the shapes of only 20 galaxies arcmin^{-2}, and extremely deep space-based observations that resolve ~ 100 galaxies arcmin^{-2} (we assume all faint galaxies have constant intrinsic shape noise, as suggested by figure 17 of Leauthaud et al. 2007). With these catalogues, we repeat the whole analysis: including the mass reconstruction and filament search (table 4.3).

The low purity and high completeness of KS93+MRLens with 20 arcmin^{-2} source galaxy is because the Q maps are filled with random noise peaks that mimic the filament signals. Some radial directions defined by the alignment of noise peaks match the true filament direction by chance and thus boost the completeness in spite of low purity. Since these maps are not informative, we show only those Q measurements using 100 arcmin^{-2} source galaxies in figure 4.10. The performance of Lenstool reconstructions with deep space-based data is impressive: thanks to the prior assumption of looking harder where there are galaxies, it finds 18 filaments around 10 clusters (45% completeness) with 82% purity. Recall that, even with noise-free data (section 4.6.3), the maximum completeness with the multipole moment technique was 55%. In general, we find that Lenstool is most appropriate for filament searches. Applied to future deep space-based surveys, the multipole moment technique should detect one or two filaments around most clusters.

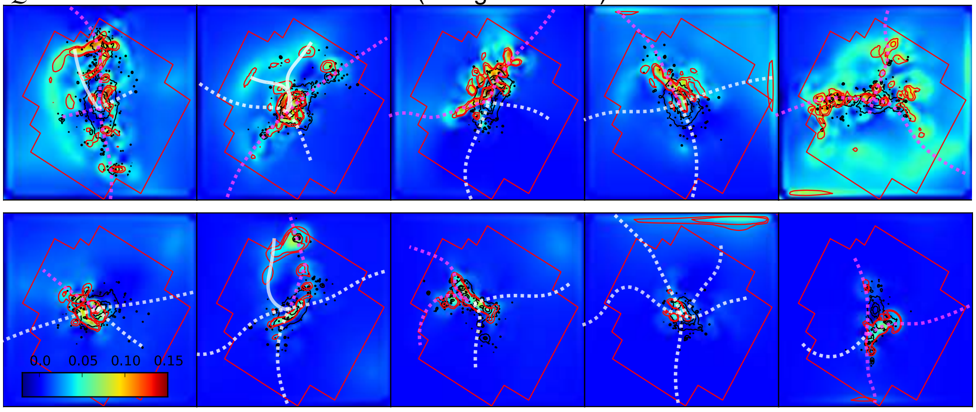
4.7 Conclusions

High-precision calibration of weak-lensing mass reconstruction techniques will be essential for the next generation of space-based surveys. Understanding methods'

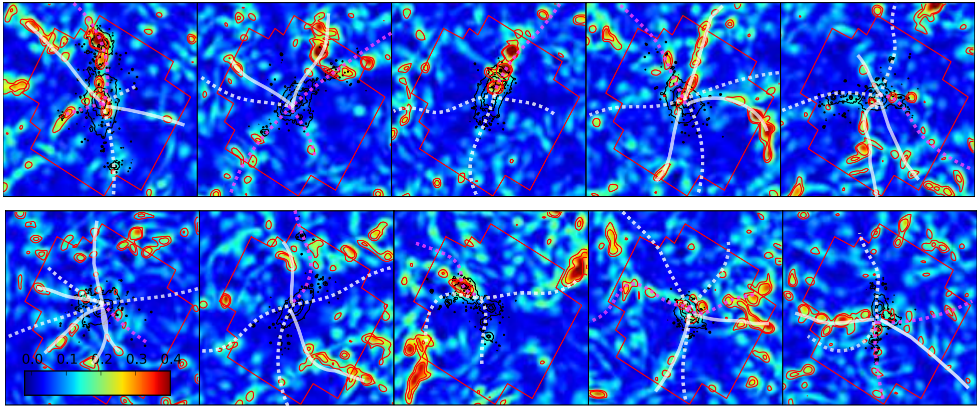
Q on the LENSTOOL Reconstructions (50 gal/arcmin^2)



Q on the LENSTOOL Reconstructions (100 gal/arcmin^2)



Projected Q on the Positive-only KS93+MRLENS Reconstructions (50 gal/arcmin^2)



Projected Q on the Positive-only KS93+MRLENS Reconstructions (100 gal/arcmin^2)

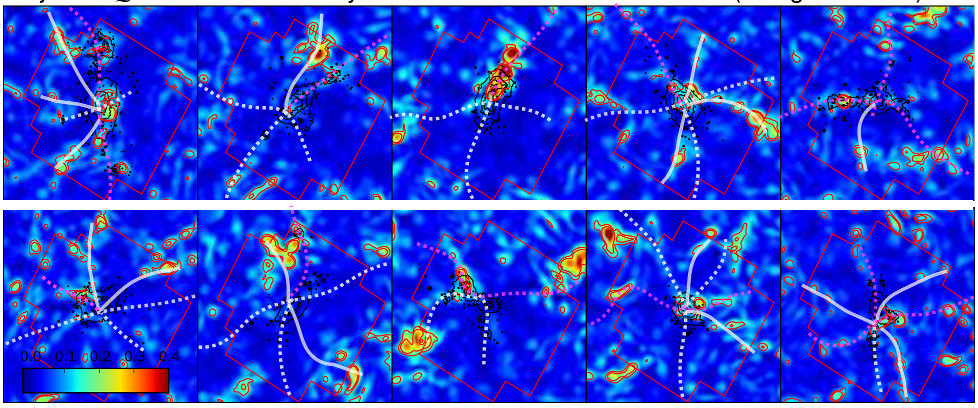


Figure 4.10: *(see next page)*

Figure 4.10: Results for the filament search around 10 simulated clusters. Colours show a linear combination of aperture multipole moments Q, calculated from the mass maps after subtracting their best-fit smooth component. Dotted lines show true filaments, reproduced from figure 4.1; those identified successfully (with $Q_{\text{threshold}} = 3\,\sigma_Q$, see section 4.6.3) are highlighted in magenta. Solid lines show false positive detections. The top and second panel use mass maps created by Lenstool (including shape noise and LSS), with $50\,\text{arcmin}^{-2}$ and $100\,\text{arcmin}^{-2}$ source galaxies, respectively. The third and bottom panel show the phase-projected version of the filter applied to the positive-only KS93+MRLens mass map (with a different colour scale to the top two panels). In all panels, red contours show $Q = 3\sigma_Q$ and $4\sigma_Q$, and black contours show the true mass distribution.

performance in different systems (such as non-linear structures or stacked clusters), and quantifying any biases they introduce, will help identify the optimal method for each scientific analysis.

In this chapter, we simulate mock observations of ten galaxy clusters from the BAHAMAS cosmological simulation. We use their known distribution of mass $4 \times 10^{14} < M_{200}/\text{M}_\odot < 2 \times 10^{15}$ to test two mass mapping methods: (1) direct KS93 inversion from lensing shear observations to the projected mass distribution, which is then denoised using MRLens; (2) the forward-fitting Lenstool technique that uses a Bayesian MCMC sampler to fit the distribution of mass in a multi-scale grid. Any mass reconstruction method must interpolate the finite resolution in a shear catalogue that samples the shear field only along the lines of sight to galaxies.

We find that MRLens is particularly efficient at suppressing noise owing to the diverse intrinsic shapes of background galaxies, whilst retaining signal from statistically significant structures on all scales. In a typical cluster field, it reduces total noise σ_κ from 0.088 ± 0.001 to 0.026 ± 0.001. However, it has no knowledge of cluster physics, and its noise suppression via smoothing softens the inferred central density profile. At large projected radii, $R > 1\,\text{Mpc}$, noise in the map of an individual cluster becomes dominated by unrelated structures at different redshifts, projected along adjacent lines of sight.

LENSTOOL incorporates physical knowledge of galaxy clusters by imposing strong priors on the distribution of mass. For example, it preserves central cusps. The method is more aggressive in denoising the reconstructed convergence field, achieving $\sigma_\kappa = 0.015 \pm 0.004$. By adjusting the grid's adaptive resolution, it is also possible to suppress the spurious signal from unrelated, isolated structures at different redshifts, once they have been identified via multiband photometry or spectroscopy. We find that this method is well-suited to reconstructions of individual clusters, or measurements of low signal-to-noise quantities, such as filaments. However, if the cluster members are misidentified in the first place, then it can cause bias in the LENSTOOL mass reconstruction due to the incorrect mass-trace-light prior.

In its standard configuration however, we find that LENSTOOL biases a mass reconstruction at large distances from the centre of a cluster, by imposing a prior that the projected density everywhere in a field of view must be positive (relative to the mean density in the Universe). This bias will need to be managed carefully when statistical errors are reduced by averaging over a population of clusters: perhaps by reconfiguring the Bayesian optimisation engine. The standard configuration of LENSTOOL also forces the mass distribution in every grid point to be spherically symmetric. In a purely weak-lensing analysis, this leads to spuriously spherical cluster cores, even when the global mass distribution is well modelled. This issue is automatically solved and irrelevant if strong gravitational lensing information is available, and used to pre-fit the axis ratio of the core. In this weak lensing-only study, we adopt a simple solution by masking the central $R < 35''$ regions of a weak-lensing-only reconstruction.

Based on the performance of these two methods, for an individual cluster, or measurements of highly nonlinear quantities such as filament detection, LENSTOOL is well-suited to applications that require as precise a reconstruction as possible. However, for high-precision analyses that stack many clusters, it would be necessary to drop LENSTOOL's positive definite constraint to reduce bias of mass over-estimation. By contrast, KS93+MRLENS retains a higher level of noise, but the positive and negative fluctuations are preserved in a manner which can reduce bias in stacked measurements.

We also develop a filter to search for filaments and measure their orientation. The low density of filaments leads to low signal-to-noise in reconstructed maps, and they can rarely be stacked usefully. To retain their individual signal whilst suppressing noise, we construct a linear combination of multipole moments. We explore two further strategies: (1) filtering on the orientations (complex phases) of higher-order moments, exploiting the prior knowledge that filaments typically extend radially out from cluster halos, and (2) replacing with the mean density of the Universe those regions inferred to have (negative) less density, which are more likely to be noise than regions inferred to have (positive) higher density. We find that it will be impossible to detect individual filaments using data from ground-based telescopes, and remains challenging with current space-based (*HST*) data. However, we find that the dominant source of noise relevant to filament detection comes from lensed galaxies' intrinsic shapes. Deeper observations with the next generation of space-based telescopes will resolve more background galaxies, and efficiently beat down this noise. Our filtering method successfully finds 45% of filaments with projected density $\Sigma > 1.7 \times 10^7\,\mathrm{M}_\odot/\mathrm{kpc}^2$ (with a false detection rate <20%), when applied to mock observations at the depth of possible future surveys.

Chapter 5

Space-based Lensing Analysis of Galaxy Cluster MS 0451

5.1 Abstract

Using the largest mosaic of *Hubble Space Telescope* images around a galaxy cluster, we map the distribution of dark matter throughout a ∼$6 \times 6\,\mathrm{Mpc}^2$ area centred on the cluster MS 0451−03 ($z = 0.54$, $M_{200} = 1.65 \times 10^{15}\,\mathrm{M}_\odot$). Our joint strong- and weak-lensing analysis shows three possible filaments extending from the cluster, encompassing six group-scale substructures. The dark-matter distribution in the cluster core is elongated, consists of two distinct components, and is characterized by a concentration parameter of $c_{200} = 3.79 \pm 0.36$. By contrast, *XMM-Newton* observations show the gas distribution to be more spherical, with excess entropy near the core, and a lower concentration of $c_{200} = 2.35^{+0.89}_{-0.70}$ (assuming hydrostatic equilibrium). Such a configuration is predicted in simulations of major mergers 2–7 Gyr after the first core passage, when the two dark-matter halos approach second turnaround, and before their gas has relaxed. This post-merger scenario finds further support in optical spectroscopy of the cluster's member galaxies, which shows that star formation was abruptly quenched 5 Gyr ago. MS 0451−03 will be an ideal target for future studies of the growth of structure along filaments, star-formation processes after a major merger, and the late-stage evolution of cluster collisions.

The following has been published in MNRAS as Tam et al. (2020b)

5.2 Introduction

In Chapter 4, we have shown the performance of two mass mapping techniques for halo shape measurement, and derived a optimal filter to search for filaments. In this Chapter, we apply these methods on observational data to constrain physical properties of a real galaxy cluster.

Higher resolution space-based imaging increases the signal-to-noise ratio (S/N) of lensing measurements. We use a wide-field *HST*/ACS imaging mosaic to conduct a combined strong+weak lensing analysis of the redshift z=0.54 galaxy cluster MS 0451-03 (hereafter MS 0451), also known as MACS J0454.1-0300 (Ebeling et al., 2001, 2007). It is the most X-ray luminous cluster in the Extended Medium Sensitivity Survey (EMSS; Gioia et al., 1990). Previous studies of MS 0451 have been conducted at optical wavelengths (Jørgensen & Chiboucas, 2013; Luppino et al., 1999; Martinet et al., 2016; Moran et al., 2007a,b; Soucail et al., 2015), in X-rays (Donahue et al., 2003; Jeltema et al., 2005; Jørgensen et al., 2018; Molnar et al., 2002) and via the Sunyaev-Zel'dovich (SZ) effect (De Filippis et al., 2005; Sayers et al., 2019). Strong gravitational lensing analyses have built a model of the cluster core (Berciano Alba et al., 2010; Borys et al., 2004; MacKenzie et al., 2014; Zitrin et al., 2011, Jauzac et al. sub.), and a ground-based weak lensing analysis detected a possible filamentary structure (Martinet et al., 2016). In 2014, MS 0451 was extensively observed with *HST* over a large area, providing the community with the largest *HST* mosaic centered on a galaxy cluster. In this Chapter, we exploit these wide *HST* observations, combining strong and weak gravitational lensing to map the mass distribution out to a projected radius of ∼3 Mpc.

Multi-wavelength data, such as X-ray imaging that traces the intra-cluster medium (ICM), is crucial to a better understanding of the dynamics in clusters. Since dark matter and baryons interact differently during a merger, a combined study of the distributions of dark matter and ICM provides an insight into clusters' evolutionary history (e.g. Bradač et al., 2006; Jauzac et al., 2015a; Merten et al., 2011; Molnar & Broadhurst, 2018; Ogrean et al., 2015). Furthermore, X-ray analyses usually assume that the ICM is in hydrostatic equilibrium (HSE) and spherically symmetric. Therefore, a comparison between the X-ray hydrostatic and lensing mass measurements can also be used as a check for deviations from HSE. Our collaborator used *XMM-Newton* observations to conduct X-ray analysis of MS 0451, which allows us to probe the dynamical state of this massive cluster.

This chapter is organised as follows. Existing multiwavelength observations of MS 0451 are summarised in Section 5.3. Our methods for measuring gravitational lensing and reconstructing the distribution of mass are described in Section 5.4. Our methods for X-ray data analysis are presented in Section 5.5. We present our measurements of the main cluster halo and surrounding large-scale structures in Section 5.6. We infer the cluster's dynamical state in Section 5.7, and conclude in Section 5.8. Throughout this chapter, we adopt a ΛCDM cosmology with $\Omega_m = 0.27$, $\Omega_\Lambda = 0.73$, and $H_0 = 70\,\mathrm{kms}^{-1}\mathrm{Mpc}^{-1}$, hence 1' corresponds to 6.49 kpc at the redshift of the cluster.

5.3 Observations

5.3.1 *Hubble Space Telescope* observations

HST/*ACS*

A $\sim 20' \times 20'$ mosaic of 41 high-resolution images around MS 0451 was obtained with the *Advanced Camera for Survey* onboard *HST* (ACS; Ford et al., 1996) between January 19 and February 3, 2014 (GO-9836, PI: R. Ellis), in the F814W pass-band, with an exposure time of 2 ks per pointing (single orbit depth). We reduced the data using the PYHST software[1] which corrects Charge Transfer Inefficiency using ARCTIC (Massey et al., 2014), removes bias and flat fields using CALACS (Miles et al., 2018), then stacks dithered images using ASTRODRIZZLE (Hoffman & Avila, 2018). These high-resolution images are used to measure the effect of weak gravitational lensing on the shapes of background galaxies.

---

[1] `https://github.com/davidharvey1986/pyHST`

HST/ *WFC3*

The cluster core ($2' \times 2'$) was imaged with the *Wide Field Camera 3* onboard *HST* (WFC3; Kimble et al., 2008) on January 13, 2010 (GO-11591, PI: J.-P. Kneib). We use observations in the F110W and F160W pass-bands, with exposure times of 17 912,s and 17 863 s respectively, for the strong-lensing analysis.

5.3.2 Ground-based Observations

Imaging Data

Multicolour imaging in the B,V,R_c,I_c and z' passbands was obtained with the 8.3 m Subaru telescope's wide-field Suprime-Cam camera for 1440 s, 2160 s, 3240 s, 1800 s, and 1620 s, respectively. Dates of these observations are December 21, 2006 (z'), December 11, 2001 (R_c,I_c) and January 23, 2009 (B,V). Near-UV imaging in the u^* passband was obtained by the 3.6 m CFHT's MegaPrime camera for 6162 s on November 27, 2006 (ID: 06BH34, PI: H. Ebeling). Near-infrared imaging in the J and K_S pass-bands was obtained with CFHT's Wide-field InfraRed Camera (WIRCam) on November 8, 2008 and October 25, 2007, respectively (ID: 08BH63, 07BH98, PI: C.-J. Ma). All observations were dithered to facilitate the removal of cosmic rays, minimising the impact of pixel defects and chip gaps; all data were reduced using standard procedures (Donovan, 2007). These data are used to measure photometric redshifts and thereby identify galaxies within, in front of, or behind the cluster.

Photometric redshifts were computed using the ground-based imaging. In order to allow a robust estimate of the spectral energy distribution (SED) to be obtained for all objects within the field of view, data from different pass-bands are seeing-matched using the technique described in Kartaltepe et al. (2008). The object catalogue is then created with the SExtractor photometry package (Bertin & Arnouts, 1996) in 'dual-image mode', with the R-band image being the reference detection image. Photometric redshifts for galaxies with magnitude $R_C < 24$ were subsequently computed using the adaptive SED-fitting code Le Phare (Arnouts et al., 1999; Ilbert et al., 2009). By comparing the spectroscopic and photometric redshifts, the typical statistical error of the photometric redshifts is found to be $\Delta z = 0.024$. For more details of this calibration, see Ma et al. (2008).

Spectroscopic observations

MS 0451 was also observed with the Multi-Unit Spectroscopic Explorer (MUSE; Bacon et al., 2010) at the VLT on January 10-11, 2016 (ID: 096.A-0105(A), PI: J.-P. Kneib) using its WFM-NOAO-N mode in good seeing conditions with full width at half maximum (FWHM) of approximately 0.8 arcsec. The MUSE observations consist of 2 pointings of 3 exposures, slightly shifted to account for systematics on the detector, covering a field of view of $\sim 2.2\,\mathrm{arcmin}^2$. These data are used for the strong-lensing analysis. They were reduced using version 1.6.4 of the MUSE standard pipeline (Weilbacher et al., 2012; Weilbacher et al., 2014), which corrects for bias and illumination; performs geometrical, astrometric and flux calibrations; then combines the 3 individual exposures of each pointing into a single data cube. The sky residuals within each data cube were finally subtracted using the Zurich Atmosphere Purge algorithm (Soto et al., 2016), which masks sources identified by SEXTRACTOR (Bertin & Arnouts, 1996), then uses principal component analysis to model the sky background.

The spectroscopic redshifts of galaxies used in this work were compiled from the literature and complemented by redshift measurements obtained by us, based on spectroscopic data obtained in September 2004 with Gemini-North/GMOS on Mauna Kea. The latter used a 1 arcsec slit, the 800 l/mm grating, and a spectral range typically from 4200 to 7000 Å. The resulting data were reduced using standard IRAF procedures.

5.3.3 *XMM-Newton* X-ray Observations

MS 0451 was observed by *XMM-Newton* (observation ID: 0205670101, PI: D. Worrall) on September 16-17, 2004 for a total of 44 ks. We reduced the *XMM-Newton*/EPIC data using the XMMSAS v16.1 software package and a pipeline developed in the framework of the *XMM-Newton Cluster Outskirts Project* (X-COP, Eckert et al., 2017). After performing the standard data reduction steps to extract calibrated event files, we used the XMMSAS tools MOS-FILTER and PN-FILTER to automatically define good-time intervals (unaffected by soft proton flares) of 24 ks (MOS1), 24 ks (MOS2), and 19 ks (PN). For more details of this procedure, see Ghirardini et al. (2019, Sect. 2 and Fig. 1). These data are used to measure the properties of the baryonic ICM.

An independent, $\sim$ 50ks *Chandra* observation provides a high-resolution X-ray view of the cluster core but is not used by us here, since the covered area does not match the extended *HST* mosaic that is the focus of this study.

5.4 Method: Gravitational Lensing Analysis

5.4.1 Weak Lensing Shear Catalogue

To detect sources and measure the shapes of galaxies in the *HST*/ACS imaging, we use the PYRRG (Harvey et al., 2019) implementation of the (Rhodes et al., 2000, hereafter RRG) shear measurement method. For the full details of this process, we refer the reader to Section 3.3.1.

Background galaxy selection

The *HST/ACS* galaxy catalogue output from PYRRG is contaminated by foreground and cluster member galaxies. These are not gravitationally lensed by the cluster, and dilute the shear signal. To eliminate these galaxies from our analysis, we use multicolour ground-based imaging.

For the 13% of galaxies in the *HST*/ACS catalogue that are brighter than $R_C <$ 24, we can assign photometric redshifts (see section 5.3.2). We thus identify and remove all cluster member galaxies with photometric redshift $0.48 < z_{\rm phot} < 0.61$, or spectroscopic redshift $0.522 < z_{\rm spec} < 0.566$.

For an additional 16% of galaxies, we obtain multicolour information in at least the B, R_C, and I_C bands. We have experimented with several criteria adopted elsewhere in the literature to identify foreground and cluster member galaxies (e.g. cuts in $B - R_C$ and $R_C - I_C$, or $B - V$ and $u - B$; see Jauzac et al. 2012; Medezinski et al. 2010, 2018). We obtain the cleanest catalogue by retaining only those galaxies

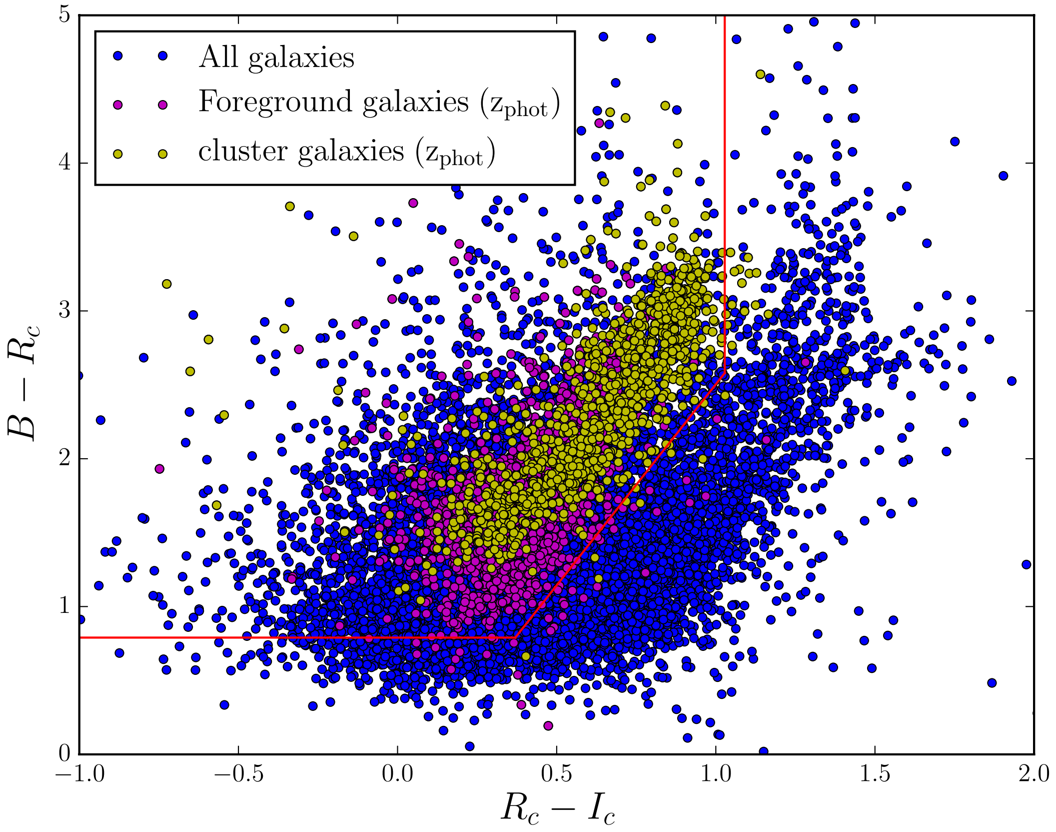


Figure 5.1: Colour-colour diagram ($B - R_C$ vs $R_C - I_C$) for objects within the *HST*/ACS mosaic of MS 0451. Blue dots represent all objects; magenta and yellow dots are galaxies classified as foreground and cluster galaxies thanks to photometric redshifts respectively. The red solid lines delineate the B, R_C and I_C colour cuts that minimize contamination in the catalogue.

with $(B - R_C)<0.79$, $(R_C - I_C)>1.03$, or $(B - R_C)<2.72(R_C - I_C) - 0.216$ (figure 5.1). After these colour cuts, the photometric redshift distribution of $R_C < 24$ galaxies suggests that the contamination from foreground and cluster members is $\sim$4% (figure 5.2 top panel). This is smaller than our statistical error budget, so we ignore this bias. We shall refer to the combined 30% of galaxies with photometric information as the 'bright sample'.

For the remaining 70% of galaxies without any ground-based photometric information, we first discard the 6% of galaxies brighter than $F814W < 24$ (which are mainly foreground or cluster member galaxies; in the bright sample, 80% of foreground galaxies and 89% of cluster members have $F814W < 24$, and their

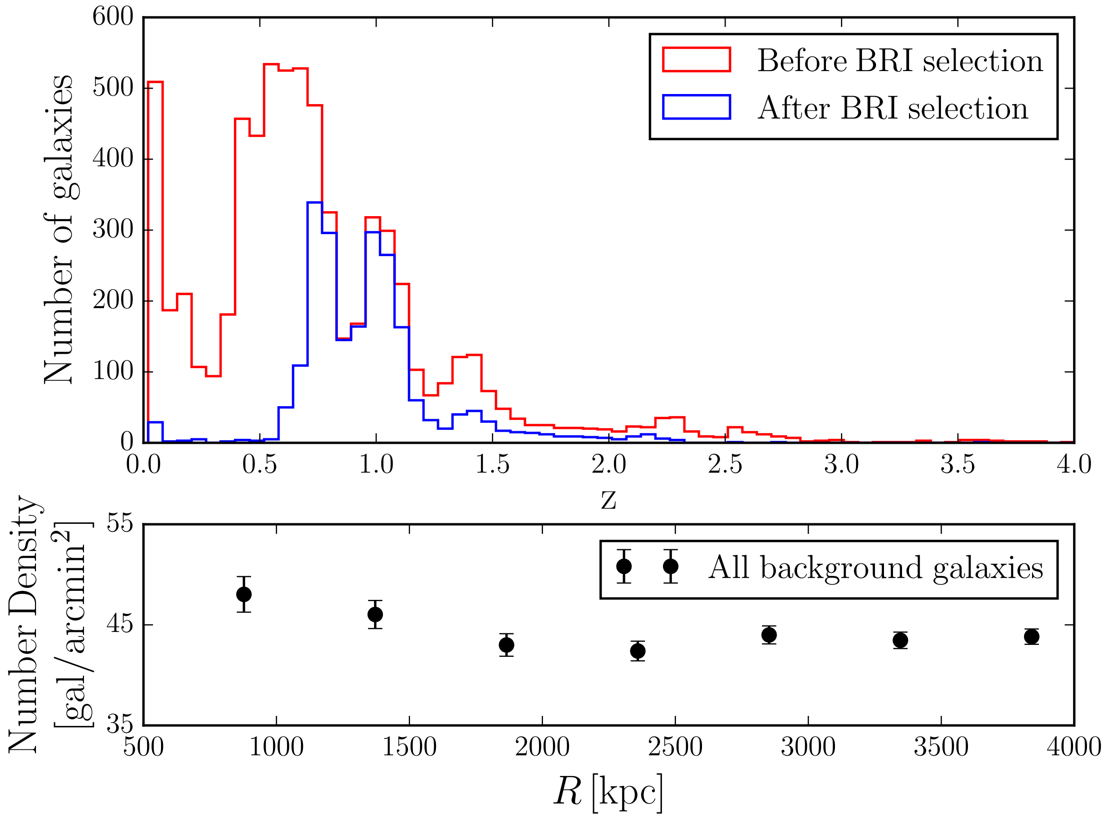


Figure 5.2: The successful identification of background galaxies. *Top:* Redshift distribution of all galaxies that have a spectroscopic and/or photometric redshift (red histogram). The blue histogram shows the redshift distribution of galaxies classified as background sources based on B, R_C, I_C colour-colour selection. *Bottom:* Number density of *all* background galaxies in the final weak-lensing catalogue (including fainter galaxies without observed colours), as a function of their projected distance from the cluster center.

combined magnitude distribution peaks at $F814W \sim 23$). For this 'faint sample' of galaxies, we assign nominal redshifts drawn at random from a distribution $N(z > 0.54) \propto (e^{-z/z_0})^{\beta}$, with $\beta = 1.8$ and median redshift $z_0 = 0.71$ (Gilmore & Natarajan, 2009; Natarajan & Kneib, 1997).

Our final weak lensing catalogue (combining the 'bright' and 'faint' samples) contains 21,232 background galaxies, corresponding to 44 galaxies arcmin^{-2}. Before cuts, an excess of ~ 35 galaxies arcmin^{-2} is present within 1 Mpc of the cluster centre; our selection process removes these galaxies, leaving an approximately constant density throughout the field (figure 5.2 bottom panel), as expected for an uncorrelated population of background galaxies. Of these background sources, 10%, 11% and 79% are selected via cuts in redshift, colour and magnitude, respectively.

5.4.2 Strong lensing constraints

For this analysis, we adopt the best-fit strong-lensing mass model from Jauzac et al. (sub.). We here only give a summary of the strong-lensing mass model, and refer the reader to Jauzac et al. (sub.) for more details. The cluster core is modeled using two cluster-scale halos and 144 galaxy-scale halos associated with cluster galaxies. All potentials are modeled using Pseudo-Isothermal Elliptical Mass Distributions (PIEMDs; Elíasdóttir et al., 2007; Kassiola & Kovner, 1993b; Limousin et al., 2005) which are described by seven parameters: position (x,y), ellipticity e, position angle θ, core radius r_{core}, truncation radius r_{cut}, and velocity dispersion σ. Best-fit parameters for the two cluster-scale components are listed in Table 5.1.

Seven cluster galaxies acting as small-scale perturbers of some of the multiple images are independently modeled as individual PIEMDs. The rest of the cluster galaxy population is modeled using scaling relations; to limit the number of free parameters, positions, ellipticities, and position angles of all galaxies are fixed to the respective values of the observed stellar component. The galaxies' velocity dispersions are scaled from the observed stellar luminosity according to the Faber & Jackson (1976) relation, which describes well the mass in early-type cluster galaxies (Jullo et al., 2007; Wuyts et al., 2004).

The strong-lensing mass model is constrained by 16 systems of multiple images (47 images in total). These include well known lensed objects such as a sub-millimeter arc at $z \sim 2.9$ (Borys et al., 2004), five other sub-millimeter systems (MacKenzie et al., 2014), a triply imaged galaxy (Takata et al., 2003), and six new systems identified with VLT/MUSE, including a quintuple image at $z = 6.7$ in the poorly constrained northern region. The latter has a redshift from VLT/XShooter observations and was previously studied by Knudsen et al. (2016). Five of these systems are spectroscopically confirmed, two of them newly identified by Jauzac et al. (sub.) using MUSE observations. The quintuple-image system in combination with the two new systems identified through MUSE observations in the northern region motivated the addition of a second cluster-scale halo in the strong-lensing mass model. Without this second large-scale halo, the geometry of the $z = 6.7$ system cannot be recovered and the root-mean-square (rms) distance between the observed and predicted locations of the multiple images of other systems is unacceptably high at $> 1.5''$. Two close groups of cluster galaxies were identified in this region. Adding a third cluster-scale mass halo did not significantly improve the model.

The resulting best-fit strong-lensing mass model has an rms separation of $0.6''$. The best-fit parameters of the two main cluster halos are given in Table 5.1. Note that the coordinates of the halos are given in arcseconds relative to the cluster center, here the BCG (α =73.545202, $\delta = -3.014386$).

5.4.3 Lensing 2D mass map

Forward fitting with Lenstool

We reconstruct the 2D distribution of mass using version 7.1 of Lenstool, whose performance has been quantified on mock *HST* data in Chapter 4. This returns 1700 Markov Chain Monte Carlo (MCMC) samples of the posterior likelihood, from which we compute the mean mass map.

The mass modeling here is different from Chapter 4 which only considered weak lensing signal. In this work, we conduct a combined strong+weak lensing analysis. In the cluster core, we fix the mass distribution to the best-fit of the strong-lensing model (see Sect 5.4.2). This includes 2 cluster-scale halos separated by 237 kpc,

Parameter	**Main halo**	**Second halo**
R. A.	$-7.5^{+0.9}_{-1.2}$	$22.3^{+3.1}_{-0.1}$
Dec.	$-2.6^{+0.6}_{-0.7}$	$19.5^{+4.8}_{-0.1}$
σ (km/s)	1001^{+30}_{-25}	810^{+210}_{-670}
e	$0.63^{+0.04}_{-0.03}$	$0.18^{+0.12}_{-0.06}$
θ (deg)	32.2 ± 0.5	147^{+9}_{-16}
$r_{\rm core}$ (kpc)	120^{+10}_{-6}	332^{+60}_{-30}
$r_{\rm cut}$ (kpc)	[1000]	680^{+200}_{-570}

Table 5.1: Best-fit parameters for the two cluster-scale PIEMD components of the strong lensing model of MS 0451. The coordinates are expressed in arcseconds relative to the location of the BCG (α =73.545202, δ =-3.014386). The truncation radius of the larger halo is outside the strong lensing information. It was thus fixed to 1 Mpc.

plus the 7 individually optimised galaxy-scale components. To extend our analysis from the cluster center to ~3 Mpc, we add a total of 1277 galaxy-scale halos at the locations of cluster member galaxies identified via spectroscopic and photometric redshifts over the entire field of view covered by the *HST* mosaic. Each of the 1277 cluster galaxies is modelled as a PIEMD potential with fixed $r_{\rm core} = 0.15\,{\rm kpc}$ and $r_{\rm cut} = 58\,{\rm kpc}$, and velocity dispersion σ that is scaled relative to an $m^*_K = 18.7$ galaxy with $\sigma^* = 163.10\,{\rm km\,s^{-1}}$ using the Faber & Jackson (1976) relation.

Throughout the rest of the mosaic area outside the strong lensing region, we add a free-form grid of mass, whose resolution is adapted to the local K-band luminosity. The details of this technique are described in Section 3.4.2. The final grid model (Figure 5.3) includes a total of 5570 individual Pseudo-Isothermal Mass Distributions. To avoid superceding the strong lens model, we prevent the mass grid from extending inside the multiple image region, defined as an ellipse aligned with the cluster core ($a = 44''$, $b = 34''$, $\theta = 30°$ counter-clockwise with respect to the East-West axis, centered on α=73.545202°, δ=−3.0143863°). We also exclude shear measurements from this region.

Uncertainty

We estimate the noise in each pixel of the Lenstool mass map via bootstrap resampling. To this end, we first select the two 2 Mpc × 2 Mpc patches of the sky[2] outside the cluster core that contain the smallest grid cells (and hence the highest K-band luminosity peaks) where substructures and filaments are most likely to be located. We choose a random orientation for each shear measurement in these two patches of sky, then reconstruct a new mass map using Lenstool. Inside an aperture of $r < 480\,\mathrm{kpc}$, the mean noise level of 100 random realisations is found to be $\langle M \rangle = 2.08 \times 10^{13}\,\mathrm{M}_{\odot}$, which is non-zero because of Lenstool's positive-definite mass prior, and its rms uncertainty is $\sigma_M = 1.64 \times 10^{13}\,\mathrm{M}_{\odot}$. We use the latter to normalise the signal-to-noise ratios of substructures detected in Sect. 5.6.5.

Mass mapping with KS93+MRLens

As an alternative to the Lenstool multiscale grid method for lensing mass mapping, we use the Kaiser & Squires (1993) direct inversion method. This converts an observed, binned shear map, $\boldsymbol{\gamma}(\boldsymbol{R})$ into a convergence $\kappa(\boldsymbol{R})$ map, via their Fourier transforms. We then denoise the convergence field using MRLens. The details of this mass mapping method are described in Section 3.4.1 and the setting of MRLens parameters are the same as in Section 4.4.1.

5.4.4 Lensing 1D density profile

We calculate the cluster's 1D radial density, $\Sigma(r)$, by (azimuthally) averaging the 2D mass distribution in logarithmically spaced annuli between 80 kpc and 4 Mpc. To enable a statistically rigorous analysis of this key characteristic, we calculate the full covariance matrix $C_{i,j}$ between measurements in each bin (see Sect. 5.4.4).

---

[2] The two patches of sky used to estimate the level of noise in the weak lensing map are centered at ($\alpha = 73.644053$, $\delta = -3.012986$) and ($\alpha = 73.426295$, $\delta = -3.089766$).

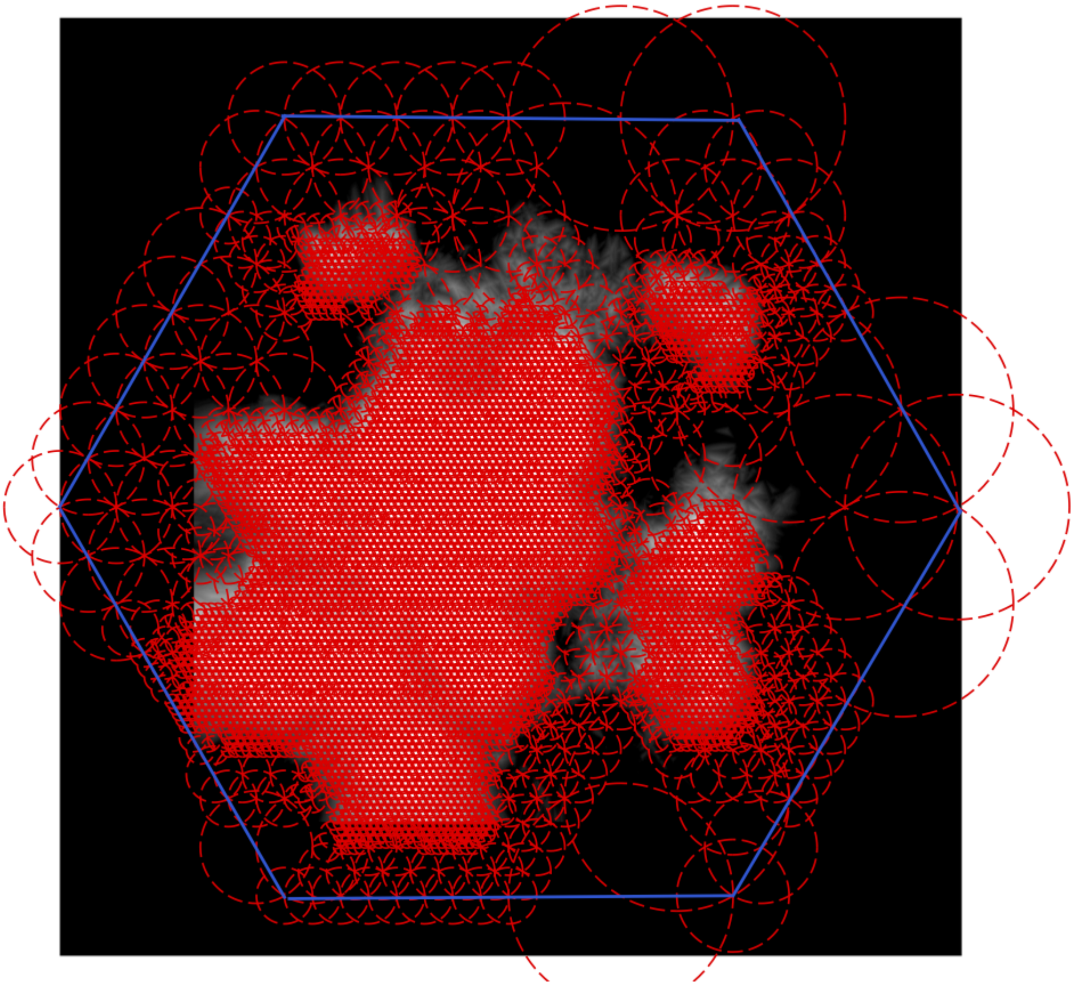


Figure 5.3: The multi-scale grid that determines the maximum spatial resolution of the Lenstool mass reconstruction. One RBF is placed at the centre of each circle, with core radius r_c equal to the radius of the circle, and a free mass normalisation. The grid is determined from (and shown overlaid upon) the cluster's K-band emission. The blue hexagon covers an area slightly larger than the *HST* field of view.

Model comparison

We compare the mean density profile to five models obtained from cosmological simulations: NFW (Navarro et al., 1996, 1997), generalised NFW (gNFW; Zhao, 1996), Einasto (Einasto, 1965), Burkert (Burkert, 1995) and DK14 (Diemer & Kravtsov, 2014). A mathematical definition and description of each halo model is given in Section 2.2.5. We optimise the free parameters of each model using EMCEE (Foreman-Mackey et al., 2013) with a likelihood function

$$\log\mathcal{L} = -\frac{1}{2}\sum_{i,j=1}^{N_{\rm bin}}(\Sigma_i - \hat{\Sigma}_i)\, C^{-1}_{i,j}(\Sigma_j - \hat{\Sigma}_j) - \frac{1}{2}N_{\rm bin}\log\left(2\pi|C|\right), \tag{5.4.1}$$

where $\hat{\Sigma}$ is the model, $N_{\rm bin} = 20$ is the total number of radial bins, and $|C|$ is the determinant of the covariance matrix. When fitting the NFW, gNFW, Einasto, and Burkert models, we adopt flat uniform priors for $M_{200} \in [5, 30] \times 10^{14}\,{\rm M}_{\odot}$, and $c_{200} \in [1, 10]$. We also adopt a flat prior for the gNFW and Einasto shape parameters, $\alpha \in [0, 3]$ and $\alpha_E \in [0.02, 0.5]$, respectively. For the Burkert model, we use a flat prior for the core radius, $r_{\rm core} \in [100, 800]\,{\rm kpc}$. For the DK14 model, following More et al. (2016) and Baxter et al. (2017), we use the priors for $\rho_{\rm s}$, $r_{\rm s}$, $r_{\rm t}$, $\log(\alpha)$, $\log(\beta)$, $\log(\gamma)$, and $s_{\rm e}$ that are listed in Table 2 of Chang et al. (2018). Because the location of MS 0451 is so well known from strong-lensing constraints, we omit their miscentering term.

To compare the goodness of fit for models with different numbers of free parameters, we calculate the Bayesian Information Criterion

$$\mathrm{BIC} = -2\log\mathcal{L} + k\log N_{\rm bin}, \tag{5.4.2}$$

the Akaike Information Criterion

$$\mathrm{AIC} = -2\log\mathcal{L} + 2k, \tag{5.4.3}$$

and the corrected Akaike Information Criterion

$$\mathrm{AICc} = \mathrm{AIC} + \frac{2\,k\,(k+1)}{(N_{\rm bin} - k - 1)}, \tag{5.4.4}$$

where k is the number of free parameters. These three information criteria include penalty terms for adding free parameters that make a model more complex. The

BIC has a larger penalty term than the AIC; the AICc approaches AIC as $N_{\rm bin}$ increases, but is more robust for small $N_{\rm bin}$. The individual information criteria are not interpretable because they contain arbitrary constants. Therefore, it is useful to define the differences in information criteria between different models. For example, $\Delta{\rm AIC_i} = {\rm AIC_i} - {\rm AIC_{min}}$, where ${\rm AIC_{min}}$ is the minimum AIC value (the best model) among various halo models. For all three criteria, lower values indicate the preferred model. We can assess the relative merits of models based on these rules (Burnham & Anderson 2002): models having $\Delta_i < 2$ have substantial evidence; models with $4 < \Delta_i < 7$ have considerably less evidence; while models with $\Delta_i > 10$ have no evidence. In this work, we use the information criteria of the best-fit NFW model to define ΔAIC (see Table 5.2).

Covariance matrix

When fitting a parametric density profile to the azimuthally averaged mass maps, a first estimate of the uncertainty on the density at a given radius can be obtained by looking at the spread of densities at that radius in the MCMC samples generated by Lenstool (or bootstrap sampling, as described in Section 5.4.3). However, the non-local mapping between observable shear and reconstructed mass, leads to covariance between adjacent pixels. To fully account for this, we calculate the covariance matrix between radial bins i and j

$$C_{\mathrm{SL+WL}\,(i,j)} = \frac{1}{N}\sum_{l=1}^{N}\big(\Sigma_{l,i} - \langle\Sigma_i\rangle\big)\big(\Sigma_{l,j} - \langle\Sigma_j\rangle\big), \tag{5.4.5}$$

where N is the number of MCMC samples generated by Lenstool, $\Sigma_{l,i}$ is the surface mass density of the $l^{\rm th}$ sample in the $i^{\rm th}$ spatial bin, and $\langle\Sigma_i\rangle$ is the mean surface mass density of MCMC samples in the $i^{\rm th}$ spatial bin.

Inside the cluster, statistical noise is dominated by the finite number and intrinsic (unlensed) shapes of background galaxies used for the weak lensing analysis. Very near the cluster core, our default strong+weak lensing analysis underestimates the noise, because we fixed the strong lensing potentials. To account for that in the covariance matrix, we reconstruct a separate mass map using only weak lensing information, reinstating the mass grid and shear measurements in the core region.

This yields a covariance matrix C_{WL} that is significant up to $R{\sim}1\,\mathrm{Mpc}$. We therefore define a combined covariance matrix

$$C_{\mathrm{shape}\,(i,j)} = \begin{cases} C_{\mathrm{SL+WL}\,(i,j)} + C_{\mathrm{WL}\,(i,j)} & \text{for } R_i \text{ and } R_j < 1\,\mathrm{Mpc} \\ C_{\mathrm{SL+WL}\,(i,j)} & \text{otherwise.} \end{cases} \tag{5.4.6}$$

Note that this procedure overestimates $C_{i,j}$ in bins close to the $R \sim 1\,\mathrm{Mpc}$ transition region. However, this effect is small and therefore negligible in our measurement.

In the outskirts of a cluster, statistical uncertainty is dominated by large-scale structure (LSS) projected by chance along the line-of-sight. The specific realisation of LSS along the line-of-sight to any real cluster is static, but we account for its contribution to the covariance matrix $C_{\mathrm{LSS}\,(i,j)}$ by analysing mock observations of clusters from the BAHAMAS simulation (McCarthy et al., 2017) along many different lines of sight (see Section 4.3.2). We combine the two components across the full range of scales,

$$C_{(i,j)} = C_{\mathrm{shape}\,(i,j)} + C_{\mathrm{LSS}\,(i,j)} \ . \tag{5.4.7}$$

5.4.5 Lensing-derived halo shape

We measure the shape of MS 0451 by fitting our reconstructed 2D mass map with elliptical NFW models (eNFW; Oguri et al., 2010) (see Section 4.5.1). We fix the centre of the eNFW halo to the location of BCG ($\alpha = 73.545202$, $\delta = -3.014386$), then optimise[3] its four free parameters (with the allowed range for the parameters within the optimisation: $M_{200} \in [0.5, 3] \times 10^{15}\,\mathrm{M}_\odot$, $c_{200} \in [1, 10]$, position angle $\phi \in [0, 180]^\circ$, and axis ratio $q = a/b \in [0.1, 0.9]$) to minimise the absolute difference between the observed and modelled mass maps, integrated inside a circular aperture. To measure the cluster's change of shape as a function of radius, we repeat this fit inside circular apertures of varying radii. We perform this fit on every mass model

---

[3] Using the L-BFGS-B algorithm (Byrd et al., 1995) from Python's SCIPY.MINIMIZE package `https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.minimize.html`

output in Lenstool's MCMC chains, and measure the mean and rms values for each free parameter, marginalising over all others.

5.5 Method: X-ray Analysis

5.5.1 X-ray imaging analysis

We extract *XMM-Newton* images and exposure maps in the [0.7–1.2] keV band from the cleaned event files. To predict the spatial and spectral distribution of the particle-induced background, we use a collection of filter-wheel-closed observations. We compute model particle background images by applying a scaling factor to the filter-wheel-closed data to match the count rates observed in the unexposed corners of the three EPIC cameras. The images, exposure maps and background maps for the three detectors are then summed to maximize the signal-to-noise ratio.

To determine the thermodynamic properties of the source, we extract spectra for the three EPIC detectors in 7 annular regions from the center of the source to its outskirts (radial range $0'$–$4'$). We also extract the spectra of a region located well outside the cluster to estimate the properties of the local X-ray background. The redistribution matrices and effective area files are computed locally to model the telescope transmission and detector response. For each region, the spectra of the three detectors are fitted jointly in `XSPEC` (Arnaud et al., 1999) with a model including the source (described as a single-temperature thin-plasma APEC model (Smith et al., 2001) absorbed by the Galactic N_H), the local three-component X-ray background as fitted in the background region, and a phenomenological model tuned to reproduce the spectral shape and intensity of the particle background. The best-fitting parameters (temperature, emission measure, and metal abundance) as a function of radius are obtained by minimizing the C-statistic.

5.5.2 X-ray 1D surface brightness profile

To measure the 1D surface brightness profile of the cluster, we use the azimuthal median technique (Eckert et al., 2015), which allows us to excise the contribution of infalling substructures and asymmetries. To this aim, we use Voronoi tessellation to construct an adaptively-binned surface brightness map of the cluster. For each annulus, we then draw the distribution of surface brightness values and estimate the median value. Uncertainties are estimated by performing 10^4 bootstrap resamplings of the distribution and computing the root-mean-square deviation of the measured medians. We measure the local background outside the cluster, where the brightness profile is flat, then subtract it from the source profile. Gas density profiles are finally determined by deprojecting the surface brightness profile, assuming spherical symmetry.

We estimate the mass profile of the cluster from the gas density and temperature profiles, assuming HSE (see Pratt et al., 2019, for a review). This asserts that, if the gas is at rest within the gravitational potential of the halo, the pressure gradient balances the gravitational force,

$$\frac{dP_{\mathrm{gas}}}{dr} = -\rho_{\mathrm{gas}} \frac{GM(< r)}{r^2}. \tag{5.5.8}$$

The gravitating mass profile can thus be inferred from the gas pressure and density profiles. To solve equation (5.5.8), we use the *backwards* approach introduced by Ettori et al. (2019). Namely, we use a parametric model for the mass profile (here, NFW) and combine it with the density profile computed through the multi-scale decomposition technique to predict the pressure (and hence, temperature) as a function of radius. The model temperature profile is projected along the line of sight and corrected for multi-temperature structure along the line of sight using the Mazzotta et al. (2004) scaling. The projected temperature profile is then compared to the data and the parameters of the mass model (i.e. mass and concentration) are optimized using MCMC. The integration constant, P_{out}, which describes the overall pressure level at the edge, is left free while fitting and determined on-the-fly. For more details on the mass reconstruction technique and a careful assessment of the level of systematic involved, see Ettori et al. (2019).

The cumulative gas mass profile is computed by integrating the gas density profile over the volume, assuming the source is spherically symmetric,

$$M_{\mathrm{gas}}(< r) = \int_0^r 4\pi r'^{\,2} \rho_{\mathrm{gas}}(r')\, dr', \tag{5.5.9}$$

where $\rho_{\text{gas}} = \mu m_p(n_e + n_H)$, with n_e and $n_H = n_e/1.17$ the number density of electrons and protons, respectively, $\mu = 0.61$ the mean molecular weight, and m_p the proton mass. Our procedure directly outputs the hydrostatic gas fraction $f_{\text{gas,HSE}}(r) = M_{\text{gas}}(r)/M_{\text{HSE}}(r)$, which traces the virialization state of the gas (Eckert et al., 2019).

5.6 Results

5.6.1 2D mass distribution

The distribution of mass around MS 0451 (mapped in figure 5.4) has a core that is elongated along an axis from South-East to North-West, and surrounded by lower-mass substructures. Our primary Lenstool method achieves higher spatial resolution in regions containing cluster member galaxies, and suppresses more noise in regions without them. From the strong-lensing constraints, two distinct mass peaks separated by 237 kpc in the elongated core are detected. This is consistent with the analysis of CFHT/Megacam ground-based weak lensing measurements which also detected two dark matter halos in the core region (Martinet et al., 2016, and shown in figure 5.6 with magenta contours provided via private communication by N. Martinet). We confirm the existence of several nearby substructures, but our higher S/N data do not show them joined up into a filament running South-West to North-East, as hypothesised by Martinet et al. (2016). X-ray emission is detected out to $R = 1.7\,\text{Mpc}$ (figure 5.6).

An alternative weak-lensing-only reconstruction using Kaiser & Squires (1993) direct inversion obtains a noisier mass map (figure 5.5) but finds consistent features. The elongated cluster core reconstructed from KS93+MRLens is still detected with more than $4\sigma_\kappa$ statistical significance, where σ_κ is the standard deviation of κ over all the pixels within the *HST* field of view. It also shows elongation along the South-East to North-West direction, consistent with other weak lensing-only analyses (Martinet et al., 2016; Soucail et al., 2015). Several weak lensing peaks are detected at lower (1–3 σ_κ) statistical significance than with Lenstool while these include all the confirmed substructures from our Lenstool analysis. Overall, this

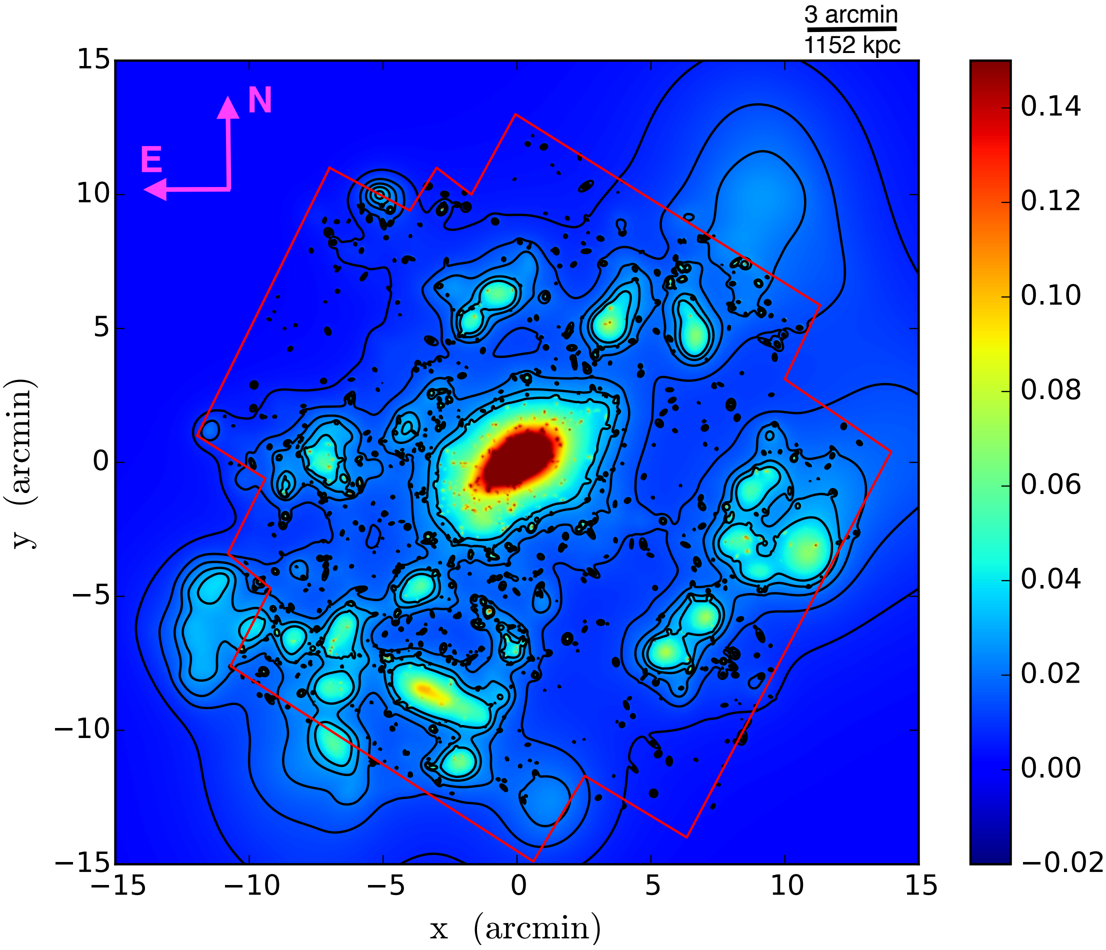


Figure 5.4: The projected distribution of mass around MS 0451, inferred from our Lenstool strong+weak lensing reconstruction and centred on (α=73.545202, δ=−3.0143863). Colours indicate the projected convergence, κ. Black contours show signal-to-noise in steps of $1\sigma_\Sigma$, measured from bootstrap re-sampling (see Sect. 5.4.3). The red polygon indicates the field of view of the *HST*/ACS imaging mosaic.

convergence map shows a similar level of noise and spurious (often line-of-sight) peaks as the maps reconstructed from mock observations of simulated clusters (see figure 4.2).

In the following, we will focus our analyses on the reconstructed mass map from Lenstool.

5.6.2 Total mass and density profile

Our combined weak- and strong-lensing reconstruction smoothly extends the surface mass density profile outside the multiple-image region (figure 5.7). We measure a projected mass $M(R < 195\,\mathrm{kpc}) = (1.85 \pm 0.87) \times 10^{14}\,\mathrm{M}_{\odot}$, consistent with previous strong-lensing measurements of $1.73 \times 10^{14}\,\mathrm{M}_{\odot}$ (Berciano Alba et al., 2010) and $1.8 \times 10^{14}\,\mathrm{M}_{\odot}$ (Zitrin et al., 2011). At larger radii, our analysis is sensitive for the first time to additional infalling or projected substructures; compared to previous models, based solely on strong-lensing features, we detect excess mass at $R > 3\,\mathrm{Mpc}$.

Theoretically motivated models to fit the 1D lensing signal (figure 5.7) are described in Section 2.2.5, and their best-fit parameters are listed in Table 5.2. For the best-fit NFW model we measure a mass $M_{200c} = (1.65 \pm 0.24) \times 10^{15}\,\mathrm{M}_{\odot}$ inside $R_{200c} = 1.99 \pm 0.06\,\mathrm{Mpc}$, or $M_{500} = (1.13 \pm 0.16) \times 10^{15}\,\mathrm{M}_{\odot}$, and concentration $c_{200} = 3.79 \pm 0.36$. Within the statistical uncertainty, this result is consistent with the ground-based weak-lensing measurement of $M_{200} = (1.84 \pm 0.35) \times 10^{15}\,\mathrm{M}_{\odot}$ for fixed $c_{200} = 4$ (Soucail et al., 2015). The Burkert profile is disfavoured by the BIC and AIC. For the best-fit gNFW and Einasto models we find masses and concentrations slightly lower than for NFW. However, their BIC and AIC differ by less than 2 from the NFW ones. Thus, we conclude that our data are unable to distinguish between these three models with statistical significance (as quantified by Burnham & Anderson 2002). We therefore adopt the NFW model as our default in the following analysis.

We note that the additional complexity of the DK14 model captures a splashback-like feature at $R \sim 2\,\mathrm{Mpc}$ (see Appendix A.1). However, the BIC and AIC both disfavour the DK14 model, and the mentioned feature might be caused by noise or the projection of unrelated large-scale structure along the line of sight.

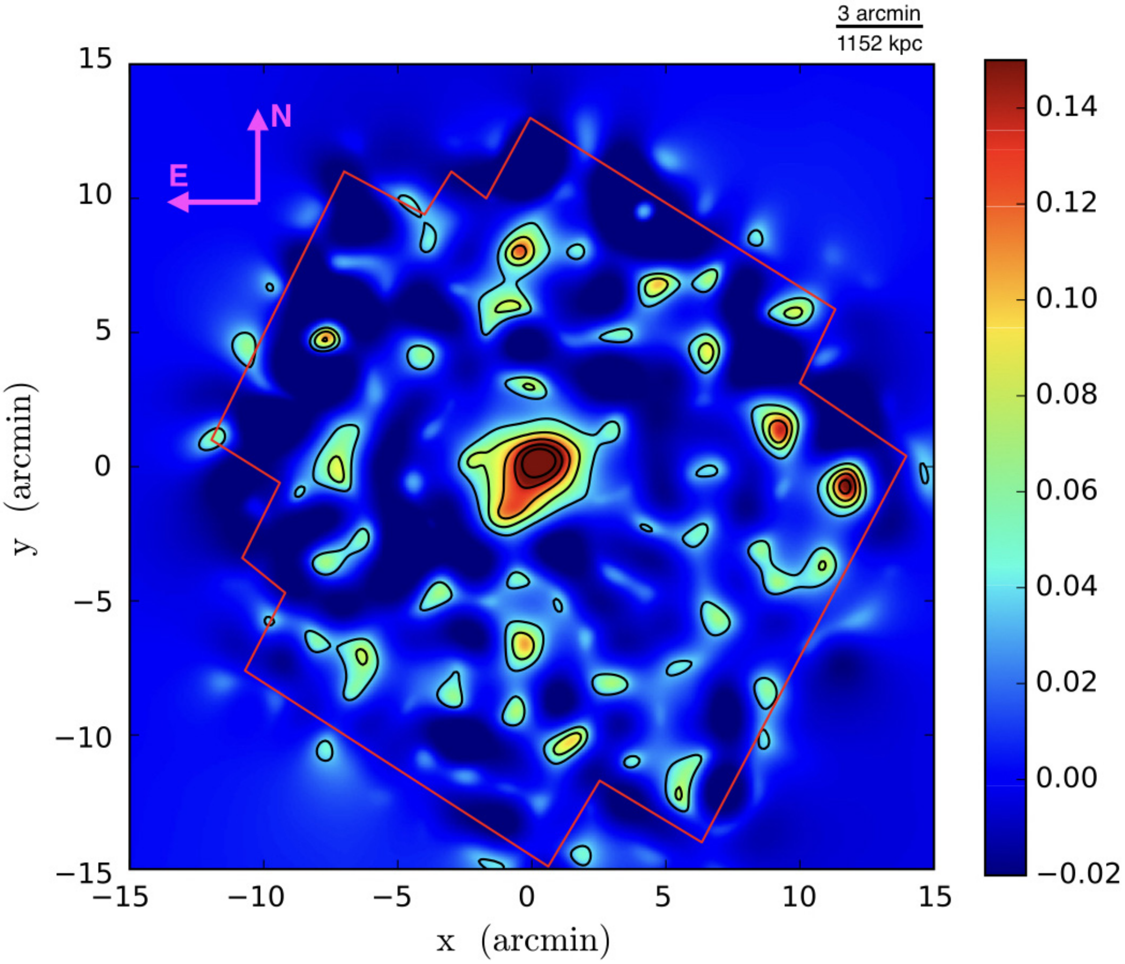


Figure 5.5: Convergence map of MS 0451 obtained with alternative method KS93+MRLens. The field is centred on (α =73.545202, δ =-3.0143863), with a red polygon indicating the extent of the *HST* imaging mosaic. Black contours show statistical significance thresholds, starting at $1\,\sigma_\kappa$ and spaced linearly in units of $1\,\sigma_\kappa$ above that. It is consistent with results of our Lenstool method (figure 5.4), but noisier and lower resolution.

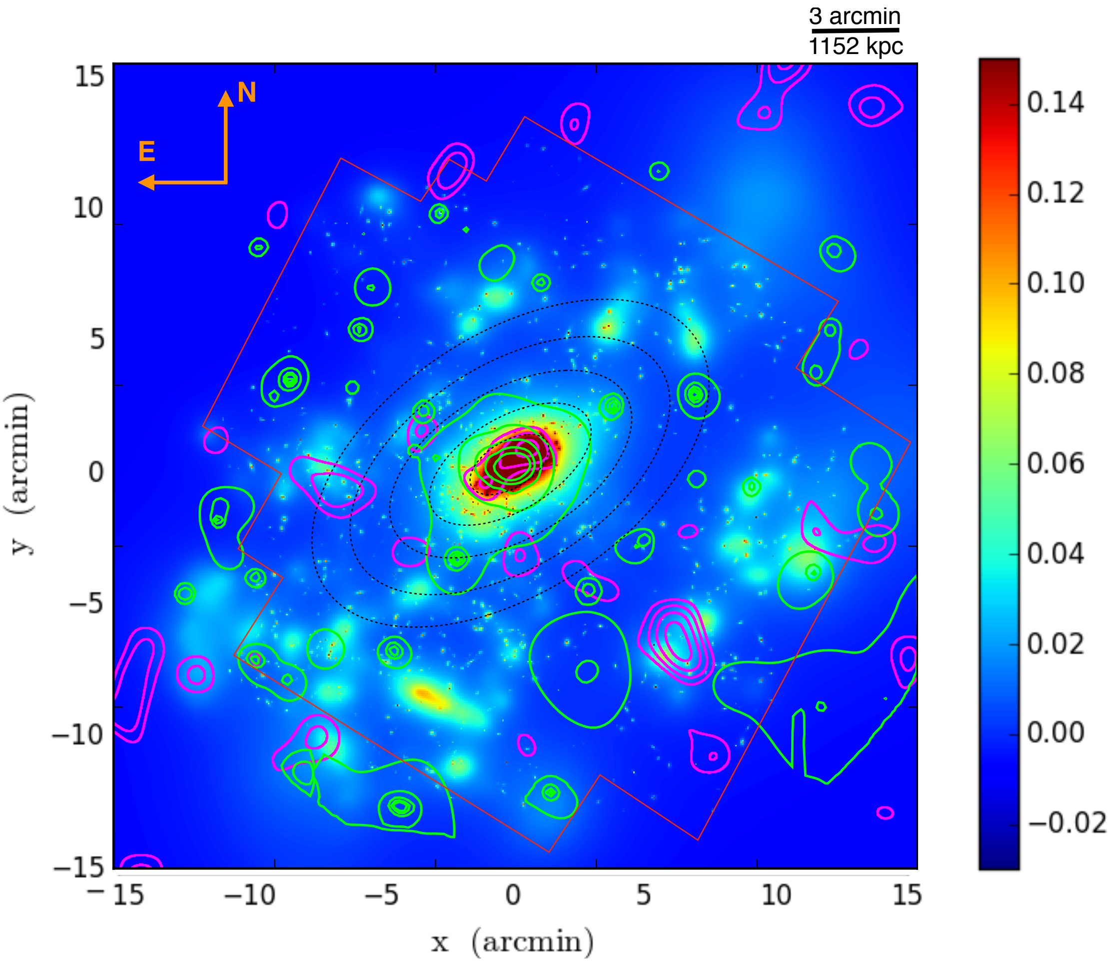


Figure 5.6: Alternative probes of the mass distribution around MS 0451, overlaid for ease of reference on the colour image from figure 5.4. Magenta contours show weak lensing measurements from ground-based observations (private communication N. Martinet), starting at $3\sigma_\kappa$ and in steps of $1\sigma_\kappa$, the rms uncertainty on convergence. Green contours show the X-ray surface brightness from *XMM-Newton*. Black ellipses show the shape of the eNFW model that best fits our LENSTOOL reconstruction within circular apertures of different radii (defined by the semimajor axis).

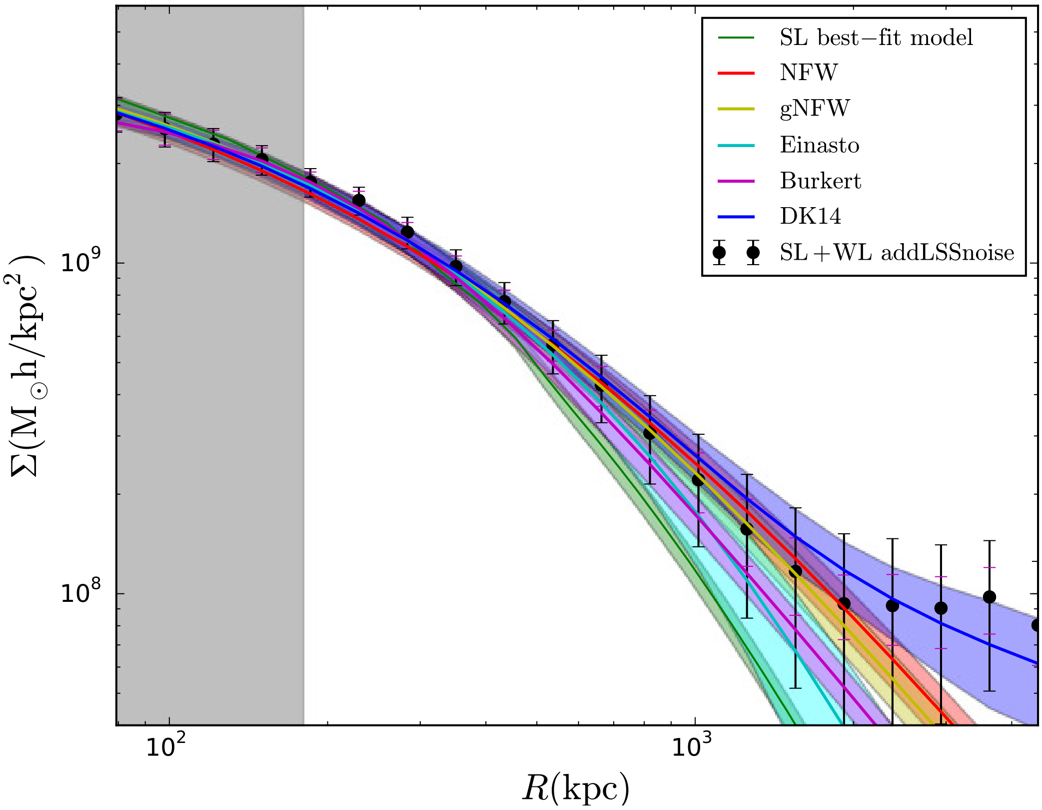


Figure 5.7: Azimuthally averaged 1D profile of mass in MS 0451 (black data points), from our combined strong+weak lensing analysis (figure 5.4). Error bars show the statistical uncertainty owing to galaxies' intrinsic shapes (inner) and also line-of-sight substructures (outer). The green curve shows the best-fit model using only strong lensing information (Jauzac et al. sub.), extrapolated beyond the multiple-image region (grey shaded area). Solid lines in other colours and their respective shaded areas show the mean and 68% confidence intervals from fits to various models.

Table 5.2: Marginalized posterior constraints on cluster model parameters, and the differences between their information criteria and those of the best-fit NFW model. The information criteria of an NFW are $\mathrm{BIC_{NFW}}$=13934.50, $\mathrm{AIC_{NFW}}$=13932.51 and $\mathrm{AICc_{NFW}}$=13933.22. Lower values indicate preferred models.

Models	$M_{200c}[10^{14}\,\mathrm{M}_\odot]$	c_{200}	Shape parameter	ΔBIC	ΔAIC	ΔAICc
NFW	16.51 ± 2.44	3.79 ± 0.36	–	0	0	0
gNFW	16.10 ± 1.94	4.47 ± 0.47	$\alpha = 0.57 \pm 0.20$	0.51	-0.48	0.02
Einasto	14.32 ± 2.67	4.26 ± 0.50	$\alpha_E = 0.42 \pm 0.11$	0.02	-1.00	-0.51
Burkert	13.62 ± 1.62		$r_\mathrm{core} = (230 \pm 20)\,\mathrm{kpc}$	7.60	6.60	7.10
DK14a				9.60	4.60	12.90
X-ray	17.47 ± 7.61	$2.35^{+0.89}_{-0.70}$	–			

Parameters of the DK14 model are excluded from this table for clarity. These are listed in Table A.1.

5.6.3 Halo shape

We measure the projected shape of MS 0451 by fitting the 2D mass distribution inside a circular aperture with an eNFW model. This approach yields results that are consistent with the previous 1D fit (Sect. 5.6.2): for the region inside $R < 3.24\,\mathrm{Mpc}$, we obtain $M_{200c} = (1.57 \pm 0.14) \times 10^{15}\,\mathrm{M}_\odot$ and $c_{200} = 3.7 \pm 0.4$. The best-fit axis ratio $q = b/a$ varies as a function of radius, from $q = 0.48 \pm 0.01$ within $R < 649\,\mathrm{kpc}$ to $q = 0.57 \pm 0.03$ within $R < 3.24\,\mathrm{Mpc}$. The cited statistical uncertainty may be an underestimate because we have neglected correlations between adjacent pixels in our error model of the mass map and, in the cluster core, because of our use of fixed strong-lensing potentials during MCMC parameter search. The axis ratio is consistent with simulations of general clusters (Jing & Suto, 2002; Suto et al., 2016, Tam et al. sub.), but smaller than the value of $q = 0.72$ ($649\,\mathrm{kpc} < R < 974\,\mathrm{kpc}$) measured from ground-based lensing observations by Soucail et al. (2015). This discrepancy might be explained by the large smoothing kernel used by Soucail and coworkers to reconstruct the mass distribution, which artificially circularises the data. Indeed, our results more closely resemble those from lensing analyses of large cluster samples, including Oguri et al. (2010), who found $\langle q \rangle = 0.54 \pm 0.04$ for 18

X-ray luminous clusters at $0.15 < z < 0.3$, and Umetsu et al. (2018), who found $\langle q \rangle = 0.67 \pm 0.07$ for the CLASH sample of 20 massive clusters.

At all radii, we find that MS 0451 is elongated roughly along a North-West to South-East axis (Figure 5.6), with a mean orientation $\sim 31.9^\circ$ counter-clockwise from East. The ~10% variation in this angle between the inner ($R < 640\,\mathrm{kpc}$) and outer halo ($R < 3.24\,\mathrm{Mpc}$) agrees well with typical clusters in both simulations (Despali et al., 2017) and observations (Harvey et al., 2019).

5.6.4 Baryonic components

Distribution of baryons

To measure the cluster's electron density profile, we apply the non-parametric "onion peeling" algorithm (Kriss et al., 1983) and the multiscale decomposition technique (Eckert et al., 2016) to the X-ray data (figure 5.8). Both methods assume spherical symmetry, and both yield consistent results. We find that the distribution of baryons is different to that found by our lensing analysis. It shows a constant-density core, flatter than both our lensing results and the distribution of gas in a typical massive cluster from the X-COP low redshift sample (Ghirardini et al., 2019).

To measure the cluster's temperature profile, we fit a single temperature plasma emission model to the X-ray spectra extracted in 6 concentric annuli spanning the radial range 0–1.5 Mpc. We find that the temperature of the X-ray emitting gas decreases from ~9 keV in the core to ~6 keV in the outskirts, consistent with the 'universal' thermodynamic profile of X-COP clusters.

In a separate analysis of the X-ray data assuming hydrostatic equilibrium and spherical symmetry, we measure a hydrostatic mass $M_{500,\mathrm{HSE}} = (1.06 \pm 0.35) \times 10^{15}\,\mathrm{M}_\odot$, and concentration $c_{200,\mathrm{HSE}} = 2.35^{+0.89}_{-0.70}$. The concentration is again lower than the one we obtain with the lensing analysis. Extrapolating the best-fit model to large radii yields a total mass $M_{200c,\mathrm{HSE}} = (1.75 \pm 0.75) \times 10^{15}\,\mathrm{M}_\odot$, which inevitably has large uncertainties because the X-ray emission at these radii is faint. The assumption of hydrostatic equilibrium may not be appropriate for this cluster. Deeper X-ray imaging and/or constraints on the Sunyaev-Zel'dovich signal are required to quantify the level of non-thermal pressure support.

The radial entropy profile of the intra-cluster gas (figure 5.9), obtained by combining the measured spectroscopic temperature with the gas density, is consistent with the 3D entropy model recovered from the *backwards NFW fit* under the assumption of HSE. We find a strong entropy excess in the cluster core, compared with the entropy of the fully relaxed gas calculated from the gravitational-collapse model (Voit, 2005). This large entropy excess confirms that MS 0451 does not contain a cool core.

Baryonic mass fraction

To measure the gas mass fraction $f_{\rm gas}$, we first integrate the non-parametric gas profiles (which do not assume hydrostatic equilibrium) and obtain a total gas mass $M_{\rm gas,500} = (1.29{\pm}0.15){\times}10^{14}\,{\rm M}_\odot$ inside a sphere of radius $R_{\rm 500,HSE} = 1.28{\pm}0.14\,{\rm Mpc}$. Dividing this by the total mass M_{500} of the NFW model that best fits the lensing data inside a sphere of radius $R_{500} = 1.30 \pm 0.06\,{\rm Mpc}$ indicates $f_{\rm gas,500} = (11.6 \pm 2.1)\%$. Our separate analysis assuming hydrostatic equilibrium yields $f_{\rm gas,500,HSE} = (12.2 \pm 4.3)\%$.

To measure the stellar mass fraction, we use the stellar mass to light ratio of quiescent galaxies

$$\log_{10}(M_*/L_K) = a\,z + b, \tag{5.6.10}$$

where $a = -0.18 \pm 0.04$ and $b = +0.07 \pm 0.04$ (Arnouts et al., 2007), assuming a Salpeter (1955) initial mass function (IMF)[4]. Applying this to all 1277 cluster member galaxies in the *HST*/ACS mosaic yields a mean value of $\langle M_*/L_K \rangle = 0.94 \pm 0.003$, or total stellar mass $M_{*,500} = (3.37 \pm 0.03) \times 10^{13} M_\odot$, where the uncertainty is the error on the mean. This is integrated inside a cylinder of radius R_{500}, rather than a sphere – but the stellar mass is so centrally concentrated that it should make little difference, to an already small number. Hence we adopt a stellar mass fraction

---

[4] To convert from Salpeter to a Chabrier (2003) IMF, we adjust stellar masses by 0.25 dex and find $f_{*,500} = (1.6 \pm 0.24)\%$, similar to $f_*{\sim}1.5\%$ measured in the wide-field *HST* COSMOS survey (Leauthaud et al., 2012).

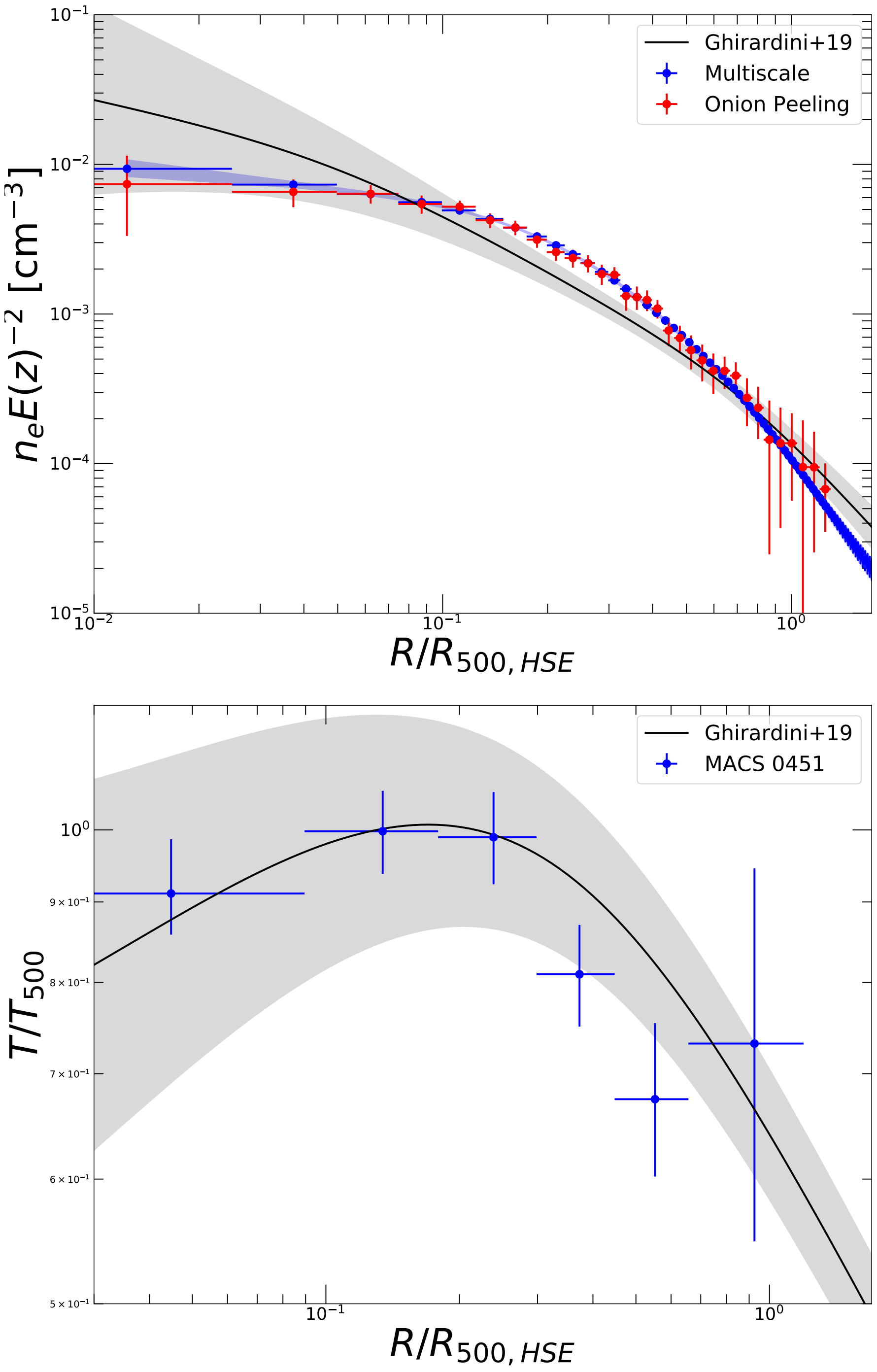


Figure 5.8: Thermodynamic profiles of MS 0451's Intra-Cluster Medium (ICM), scaled according to the self-similar model (Kaiser, 1986). *Top:* Deprojected electron density profile of the cluster computed using onion peeling (red), and multiscale decomposition (blue) methods. *Bottom:* Spectroscopic temperature profile of the cluster (blue). In both panels, the black curve and gray shaded areas show the mean profile and 1σ scatter of the X-COP sample of low redshift massive clusters (Ghirardini et al., 2019), for comparison.

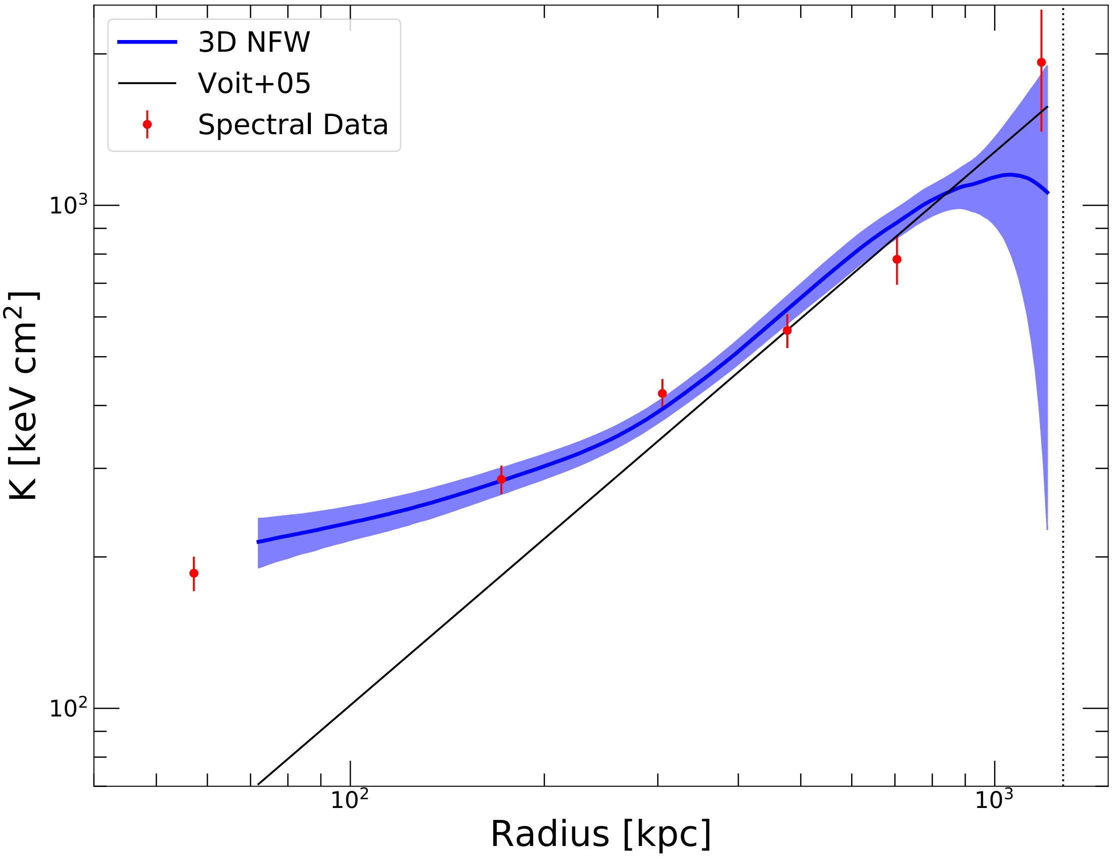


Figure 5.9: Radial profile of gas entropy. The red data points are obtained from the measured spectroscopic temperature and the gas density. The blue curve is the model optimised from the *backwards* approach. For comparison, the black curve shows the gas entropy predicted from the gravitational collapse model (Voit, 2005).

Substructure	R.A.	Dec.	$\langle z \rangle$	$M_{\rm tot}[10^{13}\,{\rm M}_\odot]$	$M_{\rm stellar}\ [10^{11}\,{\rm M}_\odot]$	Detection S/N
Sub1	4:54:26.917	-2:59:39.894	0.62	6.17±2.70	6.59±1.79	3.76
Sub2	4:54:39.389	-3:00:32.808	0.58	8.34±3.30	12.70±2.26	5.09
Sub3	4:54:15.278	-3:03:11.620	0.61	13.50±2.90	31.25±3.04	8.23
Sub4	4:54:26.088	-3:05:37.949	0.63	7.17±3.23	5.83±1.56	4.37
Sub5	4:54:11.745	-3:07:30.042	0.55	6.12±2.96	10.55±2.13	3.73
Sub6	4:54:37.972	-3:07:33.134	0.56	8.42±3.37	8.05±2.28	5.13

Table 5.3: Confirmed substructure detections in MS 0451 containing cluster member galaxies and with detection S/N>3. Columns show the location of each mass peak, the mean redshift of member galaxies within a 480 kpc aperture, the lensing and stellar masses integrated within the same aperture, and the signal-to-noise ratio of its detection, using the mean noise level of the mass map (Sect 5.4.3).

$f_{*,500} = (3.0 \pm 0.4)\%$.

Combining these measurements, we obtain a total baryonic fraction, $f_{\rm b,500} \equiv f_{*,500} + f_{\rm gas,500} = (14.6 \pm 1.4)\%$. This value is consistent with the mean cosmic baryonic fraction, measured as $f_{\rm b} = (14 \pm 2)\%$ from the outskirts of $z < 0.16$ clusters (Mantz et al., 2014), or $f_{\rm b} = (15.6 \pm 0.3)\%$ from the Cosmic Microwave Background (Planck Collaboration et al., 2016).

5.6.5 Group-scale substructures

To study the low density environment of large-scale structures surrounding MS 0451, we subtract the strong-lensing potentials from the Lenstool convergence map (figure 5.10). Outside the main halo, we detect 14 weak lensing peaks with S/N> 3 integrated within circular apertures of radius $R = 480\,{\rm kpc}$. To determine whether these 14 overdensities are at the redshift of the cluster, we assess the redshift distribution of galaxies inside those apertures with spectroscopic or (mainly) photometric redshifts (Appendix A.2). Galaxies along the line of sight to substructures 1, 2, 3, 4, 5 and 6 have a redshift distribution that peaks between $0.48 < z < 0.61$. We infer that these are part of the extended cluster, while others are projections of structures at other redshifts along a similar line of sight. Their total masses and stellar masses are listed in Table 5.3.

Previous ground-based weak lensing analyses identified only substructures 1 and 2 (Martinet et al., 2016), or substructure 2 at sub-threshold 2σ significance (Soucail et al., 2015). Our identification of 12 significant new structures, none of which have a counterpart in previous analyses, demonstrates the unique ability of space-based imaging to detect weak lensing in low-density environments.

The mismatch between our lensing and X-ray analyses is puzzling. The current depth of the *XMM-Newton* imaging should be sensitive to $\sim 10^{14}\mathrm{M}_\odot$ halos. However, we detect only faint X-ray emission for substructure 6. We detect brighter – but misaligned – emission near substructures 2 and 5, and between substructures 3 and 4. A first hypothesis that could explain this discrepancy, is that selection biases affect one or both of our lensing and X-ray analyses. Substructure 3 is the most massive ($M_{\rm tot} = (1.3 \pm 0.3) \times 10^{14}\,\mathrm{M}_\odot$ inside a 480 kpc aperture), but also the closest to the cluster core. If the main cluster is imperfectly modelled and subtracted, its residual projected mass could artificially boost the lensing signal of substructure 3. Indeed, all the substructures are closer to the cluster's major axis than to its minor axis, and the lensing signal from all of them could be biased high. Conversely, proximity to the cluster core provides a high X-ray background, which potentially lowers the signal-to-noise of the X-ray emission below our detection threshold. A second hypothesis is that substructures within $R_{\rm 200m} = (2.51 \pm 0.14)$ Mpc are probably also inside the 3D splashback radius of MS 0451 (More et al., 2015a). Therefore, they might have already passed through the pericentre. Ram-pressure stripping during their motion through the main halo could have removed much of their hot gas, and thus reduced their X-ray luminosity.

5.6.6 Filaments

Alignments of substructures

Based on the distribution of substructures around MS 0451, we propose that three filaments are connected to the cluster core (shown as green lines in figure 5.10). The first of these possible filaments extends East of the cluster, encompassing Substruc-

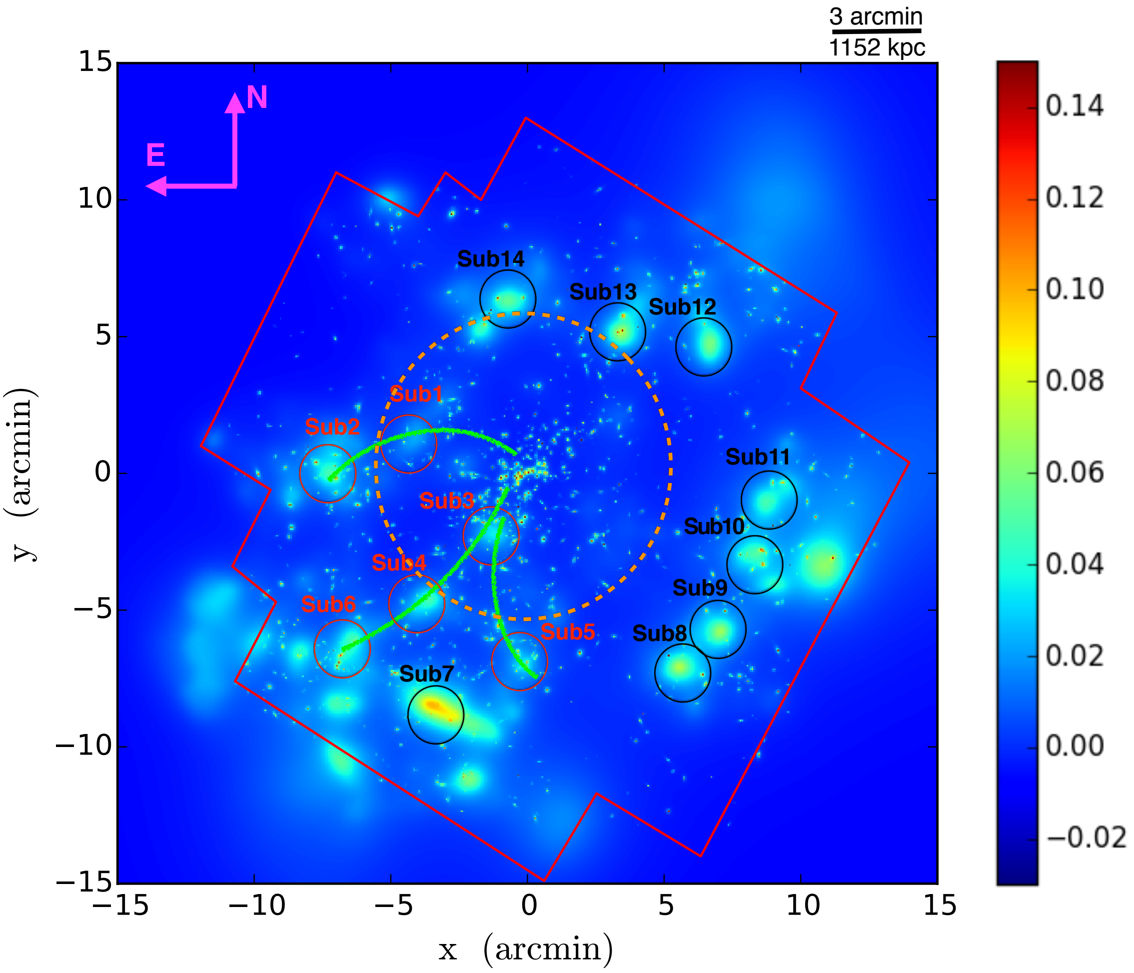


Figure 5.10: The low density environment surrounding MS 0451. The colour image shows lensing convergence with SL potentials subtracted: all the remaining signal was constrained by the potential grid and cluster member galaxies. The dashed orange circle has radius $R_{200c} = 1.99\,\mathrm{Mpc}$. Smaller circles (with radius 480 kpc) mark substructures with a projected mass $> 3\,\sigma_M$ inside that aperture; red circles have optical counterparts at the cluster redshift. Green lines suggest the extent and direction of possible large-scale filaments.

tures 1 and 2 and containing mean convergence $\langle\kappa\rangle = 0.022 \pm 0.006$. The second points South-East, encompassing Substructures 3, 4, and 6, with $\langle\kappa\rangle = 0.033 \pm 0.007$. The third, finally, turns South, from Substructure 3 to Substructure 5 and also has mean convergence $\langle\kappa\rangle = 0.033 \pm 0.007$. For each of these three candidate filaments the density contrast exceeds the threshold value of $\kappa = 0.005$ defined in Section 4.3.1 to identify filaments, and each has a mean *excess convergence* greater than 0.02, even after subtracting the smooth, cluster-scale mass distribution.

All three possible filaments point in a similar direction, close to the main cluster's South-East/North-West major axis. We detect no substructures in the opposite direction along the same axis (with the possible exception of an unconvincing feature just outside the *HST* mosaic to the North-West). This is strikingly different from the typical distribution of mass in cosmological simulations, which usually show a symmetry of infalling material along both directions of a cluster's major axis, as the system grows and becomes increasingly elongated as the result of gradual, continuous accretion along filaments.

Aperture multipole moments

Extended structures can also be identified with the measurement of aperture multipole moments (AMMs) of the 2D mass distribution (Schneider & Bartelmann, 1997). We applied the optimal filter developed in Section 4.6.2 on the strong lensing subtracted convergence map of MS 0451 (figure 5.10). The resulting Q map is shown in figure 5.11. We estimate the noise level by defining σ_Q as the standard deviation of all pixels in the Q map.

Although the signal-to-noise ratios for Q are low, the three possible filaments proposed in Sect. 5.6.6 are also highlighted by the AMM filter, with signal-to-noise ratios of $\sim 2-3$. Additional extended structures may exist at other redshifts. In particular, Substructures 8, 9, and 11 might form a linked system at $z \sim 0.7$ behind the cluster (Appendix A.2).

While the results presented in figure 5.11 are promising, the limiting, single-orbit depth of the available *HST* observations prevents us from drawing firm conclusions from this measurement. Deeper space-based observations obtained by future surveys

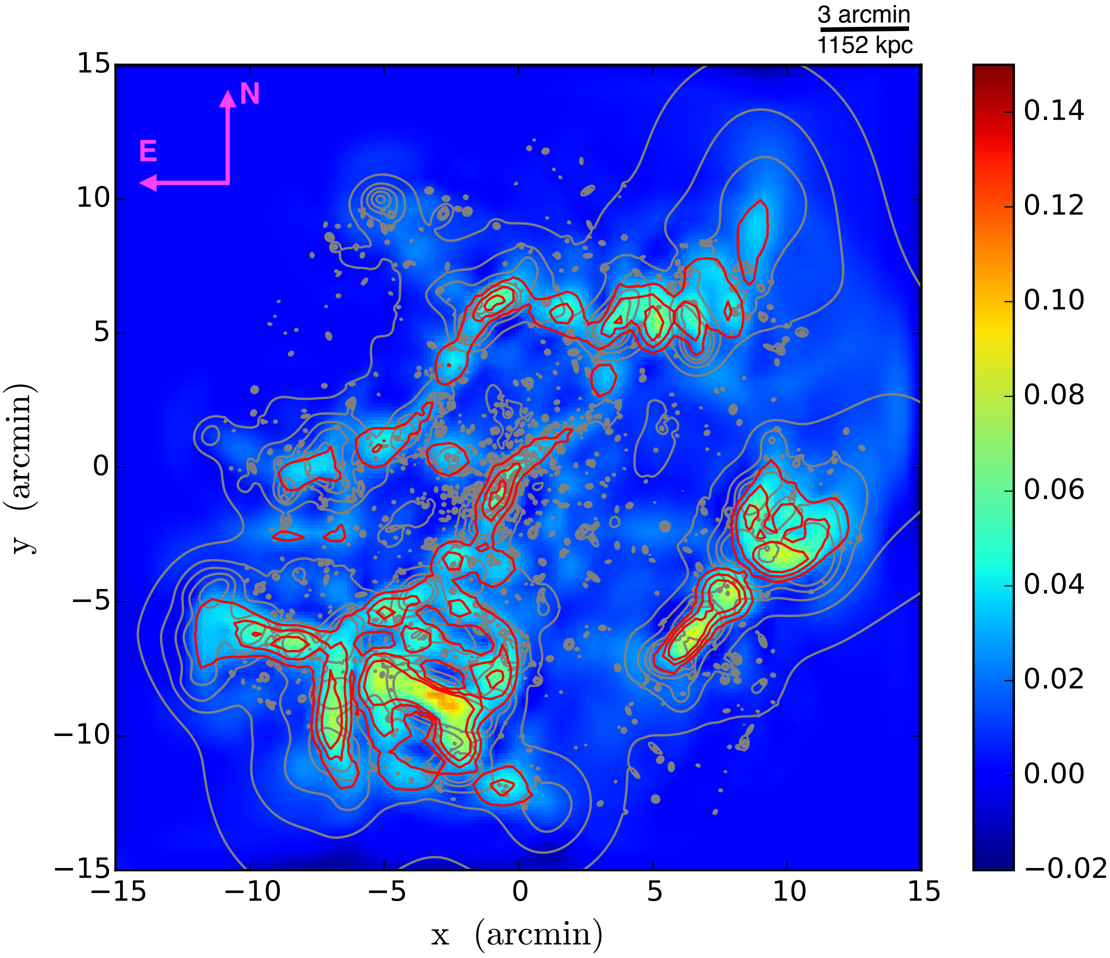


Figure 5.11: The low density environment around MS 0451 (figure 5.10), filtered using aperture multipole moments to search for extended, filamentary structures. Red lines show contours of signal-to-noise, starting at 2 and increasing in steps of 1. Grey lines are contours of the unfiltered mass distribution.

will increase the density of detected background galaxies and reduce the noise level of this technique.

5.7 Discussion: inferred dynamical state

According to N-body and hydrodynamical simulations (e.g. Nelson et al., 2011; Poole et al., 2006; Ritchie & Thomas, 2002), mergers between two clusters (classified by their masses as the primary and secondary component) proceed in five distinct stages: pre-interaction, first core-core interaction, apocentric passage, secondary core accretion, and relaxation. After the first core-core interaction, the gas of the two merging halos (including the remaining part of the secondary component's cool core) moves outward. After the two cores reach maximum separation, the secondary core falls back toward the primary core and is accreted. Finally the system evolves into a single merger remnant.

Three lines of evidence suggest that MS 0451 is in a post-collision state, approximately 2–7 Gyr after the first core passage of two progenitors that are now approaching second apocentre:

- While we observe a bimodal (Sect. 5.4.2) and elongated (Sect. 5.6.3) distribution of dark matter, we find a spherical distribution of gas with almost constant central density (Sect. 5.6.4). Such a contrasting configuration is seen in simulations of major mergers 2–7 Gyr after first core-core interaction when, "following the merger, the resultant system settles into virial equilibrium sooner than into hydrostatic equilibrium" (Poole et al., 2006). This period, in which the gas has not yet had time to fully relax and settle into the gravitational potential of the combined halo, represents the second infall phase before the system's final relaxation. These findings apply to a wide range of initial conditions regarding the progenitor mass ratio ($M_{\text{primary}} : M_{\text{secondary}}$ =1:1, 3:1, 10:1) and the ratio between the secondary's transverse velocity and the primary's circular velocity, a quantity that affects impact parameters ($v_t/v_c = 0.0, 0.15, 0.45$).
- In simulations, merging increases the entropy of gas in the cluster core, leading to a large core radius and low concentration (Ritchie & Thomas, 2002), exactly

as seen in MS 0451 (figure 5.9).

- The cluster is connected to its large-scale environment through a number of substructures (Sect. 5.6.5) and possible filaments (Sect. 5.6.6). The distinctly asymmetric distribution of these features differs starkly from that of simulated clusters which grow through smooth, continuous accretion without strong directional preferences, and whose major axes are aligned with filaments in opposite directions.

The dynamical history inferred from our analysis of MS 0451 provides a possible explanation for the star-formation history observed in this system. In a comparative study of massive clusters, Moran et al. (2007a,b) used optical and near-UV spectroscopy of passive spirals within 1.5 Mpc of the cluster core to conclude that the star formation in MS 0451 was abruptly quenched at a redshift of $z = 2$, i.e., $\sim$5 Gyr before the redshift of observation, consistent with our estimate of the temporal evolution of the merger event. Moran et al. (2007b) ascribe the sudden cessation of star formation to ram-pressure stripping by a particularly dense ICM, as evidenced by MS 0451's bright, extended X-ray emission observed today. The merger scenario proposed by us lends strong support to this explanation, by adding a contemporaneous second component of intra-cluster gas, moving at high relative velocity through the cluster core and thus dramatically increasing the ram pressure (c.f. Fujita et al., 1999; Kronberger et al., 2008; Vijayaraghavan & Ricker, 2013). Beyond 1.5 Mpc, the passive spirals in MS 0451 show a "starvation-like" gradual cessation of star formation, consistent with secular pre-processing in infalling groups. From their classification of galaxy morphologies in the *HST* imaging data, Moran et al. (2007b) also concluded that passive spirals are all but absent inside the central 600 kpc, having evolved into S0 galaxies. This finding too could be the result of ongoing, enhanced ram-pressure stripping by the current ICM as galaxies fall towards the cluster core. Alternatively, the lack of passive spirals within the inner regions of MS 0451 could be a residual indication of the merging subhalo's trajectory, as it takes $\sim$ 400 Myr for a galaxy to travel $\sim$600 kpc across MS 0451. Our inferences about MS 0451's merger dynamics thus complement and support the conclusions of previous studies of its star-formation history.

5.8 Conclusions

We present the first combined strong- and weak-gravitational lensing analysis of the massive galaxy cluster MS 0451, exploiting the largest mosaic of *HST* imaging around any massive cluster: 41 ACS pointings covering an area of $\sim 20 \times 20\,\mathrm{arcmin}^2$ ($\sim 6 \times 6\,\mathrm{Mpc}^2$). The strong-lensing model exploits 16 multiple-image systems, and our weak-lensing analysis uses a catalogue of 20,138 background galaxies ($\sim$44 arcmin^{-2}). We combine these constraints using the LENSTOOL multi-scale grid technique.

The reconstructed mass distribution of MS 0451 reveals a bimodal cluster core, elongated along the South-East to North-West axis and surrounded by six substructures, as well as eight weak-lensing peaks created by mass concentrations at other redshifts projected along our line of sight. We find a total mass of the system of $M_{200} = (1.65 \pm 0.24) \times 10^{15}\,\mathrm{M}_\odot$ with an NFW concentration of $c_{200} = 3.79 \pm 0.36$ (the gNFW and Einasto models yield similar results, while a Burkert model is disfavoured). Our mass map is consistent with that of the most recent ground-based weak-lensing analysis (Soucail et al., 2015) but resolves three times more substructures at equivalent significance of detection. The mass distribution of MS 0451 becomes more circular at large radii, parameterized by an axis ratio that decreases from $q = b/a = 0.48 \pm 0.01$ within a projected radius of $R = 640\,\mathrm{kpc}$ to $q = 0.57 \pm 0.03$ inside $R = 3.2\,\mathrm{Mpc}$. A flattening of MS 0451's density profile at $R \approx 2\,\mathrm{Mpc}$ is well fit by the splashback feature in the DK14 model. However, this model's additional complexity negates the improved fit according to Bayesian Information Criteria; the aforementioned flattening may thus just be noise or due to large-scale structure projected from other redshifts.

In our X-ray analysis, we measure a baryonic-mass fraction of $f_{\mathrm{b},500} = (14.6 \pm 1.4)\%$ for MS 0451, consistent with the cosmic baryon fraction (Planck Collaboration et al., 2016), and a total mass of $M_{200} = (1.75 \pm 0.75) \times 10^{15}\,\mathrm{M}_\odot$, in good agreement with the lensing estimate. We note though that the assumption of hydrostatic equilibrium underlying the X-ray mass measurement is unlikely to be valid, given that the cluster's gas is distributed very differently than the dark matter. The distribution of gas is circularly symmetric, with a constant-density core and low concentration, $c_{200,\mathrm{HSE}} = 2.35^{+0.89}_{-0.70}$. We also find a strong excess of gas entropy in the cluster's central 300 kpc.

Similarly contrasting distributions of gas and dark matter are seen in simulations of post-merger clusters (Poole et al., 2006). The matter distribution observed in MS 0451 suggests that the cluster underwent a major merger ~2–7 Gyr ago, and that the two dark-matter halos in the centre are now approaching second apocentre. This merger would have quenched star formation, as ram-pressure from the dense ICM stripped cold gas from cluster member galaxies. Thus the evidence from gravitational lensing, X-ray emission, optical photometry, and spectroscopy all point to a consistent dynamical history.

We find tentative evidence of three filaments extending from the cluster. The distribution of substructures and a noisy measurement of aperture multipole moments indicate that all three point in similar directions, between East and South. Interestingly, their distribution is asymmetric, with no counterparts to the North or West. Aperture multipole moments appear to be a promising method to detect extended filaments. However, our measurements based on single-orbit *HST* data are dominated by shape noise, and deeper space-based observations will be necessary to robustly test this method.

In the next decade, wide-field, space-based surveys at high resolution are planned as part of the *Euclid* and the *Nancy Grace Roman Space Telescope* missions, as the Vera C. Rubin Observatory becomes operational on the ground. MS 0451 will be an ideal target for future studies to characterise infalling substructures along filaments, the timing of star-formation processes after a major merger, and the late-time evolution of cluster collisions.

Chapter 6

Ongoing and Future Work

6.1 Introduction

This chapter presents two ongoing projects that will be useful for future studies. In the near future, MS 0451 will not be the unique cluster having wide field space-based imaging because there will soon be similar imaging taken around 6 more clusters through the *HST*/BUFFALO survey (Steinhardt et al., 2020), 200 more from the balloon-borne telescope SuperBIT (Super-pressure Balloon-borne Imaging Telescope; Romualdez et al., 2019), and 10,000 from *Euclid* (Laureijs et al., 2011). In the next decade, 40,000 clusters will be observed to even greater depth by the *Nancy Grace Roman Space Telescope* (Spergel et al., 2013). Such a large amount of data can enhance the statistical significance of cluster studies and lead to significant contribution to the cluster community.

Astronomy is therefore entering the big data era. To properly deal with this large amount of information, fast and sophisticated methods will be needed. Machine learning offers us an automated approach to perform pattern recognition. The first part of this chapter will present a new technique of denoising weak lensing maps using Convolution Neural Networks (CNN) that can be applied to future sky surveys.

The second part of this chapter is devoted to one of the potential future surveys that will be provided by SuperBIT. Traditional space-based telescopes are extremely expensive. Alternatively, balloon-borne telescopes operating above 99.7% of the Earth's atmosphere provide space-like resolution, at a fraction of the economic cost typical of space-based instrumentation. The low cost and effective observational capabilities of these balloon-borne platforms will benefit the astronomy community. After introducing SuperBIT, we also show its preliminary observational results.

6.2 Denoising Lensing Mass Maps with Deep Learning Approach

As shown in Chapter 2, the reconstruction of precise mass distributions from real observations is still a challenging task. The main obstacle is the random noisy weak lensing peaks induced by the intrinsic shape noise of background galaxies dominant in weak lensing analyses. These spurious peaks could bias the inference of the properties of galaxy clusters (e.g. Hamana et al., 2012), the identification of subhalos and diffuse structures of filaments, and weak lensing peak statistics (Bard et al., 2013). There are many mass mapping techniques which attempt to suppress these noises, with different degrees of success, but falling short of producing a thoroughly noiseless mass map. Today, machine learning algorithms are widely used for a variety of tasks in astronomy (Baron, 2019), especially in the lensing community. For strong lensing, machine learning algorithm are widely adopted to search for strong-lens systems (Metcalf et al., 2019) and detect subhalos (Brehmer et al., 2019). For weak lensing, deep learning approach is used in galaxy shape measurements (Tewes et al., 2019) and in automatically denoising weak lensing mass maps for cosmological constraints (Jeffrey et al., 2020; Shirasaki et al., 2019). The latter is first proposed by Shirasaki et al. (2019), who developed the Conditional Adversarial Networks (CANs) to reduce shape noise. Their reconstructed one-point probability density function of the convergence and power spectra are in good agreement with the true counterparts, which means that the cosmological information imprinted in the convergence maps can partially be extracted by their networks. However, their networks are not able to disentangle galaxy clusters from noisy peaks, which could cause bias(es) on the estimation of local convergence. This thesis concentrates on cluster lensing, where local properties of the cluster field are important. We have therefore started to develop a denoising method for weak lensing mass maps with a deep learning approach. In particular, the underlying gaussian noise peaks can be recognised and removed by a convolution neural network (CNN) that has been trained on a number of noisy mass maps beforehand. The deep learning networks on which I am working now mainly focus on denoising cluster field lensing maps, which provide us with an unbiased estimation of cluster properties including radial density profiles and the morphology of dark matter distributions. In this section, I present how we create the training and testing samples, and describe the architecture of the network. We will then show some preliminary results.

6.2.1 Data: Elliptical NFW Halos

To show the feasibility of this project, we first create a data set using a simple analytical halo model. The data set consists of 8,000 mock eNFW halos (see Section 4.5.1) with mass range $M_{200} \in [0.1, 10] \times 10^{15}\,\mathrm{M}_{\odot}$, and concentration $c_{200} \in [2, 8]$. Their axis-ratio and orientation are randomly selected from flat distributions: $q \in [0.3, 1]$ and $\phi \in [0, 2\pi]$. The shear fields of eNFW halos require additional numerical integrals corresponding to the 2D Poisson equation (Keeton, 2001b), which is time-consuming for 8000 samples. We therefore directly convert the analytical convergence field into a shear field using equation 3.4.35 with zero-padding. We then separately apply two sets of random shape noises, σ_γ=0.26 and 0.36, on the shear fields assuming space-based resolution, 50 galaxies arcmin^{-2}. Shape noise with $\sigma_\gamma = 0.36$ is similar to the *HST* measurements near MS 0451-03 (see Section 4.3.4). We also apply a smaller level of shape noise, $\sigma_\gamma = 0.26$, to investigate its impact on the performances of the neural network. To convert the noisy mock lensing shear into convergence, we perform a simple KS93 inversion without applying any filters to suppress the noises. Two set of resulting pixelated convergence maps, dominated by random noise peaks, are then entered into the convolutional neural network.

We split the whole halo sample into a training set with 80% of the whole sample (6,400 clusters), and a testing set with 20% of the whole sample(1,600 clusters). The training set is first input to the network to build the pre-trained model while the testing set is used to evaluate the performance of this model.

6.2.2 Convolutional Neural Networks

Convolutional neural network (CNN) is a class of deep neural networks, which is particularly suitable for image recognition and image classification. CNN can capture the spatial dependencies in an image, without losing the local properties. Traditionally, a fully connected network would flatten a two-dimensional image to a long one-dimensional array, which destroys the spatial arrangement of the image. A fully connected network therefore needs more parameters to extract features from 2D/ 3D images. On the other hand, CNN can deal with images, retaining their spatial arrangement without flattening them. It defines a weight matrix (also called a filter), and the input image is convolved with this 2D matrix to extract specific features without losing the spatial information. In the meantime, CNN dramatically reduces the number of parameters we need to train for the network, and downsamples the input image.

CNN typically consists of a series of convolutional layers with filters performing linear operations; pooling layers to reduce the dimensionality of each map; and activation functions for non-linear operations. The convolutional layer is the first layer to extract features from an input image. It performs the convolution of the image matrix and a filter, resulting in a *feature map*. The pattern (parameters of the filter) is optimised by the network. A non-linear activation function is then applied to the *feature map* to introduce non-linearity in the network. Several activation functions are widely used, and we adopt the Leaky Rectified Linear Unit in our CNN. Rectified Linear Unit (ReLU), $f(x) = \max(0, \mathrm{x})$, is computationally efficient but could cause the *Dying ReLU problem*. When inputs approach zero, or are negative, the gradient of ReLU becomes zero, and the weight is therefore not updated. This ReLU neuron becomes inactive and only yields zero for any input. This situation is unlikely to recover. An alternative function that can solve this problem is Leaky ReLU, $f(x) = \alpha x$ for $x < 0$, where α is the parameter to be decided. It has a small slope for negative values instead of zero, precluding the possibility of zero gradients.

There are two other parameters for optimising a network: the *batch size* and the *epoch*. In machine learning, it is common to randomly divide a dataset into a number of batches, rather than pass the entire dataset into the network at once.

Batch size is the total number of training samples present in a single batch. One epoch means the passing of an entire training set (divided into smaller batches) through the neural network once. Updating the weight with a single pass/epoch is not enough because it would cause the underfitting problem, where a solution is too simple and the network has not been trained long enough to learn the relevant patterns in the training data. We therefore optimise the model with limited data sources by increasing the number of epochs. This can solve the underfitting problem, but training with too many epochs could also cause an overfitting, meaning that the network learns a complex solution from the training data which does not generalize to the testing data.

The Architecture of CNN

We use the U-Net encoder-decoder structure network. The noisy convergence map calculated from KS93 is entered to the encoder part which is composed of eight convolution layers. Each convolution layer consists of a convolution with a specific kernel size (32×32 for the first two layers, 16×16 for the 3rd and 4th ones, 8×8 for the 5th and 6th, and 4×4 for the final two layers), followed by an activation function of Leaky ReLU with $\alpha = 0.2$. We also include max pooling operation which takes the largest element from the feature map in order to downsample the images and dropout layers to prevent overfitting. When an input image passes through these layers, the encoder learns and extracts the important features at different scales and removes the underlying gaussian noise peaks from it. In the final layer of the encoder, the input image is compressed into a 16×16 matrix with 128 channels, which is then passed to the decoder. The decoder is the inverse operation of the convolution layer, expanding the compressed feature to the original dimension. With the process of size decreasing in encoder, some small-scale information would be lost. Therefore the skip connection of U-Net, which links the mirrored layers in the encoder with those in the decoder, is needed for transferring the small-scale information to the reconstruction of an output image. The network iteratively learns the parameters by minimising a mean-square-error (MSE) loss function, calculated as the average of the squared differences between the predicted and ground truth (noise-free) maps.

The optimization is performed with a batch size of 40 and 100 epochs.

6.2.3 Preliminary Results

In this section, we present the results of denoising weak lensing mass maps using our U-Net CNN. Figure 6.1 shows the convergence of an example of a cluster selected from the testing set. The prediction from our network can effectively remove the underlying noise fields (for both $\sigma_\gamma = 0.26$ and $\sigma_\gamma = 0.36$) and reconstruct the cluster in the corrected shape and orientation. Their corresponding radial density profiles are also shown in figure 6.1. The density profile calculated from the prediction is in good agreement with the true counterparts while the observed fields underestimate the mass due to the presence of random noise. To quantify the performance of the whole training and testing samples, we define two quantities: σ_κ and MSE. For σ_κ, as we defined in Chapter 2, we first calculate the residual maps $\kappa_{\rm res} = \kappa - \kappa_{\rm true}$. For each $\kappa_{\rm res}$, we then compute the noise level σ_κ, defined as the root mean square (rms) deviation from the mean of $\kappa_{\rm res}$, over all pixels (128×128) in the field. The distribution of σ_κ of the training sample and testing sample are shown in figure 6.2. The noise level in the reconstructed convergence map of cluster fields with $\sigma_\gamma = 0.36$ are higher than those obtained with $\sigma_\gamma = 0.26$. We also compute the MSE between the predicted mass map and the ground-truth,

$$\mathrm{MSE} = \frac{1}{\mathrm{N_{pix}}} \Sigma_{\mathrm{i=1}}^{\mathrm{N_{pix}}} (\kappa_{\mathrm{true,i}} - \kappa_{\mathrm{pred,i}})^2, \tag{6.2.1}$$

where $N_{\rm pix} = 128 \times 128$. The distribution of MSE for different samples are shown in figure 6.3. They have similar trends as σ_κ, where the networks pre-trained with training data added by higher shape noise $\sigma_\gamma = 0.36$ obtain higher MSE values. On this basis, we can predict that the network applied on lower resolution imaging data (such as ground-based images) which resolve fewer background galaxies will result in a less accurate reconstruction.

To quantify these comparison between different samples, we separately fit the distribution of σ_κ and MSE with a Gaussian function to obtain the mean value and the spread of the noise level for different cluster samples. The mean and standard deviation of the best-fit Gaussian are listed in Table 6.1. We find that, for a given shape noise, the noise level and MSE measured from the predictions of the testing set have a higher mean value and a wider distribution compared to those obtained from the training set. This result is expected as the parameters in the networks are optimized using the training sample. In addition, a larger shape noise does impact the performance of the reconstruction, resulting in a less accurate prediction.

6.2.4 Future plans

The network trained by simple eNFW halos shows the feasibility of this deep learning approach. However, to apply this method to real observational data, a more realistic dataset is needed. In the near future, we will use real projected cluster fields from N-body simulations (C-EAGLE) as a training set, with projected large-scale structures along the line of sight. Other weak lensing systematics, such as biases in galaxy shape measurements caused by instrumental effects (Massey et al., 2013; Tewes et al., 2019), and selection bias from real observations could also be taken into account in the application to real observations.

Generative Adversarial Networks (GAN) will be tested on the same subject. In GAN, the structure of the CNN will be improved by adding a discriminator which classifies the input image (an output image from the encoder-decoder network) as a ground-true or an output image generated from the encoder-decoder network. At each iteration, the generator (referring to encoder-decoder network) learns to make the discriminator classify its output as real, while the discriminator learns to distinguish the ground-true and the new output from the generator. Both networks try to beat each other. Eventually, the final output image becomes better and more realistic as the ground-true. The matured deep-learning network can be tested and applied to large datasets from current and future sky surveys (e.g. *HSC*, *LSST*, *Euclid*), constituting a new approach to precisely map lensing mass distributions and identify subhalos.

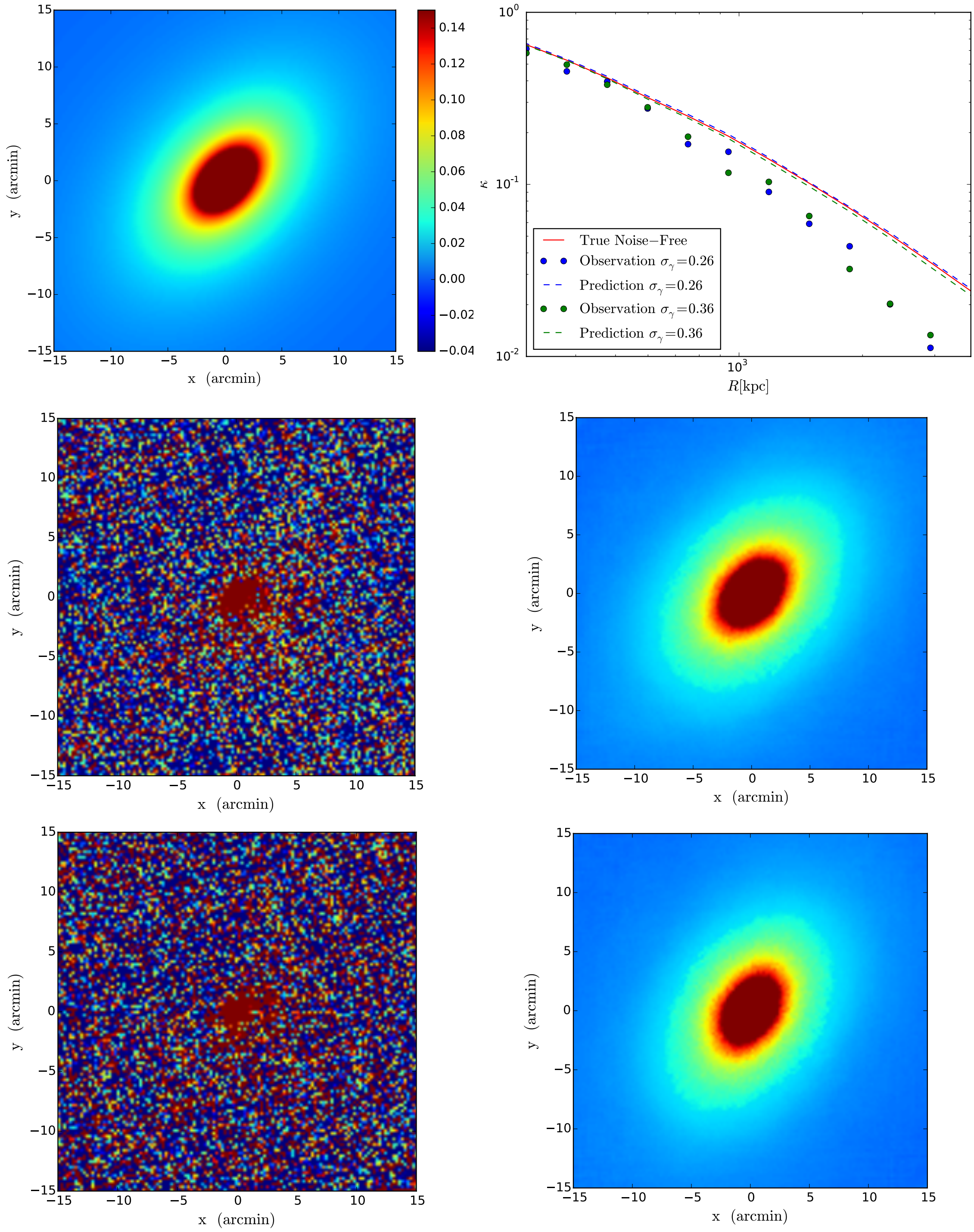


Figure 6.1: Convergence maps of a cluster from the testing set. *Top left*: the true (noise-free) convergence map. *Middle*: the observed (with shape noise $\sigma_\gamma = 0.26$) convergence map on left panel is entered into the CNN which outputs the prediction shown on the right panel. *Bottom*: the observed (with shape noise $\sigma_\gamma = 0.36$) convergence map on the left panel is input into the CNN which outputs the prediction on the right panel. *Top right*: the corresponding radial density profiles.

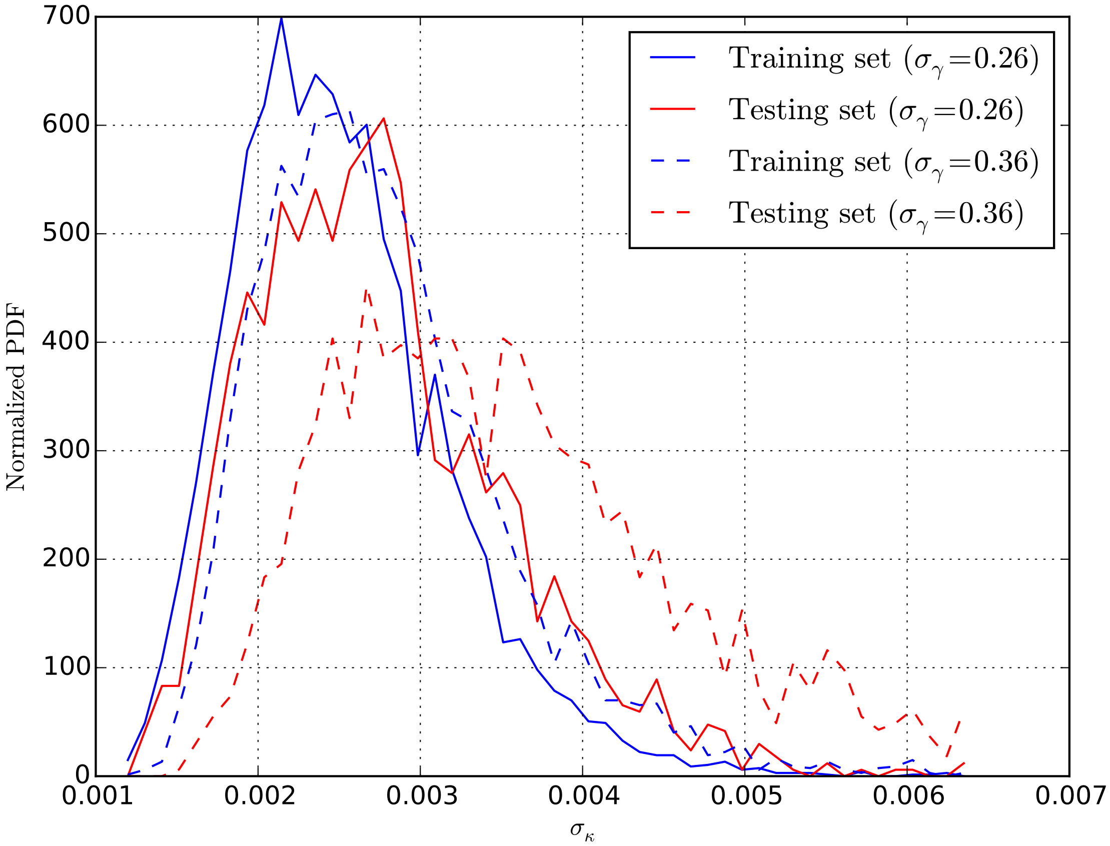


Figure 6.2: The normalized PDF distribution of noise level σ_κ. Solid (dashed) lines are the cases with shape noise $\sigma_\gamma = 0.26$ ($\sigma_\gamma = 0.36$). Blue represents the training set, and red represents the testing set.

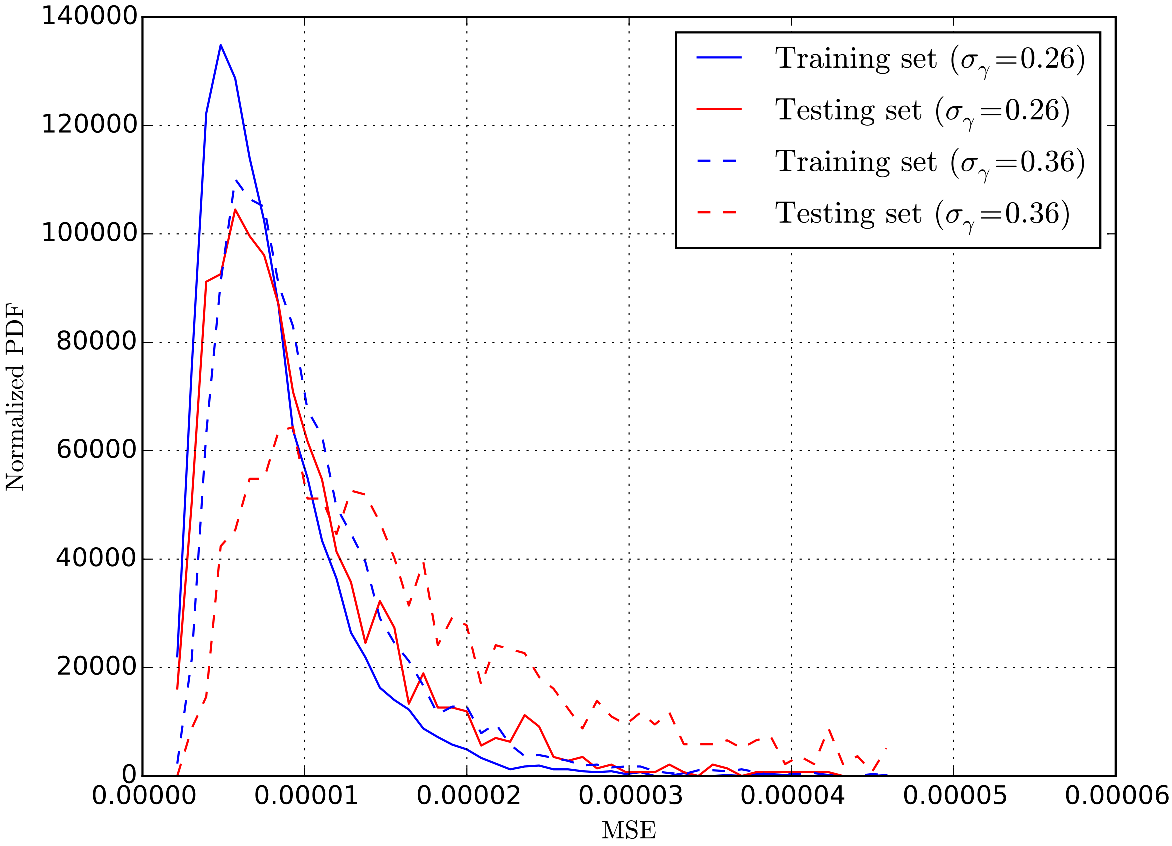


Figure 6.3: The normalized PDF distribution of MSE. Solid (dashed) lines are the cases with shape noise $\sigma_\gamma = 0.26$ ($\sigma_\gamma = 0.36$). Blue represents the training set, and red represents the testing set.

6.3 Future Wide-field Survey: SuperBIT

SuperBIT (Super-pressure Balloon-borne Imaging Telescope) is a diffraction-limited, wide-field (0.4 deg^2), 0.5m telescope operating in the visible-to-near-UV bands (300-900 um). Its main goal is to image up to 200 clusters over a single super-pressure balloon flight for strong and weak lensing studies. SuperBIT at 36 km altitude is above 99.7% of the Earth's atmosphere, offering a diffraction-limited observations and delivering space-like imaging at a cost which is economically efficient. Within the wavelength range, the high resolution and depth of SuperBIT imaging are sufficient to measure the weak lensing signal of distant ($z \sim 1$) galaxies behind foreground ($z \sim 0.3$) galaxy clusters. It also provides wide-field imaging data (25'×17'), $\sim$ 36 times larger than each *HST*/ACS pointing, allowing us to observe an entire cluster, including its connection to surrounding large-scale structure, in a single pointing.

On September 17, 2019, the 2019 SuperBIT science telescope commissioning launch took place with the *Centre National d'Etudes Spatiales* (CNES) through the Canadian Space Agency (CSA) from the launch at the Timmins (Ontario) site. The performance of this test flight shows that SuperBIT has a pointing stability at 48 milliarcseconds over multiple 1 hour observations at float (Romualdez et al., 2019). During this test flight, we observed the galaxy cluster Abell 2218 ($z = 0.175$) in Lum band with a total (stacked) of 1950s exposure time, in U, B and G, and IR bands with 300s for each. These are shorter than exposures anticipated for the science flight, but useful for SuperBIT's PSF analysis, and the calibration of its weak lensing analysis pipeline.

A long duration science flight is scheduled in 2021. At that time, 200 galaxy clusters are expected to be observed. To prepare for automated analyses of these wide field images, a weak lensing shear measurement pipeline for SuperBIT has to be developed. In the following section, I present the process of designing a suitable weak lensing shape measurement code and the preliminary shear profile of Abell 2218.

6.3.1 Weak Lensing Shear From SuperBIT

We adapt the *HST* shear measurement code pyRRG to SuperBIT. Images of Abell 2218 are first passed through pre-processing, including bias, dark, flat correction and co-adding images. Then we run SExtractor to detect sources. As we mentioned in Section 3.3.1, pyRRG can classify objects automatically using Random Forest approach. However, since the training sets for star-galaxy classification in pyRRG are based on *HST* images, which is not suitable for SuperBIT, we therefore manually conduct the classification by examining the distribution of magnitude and peak surface brightness. We then measure the moments of stars and galaxies to calibrate their shapes. The PSF modelling here is slightly different from the process for *HST* data. First of all, the focus of SuperBIT is stable; it does not change with time. Secondly, *HST*'s PSF has been well-studied and it can be described by simulations from the tinytim model (Rhodes et al., 2007). To model SuperBIT's PSF, we make use of the shapes of bright isolated stars (see Equation 3.3.24). We perform a 2 dimensional 2-order polynomial function fitting with each moment of the stars. The original process in pyRRG takes the focus as one of the fitting inputs omitted here. The best-fit models are interpolated to any positions in the field. Figure 6.4 shows the pattern of stellar ellipticities, and the best-fit PSF model. The PSF of SuperBIT varies dramatically over the field. Close to the the boundary, the ellipticity of the PSF is getting larger and stretched in the same direction to the North/South edge. The best-fit model is able to describe the shape of the majority of stars. There are a few objects with irregular and large ellipticities, which could be due to the noise and cannot be described by the best-fit model.

Based on this shear measurement, we obtain a preliminary shear catalogue of Abell 2218 which is contaminated by unlensed objects, such as cluster members. We therefore make use of their colors to identify the red sequence galaxies, and exclude them from weak lensing analysis. There is a total of five band (Lum, U, B, G and IR) photometries available from the SuperBIT flight. Their instrumental magnitude are output directly from SExtractor, without zero point calibration. In the long-duration flight, they will be calibrated from observations of (specto)photometric standard stars. For an approximate solution for data from this test flight, we conduct

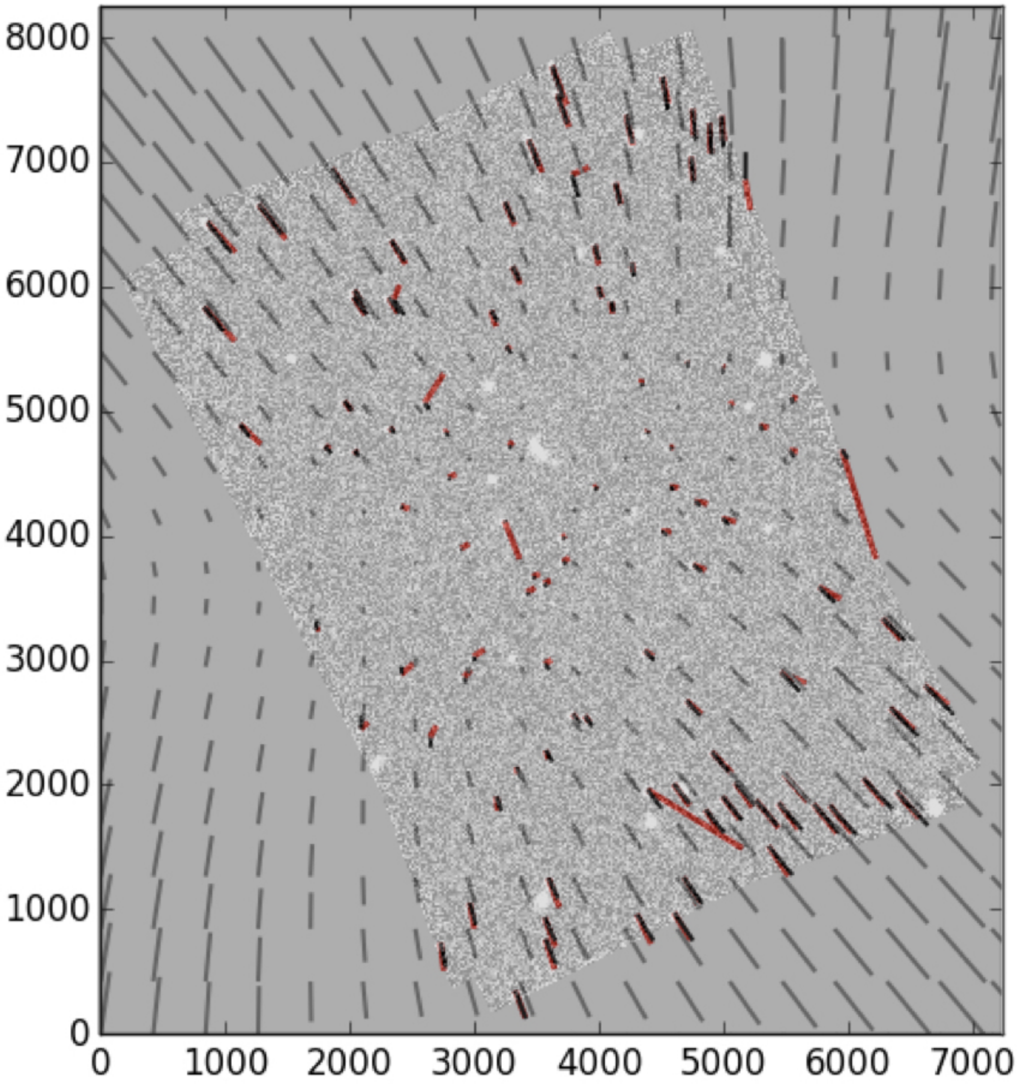


Figure 6.4: Best-fit PSF's ellipticity for SuperBIT. The stacked image of Abell 2218 is overlaid. Red lines represent the stellar ellipticity, and the black lines are the predictions from the best-fit PSF model which is interpolated to the whole field of view (also extended outside the field of view).

object matching with Abell 2218 catalogue from VizieR which have magnitudes of B, G, R, I and Z bands with zero point correction. There are a total of 213 galaxies common to the SuperBIT and VizierR catalogues. We then derive the mean value of zero point magnitude from these overlaped objects. Note that these values may not be accurate because of differences in filter shapes. A serious filter matching need a filter correction to account for the flux mismatch. However, since in this work we hope to identify red sequence galaxies in the color-magnitude diagram, a roughly correct magnitude is enough for our first pass estimate.

Figure 6.5 shows the colour-magnitude diagram of SuperBIT detected galaxies in Abell 2218. To properly select the red sequence, we follow Medezinski et al. (2007), defined a boundary in V-I vs. I diagram for Abell 1689 cluster galaxies (z=0.18) to separate the E/SO sequence from the background objects. Since Abell 1689 has a similar redshift as Abell 2218, we therefore adopt their linear relation, $(G-IR) > 0.03525\,IR + 1.505 - 0.45$ and $(G-IR) < 0.03525\,IR + 1.505 + 0.2$, to identify cluster members. Cluster galaxies are labeled in green and background samples are labeled in red in figure 6.5. Finally we further perform a lensing cut, which restricts $e < 1$ and $S/N > 3$, to exclude objects with less accurate shape measurements, and a magnitude cut, to exclude very bright objects. The final shear catalogue contains 146 background galaxies which are labeled in yellow in figure 6.5. This low number density of background galaxies is due to a short exposure during the engineering flight. The tangential weak lensing shear profile (defined in equation 3.2.16) calculated from the shear signal of the identified background galaxies is shown in figure 6.6, compared with the shear profile calibrated from ground-based CFHT observations (Herbonnet et al., 2019, shown with black data points provided via private communication by H. Ricardo). Even though we do not detect a high density of background objects in this test flight, the shear profile is well consistent with the one obtained with CFHT, showing that SuperBIT is a promising telescope for future lensing studies. We also measure the shear profile from cluster members and find that they are consistent with zero, agree with our expectation that cluster members do not contain tangential shear signal.

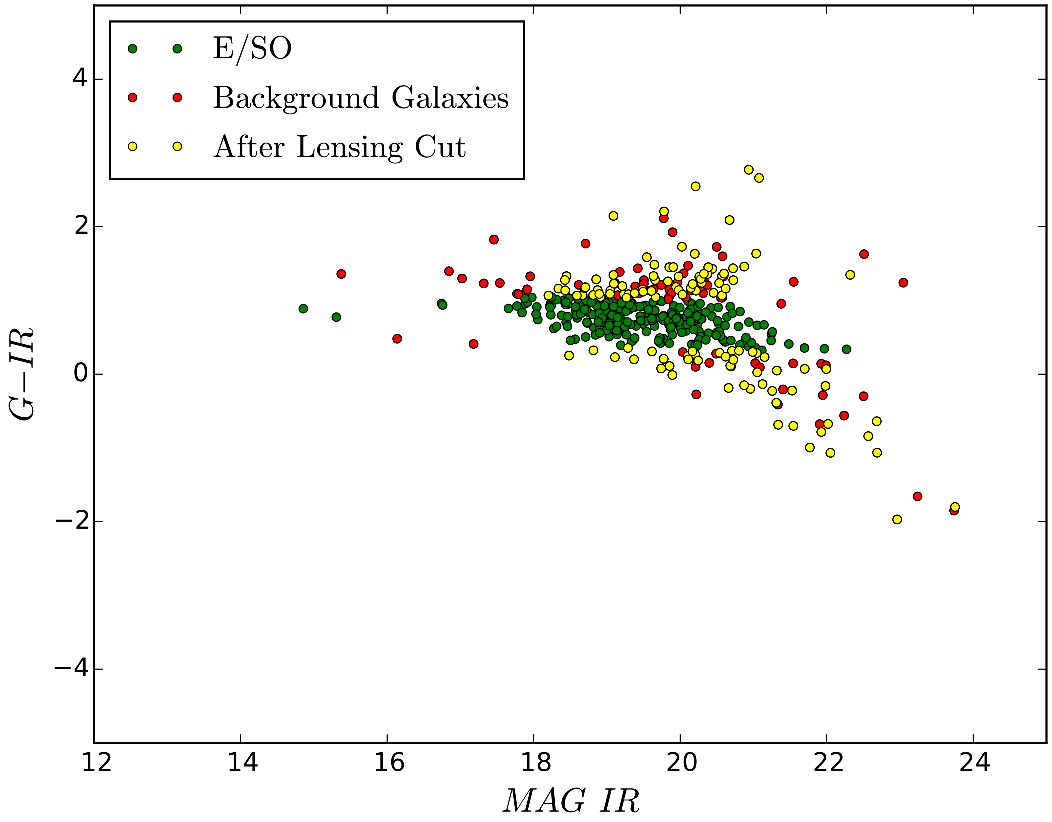


Figure 6.5: Colour vs. magnitude diagram for Abell 2218 cluster galaxies. The green and red points represent the cluster members and the background galaxies, respectively. The yellow points represent the background galaxies after applying the lensing cut, which are finally used for our weak lensing analysis.

		σ_κ	MSE
Training Set	$\sigma_\gamma = 0.26$	0.0024 ± 0.0006	$(5.90 \pm 3.20) \times 10^{-6}$
	$\sigma_\gamma = 0.36$	0.0026 ± 0.0006	$(7.68 \pm 3.84) \times 10^{-6}$
Testing Set	$\sigma_\gamma = 0.26$	0.0026 ± 0.0007	$(6.81 \pm 4.15) \times 10^{-6}$
	$\sigma_\gamma = 0.36$	0.0032 ± 0.0010	$(11.5 \pm 8.33) \times 10^{-6}$

Table 6.1: Statistics of the CNNs' performances. Third (Fourth) column shows the mean value and the standard deviation of the best-fit Gaussian function applied on the σ_κ (MSE) distribution for different samples.

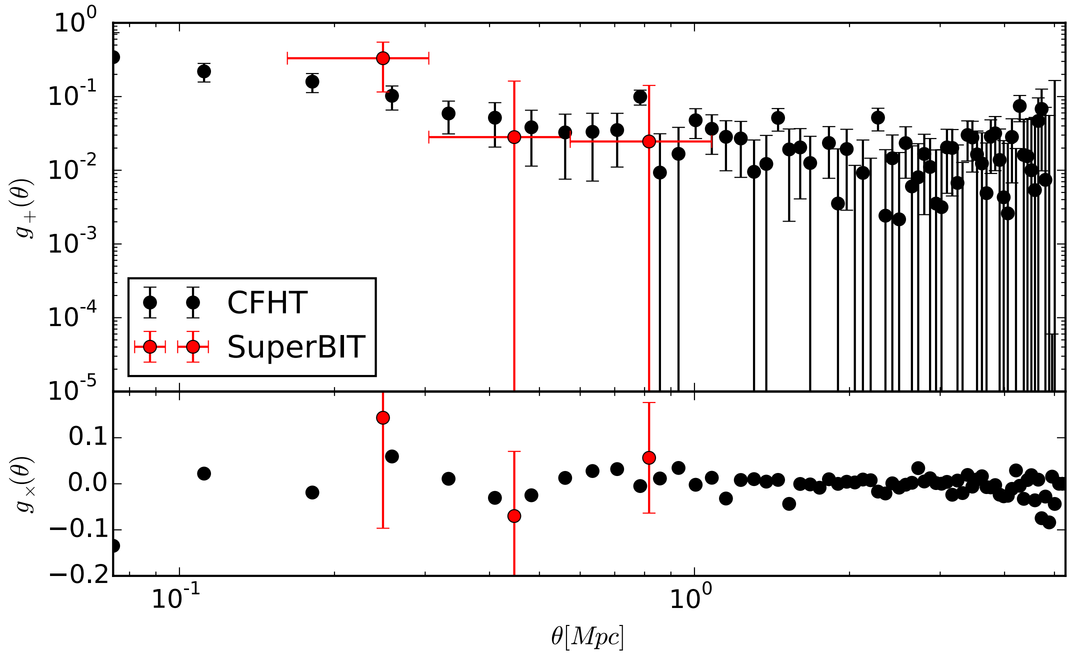


Figure 6.6: Shear profiles of Abell 2218 observed from SuperBIT (this work) and CFHT (Herbonnet et al., 2019, private communication by H. Ricardo). The red errorbars are calculated from bootstrap resampling. The y axis in the upper panel represents the magnitude of the tangential component of the shear, and the bottom panel shows the 45 degree rotated component of the shear, which should be consistent with zero in cluster lensing.

6.3.2 Future Plans

During the 2019 SuperBIT's test flight, the observations of Abell 2218 were unfortunately obtained while the telescope was away from optimal focus. This enlarged the PSF. It is therefore not trivial to measure shear using pyRRG, which is designed for diffraction-limited images. Especially when the size of the PSF is larger or equivalent to the size of the object, the assumption of pyRRG breaks and may lead to an unreliable correction. In the future, we will try to use other software, such as NGMIX (Sheldon, 2015), to improve the performances of the shear estimation. Moreover, in order to accurately calibrate the weak lensing shear from observations, a set of simulated images convolved with a similar PSF model as SuperBIT's and known weak lensing shear, is needed. These simulated images with the known shear signal can be used to quantify the precision of estimated shear from different shear measurement codes. Similar calibrations were successfully conducted by the STEP (Shear TEsting Programme; Heymans et al., 2006; Massey et al., 2007b).

SuperBIT is scheduled for a long duration super-pressure balloon flight from Wanaka, New Zealand in 2021. We will extend the exposure times in order to detect enough background galaxies. The matured weak lensing analysis pipeline will be used to calibrate the weak lensing shear signal for these observations. Together with the weak lensing mass reconstruction methods described in Section 3.4, detailed mass distribution of more than 150 galaxy clusters will be mapped. These wide-field dark matter distributions will therefore be useful in detecting large-scale structures, probing the evolutionary history of structures, and constraining the nature of dark matter. Furthermore, this large sample of galaxy clusters will allow us to probe the abundance of galaxy clusters as a function of redshift and mass, which provides a potentially sensitive test of cosmological models (e.g. Holder et al., 2001).

Chapter 7

Conclusions

In this thesis, we have focused on the gravitational lensing analyses of massive galaxy clusters in order to map the dark matter distribution and its surrounding large-scale structures. To find the optimal method that is suitable for space-based weak lensing data, we first tested different mass mapping techniques using simulation data and developed an aperture moment filter to search for filaments. We then applied these optimal measurements on the massive galaxy cluster MS 0451-03 to conduct a comprehensive cluster study. In this concluding chapter we summarise our findings and suggest future prospects for these areas of research.

7.1 Calibrating Weak Lensing Methods on Simulated Data

High-precision calibration of weak-lensing mass mapping techniques will be needed for upcoming space-based surveys. We have used mock observations of ten massive galaxy clusters from the BAHAMAS simulation to find the optimal methods for various cluster analyses. To mimic real observational imaging data, we included two sources of noise: intrinsic shapes of background galaxies, and large-scale structures along the line of sight.

7.1.1 Comparison of Two Mass Reconstruction Methods

We first quantified the performance of two mass mapping techniques: 1) direct inversion method (KS93), denoised by MRLens; 2) the Lenstool multi-scale grid technique. We found that MRLens is efficient at suppressing shape noise, whilst retaining signal from statistically significant structures on all scales. The recovered density profile from KS93+MRLens is unbiased, except the smooth central profile which causes difficulty in identifing the cuspy core. Its noise suppression via smoothing also makes the shape of the galaxy cluster rounder. The KS93+MRLens method will be suitable in stacked analysis for constraining the mean properties of an ensemble of galaxy clusters.

Lenstool suppresses noise much further thanks to its physical prior on the mass distribution. This method is thus ideal for precise reconstructions of individual clusters or detections of irregular, low signal-to-noise quantities, such as filaments. The shortcoming of this technique is the over-estimation of mass. Since Lenstool imposes a strong prior that the mass density is positive over the field of view, the recovered density profile has an excess of mass at large radii. This changes the shape of the density profile, which could cause bias in the inferred properties of the cluster. This issue will need to be managed carefully when people perform profile fitting with halo models.

7.1.2 Searching for Filaments

We also developed an optimal filter to search for filaments in the reconstructed maps. In principle, multipole aperture moments can be used to detect filaments. In practice, the signals of filaments are mimicked by background galaxy shape noise and projected LSS noise. Our optimal method allows us to detect narrow filaments from space-based data, with a purity $> 75\%$ and completeness $> 40\%$. However, finding filaments remains challenging, because they have low signal-to-noise ratio and the signal from many filaments can rarely be stacked. The dominant source of noise relevant to filament detection comes from intrinsic shapes of background galaxy. It is therefore impossible to detect individual filaments using data from ground-based telescopes, and remains challenging with current space-based (*HST*) data. Fortunately, upcoming space-based observations will resolve a higher density of background galaxies which can reduce the noise level and make this method potentially useful for filament searches.

7.2 Lensing Analysis of the Post-Merger Cluster MS 0451-03

We then applied these measurements on a massive galaxy cluster, MS 0451-03 to conduct a combined strong+weak lensing analysis. MS 0451 is an unique cluster that has the largest mosaic of *HST* imaging today, covering an area of ~20×20 arcmin^2. Thanks to these wide-field high-resolution imaging data, we were able to map the detailed dark matter distribution up to $R \sim 3\,\mathrm{Mpc}$. Using the Lenstool multi-scale grid reconstruction, we discovered six substructures, constituting three possible filaments with mean convergence $\langle\kappa\rangle \sim 0.03$.

We found that the cluster core is a bimodal mass distribution elongated along the South-East to the North-West with total mass $M_{200} = (1.65 \pm 0.24) \times 10^{15}\,\mathrm{M}_\odot$ and NFW concentration $c_{200} = 3.79 \pm 0.36$. The MS 0451's baryonic fraction $f_{\mathrm{b},500} = (14.6 \pm 1.4)\%$ is well consistent with the cosmic baryonic fraction (Planck Collaboration et al., 2016). We detected the flattening features of MS 0451's density profile at $R \approx 2\,\mathrm{Mpc}$, which is well fitted by the splashback feature in the DK14 model. However, since this feature is very close to the boundary of the field of view, this could be just noise corresponding to edge effects or large-scale structures projected from other redshifts.

We compared our lensing results with an X-ray analysis which measured a total mass of $M_{200,\mathrm{HSE}} = (1.75 \pm 0.75) \times 10^{15}\,\mathrm{M}_\odot$ and $c_{\mathrm{HSE}} = 2.35^{+0.89}_{-0.70}$, assuming that the system is in hydrostatic equilibrium (HSE). The distribution of gas is rounder than that of dark matter and there are also strong excesses of gas entropy in the cluster core. These findings suggest that MS 0451 is an unrelaxed system which underwent a major merger ~2–7 Gyr ago, and the two dark matter components in the centre are now approaching second apocentre. The post-merger stage of MS

0451 we inferred is consistent with the quenched star-formation history discovered by Moran et al. (2007b). Therefore, MS 0451 will be a useful cluster for future studies to characterise infalling substructures along filaments and the star-formation processes after a major merger. The analyses we presented here can be extended and applied to other systems observed in future surveys.

7.3 Future prospects

Finally we presented some preliminary results of two ongoing projects. We reconstructed the mass maps of NFW halos using a U-Net Convolutional Neural Network. The pre-trained network effectively removes the random noise peaks caused by background galaxies' shape noise, and recover well the mass distribution and density profile. In the future, a more realistic training dataset with complex cluster mass distributions from simulations is needed for this approach to be applied on real observational data, including irregular and unrelaxed clusters that cannot be described by simple analytic halo models.

We also showed the preliminary shear profile of Abell 2218 observed from SuperBIT during the test flight in 2019, which is consistent with what is observed from CFHT. This suggests that SuperBIT is a promising telescope which can deliver high-resolution imaging at a fraction of the cost of a space-based mission.

Appendix A

Auxiliary Results

A.1 Splashback radius

Diemer & Kravtsov (2014) shows that in N-body simulations density profiles of dark matter halos exhibit a sharp steepening at radii comparable to the virial radius. This feature depends on the accretion history of the cluster, resulting from an absence of particles orbiting beyond the radius of second turnaround. It gives us a physically motivated definition for the boundary of dark matter halos. Here we investigate the splashback feature of MS 0451 by fitting its density profile with a DK14 profile (Sect. 5.6.2). The marginalized posterior constraints are listed in Table A.1 where we employ the biweight estimators of Beers et al. (1990) for the center and dispersion of the marginalized posterior distributions (e.g. Chiu et al., 2018; Sereno & Umetsu, 2011; Umetsu et al., 2014a).

We follow More et al. (2015b) to define the splashback radius, $r_{\rm sp}$, as the radius of a local minimum in the logarithmic slope of the density profile, $\gamma \equiv {\rm d}\log\rho/{\rm d}\log{\rm r}$. Figure A.1 shows the mean and 68 per cent confidence intervals of γ, inferred from the DK14 fit, together with the posterior probability distribution of the splashback radius, $r_{\rm sp}$, and the posterior probability distribution of $\gamma(r_{\rm sp})$. The biweight central location of $\gamma(r_{\rm sp})$ is -3.10 ± 0.74, at $r_{\rm sp} = 1.49 \pm 0.57\,{\rm Mpc}$. This is not significantly different from the value of $\gamma \sim -2.4$ for the best-fit NFW at this radius, and the Bayesian and Akaike Information Criteria disfavour the increased complexity of the DK14 model. Furthermore, our best-fit $r_{\rm sp}$ is lower than predictions from

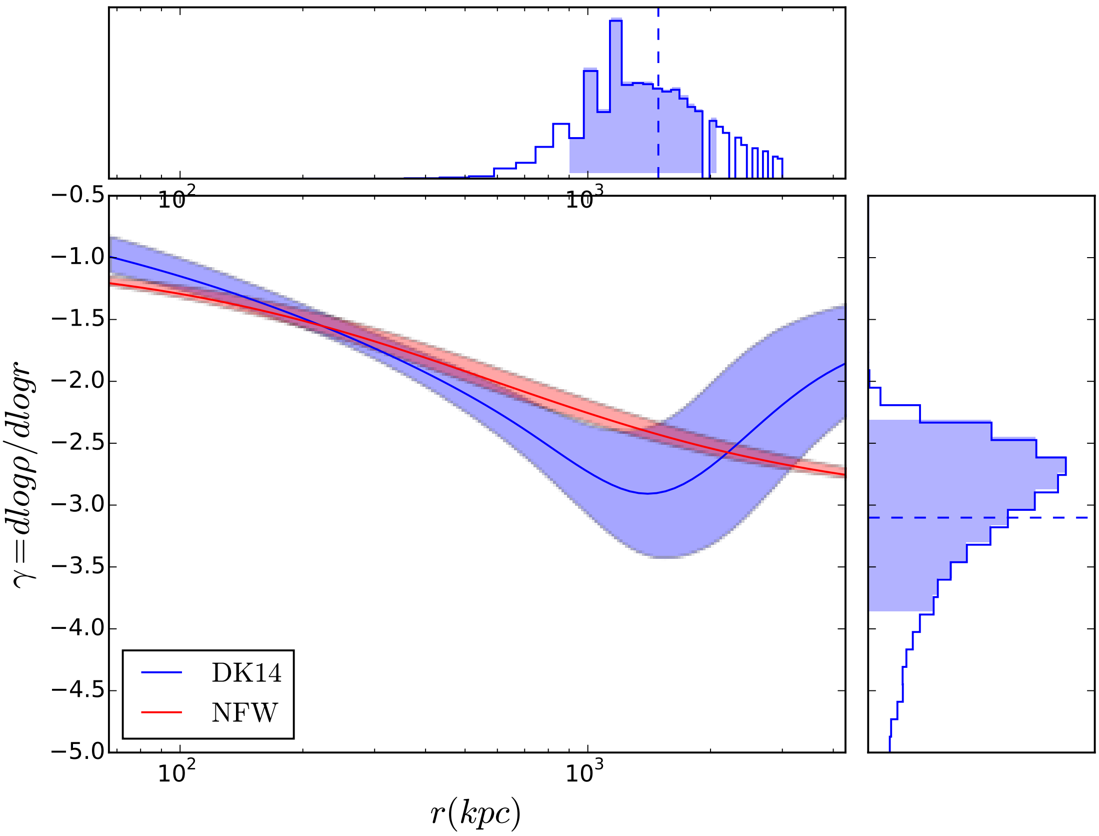


Figure A.1: Radial gradient of the total mass distribution, from fitted NFW (red) and DK14 (blue) models. Solid lines show the mean inferred values; shaded regions show 68% confidence intervals The upper and right panels show the posterior probability distributions of the splashback radius $r_{\rm sp}$, and the gradient at the splashback radius $\gamma(r_{\rm sp})$. Dashed lines and shaded regions indicate the mean and 68% confidence intervals respectively.

Parameter	**Constraint**
$\rho_s\,[10^3\mathrm{M}_\odot/\mathrm{kpc}^3]$	598.32 ± 6.41
$r_s\,[\mathrm{Mpc}]$	0.33 ± 0.12
$r_t\,[\mathrm{Mpc}]$	1.76 ± 1.34
$\log\alpha$	-0.78 ± 0.39
$\log\beta$	0.87 ± 0.45
$\log\gamma$	0.70 ± 0.45
b_e	1.96 ± 1.00
s_e	1.75 ± 0.54

Table A.1: Marginalized posterior constraints on the DK14 model.

cosmological simulations and other observational analyses (e.g. Contigiani et al., 2018). One explanation could be that a true splashback feature is close to (or ouside) the edge of the *HST* field of view, where 'noise' in the form of lensing signal from projected substructures exceeds the lensing signal of the cluster, and is correlated between radial bins. Hence, similarly to Umetsu & Diemer (2017) on a different cluster sample, our measurement of MS 0451 places only a lower limit on the splashback radius $r_{\mathrm{sp}} > 1.49\,\mathrm{Mpc}$.

A.2 Redshift distribution of detected weak lensing peaks in MS 0451's field

The summed probability density functions (PDFs) of photometric redshifts z_{phot} for all galaxies within $R = 480\,\mathrm{kpc}$ of each substructure are shown in figure A.2. For comparison, the dashed red line shows the redshift distribution of galaxies observed in the HST imaging with the same pass-band and to the same depth, but in a blank region of the sky (the COSMOS field). Its normalisation has been rescaled to the same number density of galaxies in our catalogue that have photometric redshifts (a higher fraction of COSMOS galaxies have photometric redshifts, particularly at high redshift, thanks to the deeper Subaru imaging). The grey band shows the

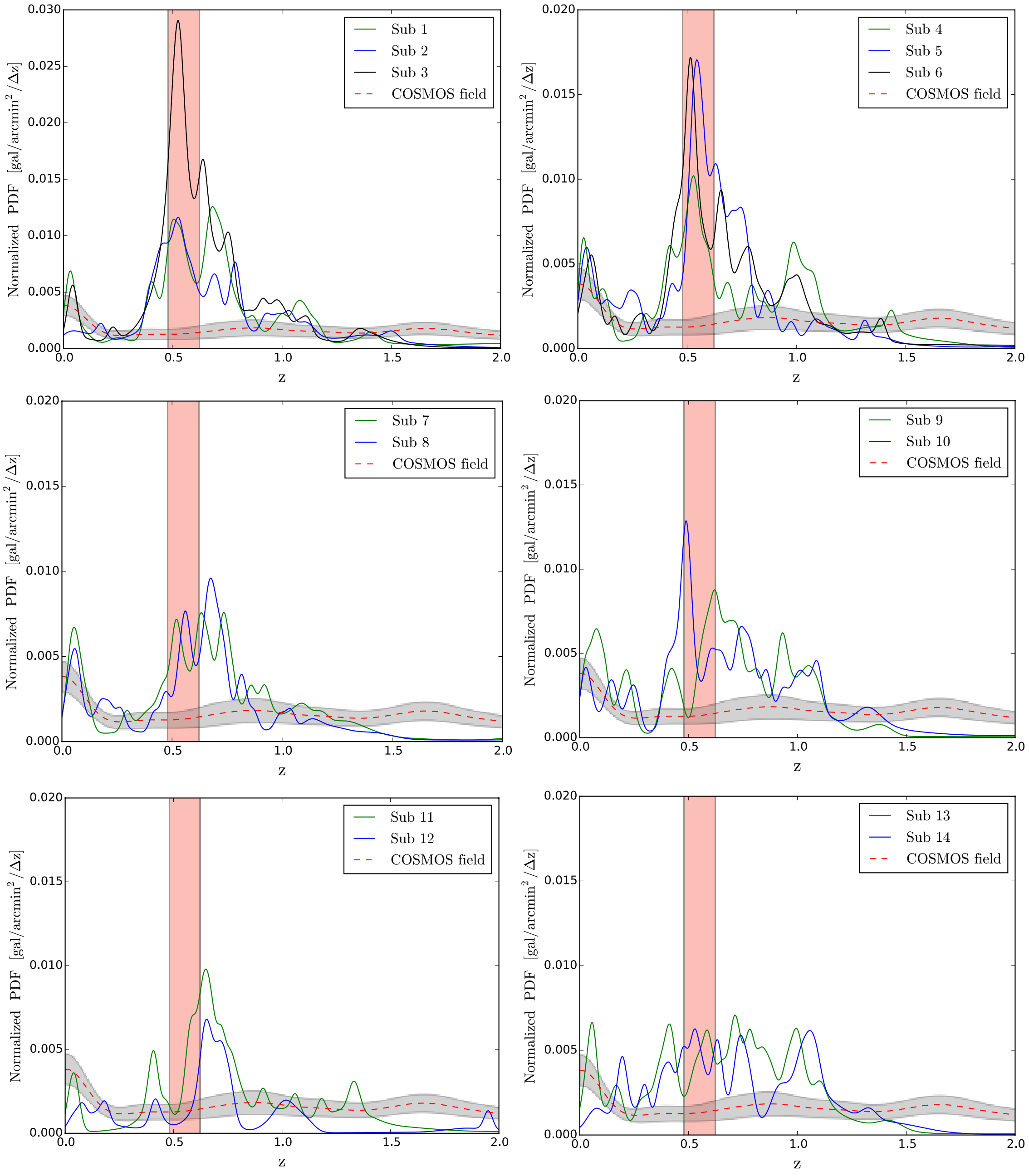


Figure A.2: Normalized PDF of photometric redshifts z_{phot} for all galaxies within an aperture ($R = 480\,\text{kpc}$) for each of the 14 detected weak-lensing peaks. Galaxy overdensities within $0.48 < z_{\text{phot}} < 0.61$ (vertical red bands) are consistent with being at the redshift of MS 0451, within typical z_{phot} uncertainties. For comparison, the red dashed line is the redshift distribution of galaxies detected in comparable *H*ST imaging of a blank patch of sky (the COSMOS field), and the grey band shows the 1σ scatter in this.

1σ scatter in the redshift distribution of COSMOS galaxies, calculated using the bootstrap method. The width of the grey band is unchanged by the higher precision of $z_{\rm phot}$ measurements using many more colours.

Substructures 1, 2, 3, 4, 5 and 6 are dominated by galaxies at the same redshift as the main cluster, and must therefore be physically associated. Indeed, there are some galaxies at this redshift throughout the entire *H*ST mosaic. However, substructures 7 to 14 (marked by black circles in figure 5.10) are dominated by galaxies at a different redshift or at a mixture of redshifts. We therefore do not consider these to be associated to MS 0451. Substructures 8, 9 and 11 might be a linked system behind the cluster ($z \sim 0.7$), and even appear as an extended mass distribution in the AMM map (figure 5.11).

Bibliography

Adhikari S., Chue C. Y. R., Dalal N., 2015, JCAP, 1501, 009

Alpaslan M., et al., 2015, MNRAS, 451, 3249

Alpaslan M., et al., 2016, MNRAS, 457, 2287

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

Aragón-Calvo M. A., van de Weygaert R., Jones B. J. T., van der Hulst J. M., 2007, ApJl, 655, L5

Arnaud K., Dorman B., Gordon C., 1999, XSPEC: An X-ray spectral fitting package (ascl:9910.005)

Arnouts S., Cristiani S., Moscardini L., Matarrese S., Lucchin F., Fontana A., Giallongo E., 1999, MNRAS, 310, 540

Arnouts S., et al., 2007, A&A, 476, 137

Bacon R., et al., 2010, in Ground-based and Airborne Instrumentation for Astronomy III. p. 773508, doi:10.1117/12.856027

Bahcall Bode P., 2003, ApJl, 588, L1

Bahcall N. A., Cen R., 1993, ApJl, 407, L49

Bard D., et al., 2013, ApJ, 774, 49

Baron D., 2019, arXiv e-prints, p. arXiv:1904.07248

Bartelmann M., 1996, AA, 313, 697

Bartelmann Maturi M., 2017, Scholarpedia, 12, 32440

Bartelmann M., Schneider P., 2001, , 340, 291

Baxter E., et al., 2017, Astrophys. J., 841, 18

Beers T. C., Flynn K., Gebhardt K., 1990, The Astrophysical Journal, 100, 32

Berciano Alba A., Koopmans L. V. E., Garrett M. A., Wucknitz O., Limousin M., 2010, A&A, 509, A54

Bertin E., Arnouts S., 1996, A&As, 117, 393

Bond J. R., Kofman L., Pogosyan D., 1996, Nature, 380, 603

Borys C., et al., 2004, MNRAS, 352, 759

Bradac M., et al., 2006, Astrophys. J., 652, 937

Bradač M., et al., 2006, ApJ, 652, 937

Brehmer J., Mishra-Sharma S., Hermans J., Louppe G., Cranmer K., 2019, ApJ, 886, 49

Burkert A., 1995, ApJl, 447, L25

Burnham K., Anderson D., 2002, New York: Springer

Byrd R., Lu P., Nocedal J., Zhu C., 1995, SIAM Journal of Scientific Computing, 16, 1190

Cautun M., van de Weygaert R., Jones B. J. T., 2013, MNRAS, 429, 1286

Chabrier G., 2003, , 115, 763

Chang C., et al., 2018, Astrophys. J., 864, 83

Chiu I.-N., Umetsu K., Sereno M., Ettori S., Meneghetti M., Merten J., Sayers J., Zitrin A., 2018, Astrophys. J., 860, 126

Clampitt J., Jain B., 2016, Mon. Not. Roy. Astron. Soc., 457, 4135

Clowe D., Gonzalez A., Markevitch M., 2004, ApJ, 604, 596

Clowe D., et al., 2006, A&A, 451, 395

Colless M., et al., 2001, MNRAS, 328, 1039

Contigiani O., Hoekstra H., Bahé Y. M., 2018,] 10.1093/mnras/stz404

Crain R. A., et al., 2009, MNRAS, 399, 1773

De Filippis E., Sereno M., Bautz M. W., Longo G., 2005, ApJ, 625, 108

Despali G., Giocoli C., Bonamigo M., Limousin M., Tormen G., 2017, Mon. Not. Roy. Astron. Soc., 466, 181

Diemer B., 2018, Astrophys. J. Suppl., 239, 35

Diemer B., Kravtsov A. V., 2014, Astrophys. J., 789, 1

Dietrich J. P., Schneider P., Clowe D., Romano-Diaz E., Kerp J., 2005, Astron. Astrophys., 440, 453

Dietrich J. P., Werner N., Clowe D., Finoguenov A., Kitching T., Miller L., Simionescu A., 2012, Nature, 487, 202

Donahue M., Gaskin J. A., Patel S. K., Joy M., Clowe D., Hughes J. P., 2003, ApJ, 598, 190

Donovan D. A. K., 2007, PhD thesis, University of Hawai'i at Manoa

Dutton A. A., Macciò A. V., 2014, MNRAS, 441, 3359

Dyson F. W., Eddington A. S., Davidson C., 1920, Philosophical Transactions of the Royal Society of London Series A, 220, 291

Ebeling H., Edge A. C., Henry J. P., 2001, ApJ, 553, 668

Ebeling H., Barrett E., Donovan D., Ma C.-J., Edge A. C., van Speybroeck L., 2007, ApJl, 661, L33

Eckert D., Roncarelli M., Ettori S., Molendi S., Vazza F., Gastaldello F., Rossetti M., 2015, MNRAS, 447, 2198

Eckert D., et al., 2016, A&A, 592, A12

Eckert D., Ettori S., Pointecouteau E., Molendi S., Paltani S., Tchernin C., 2017, Astronomische Nachrichten, 338, 293

Eckert D., et al., 2019, A&A, 621, A40

Einasto J., 1965, Trudy Astrofizicheskogo Instituta Alma-Ata, 5, 87

Einasto M., et al., 2007, A&A, 464, 815

Einstein A., 1915, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin, pp 844–847

Elíasdóttir Á., et al., 2007, preprint, (`arXiv:0710.5636`)

Ettori S., et al., 2019, A&A, 621, A39

Evans A. K. D., Bridle S., 2009, Astrophys. J., 695, 1446

Faber S. M., Jackson R. E., 1976, ApJ, 204, 668

Fluri J., Kacprzak T., Lucchi A., Refregier A., Amara A., Hofmann T., Schneider A., 2019

Ford H. C., et al., 1996, in Bely P. Y., Breckinridge J. B., eds, Vol. 2807, Space Telescopes and Instruments IV. pp 184–196, doi:10.1117/12.255099

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, , 125, 306

Fujita Y., Takizawa M., Nagashima M., Enoki M., 1999, PASJ, 51, L1

Galárraga-Espinosa D., Aghanim N., Langer M., Gouin C., Malavasi N., 2020, arXiv e-prints, p. arXiv:2003.09697

Gavazzi R., Mellier Y., Fort B., Cuillandre J. C., Dantel-Fort M., 2004, A&A, 422, 407

Geller M. J., Huchra J. P., 1989, Science, 246, 897

Ghirardini V., et al., 2019, A&A, 621, A41

Gilmore J., Natarajan P., 2009, MNRAS, 396, 354

Gioia I. M., Maccacaro T., Schild R. E., Wolter A., Stocke J. T., Morris S. L., Henry J. P., 1990, ApJs, 72, 567

Golse G., Kneib J.-P., 2002, A&A, 390, 821

Graham A. W., Merritt D., Moore B., Diemand J., Terzic B., 2006, Astron. J., 132, 2685

Gray M. E., Taylor A. N., Meisenheimer K., Dye S., Wolf C., Thommes E., 2002, ApJ, 568, 141

Gunn J. E., Gott J. Richard I., 1972, ApJ, 176, 1

Guzzo L., et al., 2007, ApJs, 172, 254

Hamana T., Oguri M., Shirasaki M., Sato M., 2012, MNRAS, 425, 2287

Harvey D., Massey R., Kitching T., Taylor A., Tittley E., 2015, Science, 347, 1462

Harvey D., Tam S.-I., Jauzac M., Massey R., Rhodes J., 2019, arXiv e-prints, p. arXiv:1911.06333

Herbonnet R., et al., 2019, arXiv e-prints, p. arXiv:1912.04414

Heymans C., et al., 2005, Mon. Not. Roy. Astron. Soc., 361, 160

Heymans C., et al., 2006, MNRAS, 368, 1323

Heymans C., et al., 2008, MNRAS, 385, 1431

Hinshaw G., et al., 2013, ApJs, 208, 19

Ho S., Bahcall N., Bode P., 2006, Astrophys. J., 647, 8

Hoekstra H., 2013, arXiv e-prints, p. arXiv:1312.5981

Hoffman S., Avila R., 2018, STScI ISR, 2017-02

Holder G., Haiman Z., Mohr J. J., 2001, ApJl, 560, L111

Hopkins P. F., Bahcall N. A., Bode P., 2005, The Astrophysical Journal, 618, 1

Ilbert O., et al., 2009, ApJ, 690, 1236

Jauzac M., et al., 2012, MNRAS, 426, 3369

Jauzac M., et al., 2015a, MNRAS, 446, 4132

Jauzac M., et al., 2015b, Mon. Not. Roy. Astron. Soc., 452, 1437

Jauzac et al., 2016, Mon. Not. Roy. Astron. Soc., 463, 3876

Jee M. J., Hughes J. P., Menanteau F., Sifón C., Mandelbaum R., Barrientos L. F., Infante L., Ng K. Y., 2014, ApJ, 785, 20

Jeffrey N., Lanusse F., Lahav O., Starck J.-L., 2020, MNRAS, 492, 5023

Jeltema T. E., Canizares C. R., Bautz M. W., Buote D. A., 2005, ApJ, 624, 606

Jing Y. P., Suto 2002, Astrophys. J., 574, 538

Jørgensen I., Chiboucas K., 2013, The Astrophysical Journal, 145, 77

Jørgensen I., Chiboucas K., Hibon P., Nielsen L. D., Takamiya M., 2018, ApJs, 235, 29

Jullo E., Kneib J.-P., 2009, MNRAS, 395, 1319

Jullo E., Kneib J.-P., Limousin M., Elíasdóttir Á., Marshall P. J., Verdugo T., 2007, New Journal of Physics, 9, 447

Jullo E., Pires S., Jauzac M., Kneib J.-P., 2014a, MNRAS, 437, 3969

Jullo E., Pires S., Jauzac M., Kneib J.-P., 2014b, Mon. Not. Roy. Astron. Soc., 437, 3969

Kaiser N., 1984, ApJl, 284, L9

Kaiser N., 1986, MNRAS, 222, 323

Kaiser N., 1995, ApJl, 439, L1

Kaiser N., Squires G., 1993, ApJ, 404, 441

Kaiser N., Wilson G., Luppino G., Kofman L., Gioia I., Metzger M., Dahle H., 1998

Kartaltepe J. S., Ebeling H., Ma C. J., Donovan D., 2008, MNRAS, 389, 1240

Kassiola A., Kovner I., 1993a, ApJ, 417, 450

Kassiola A., Kovner I., 1993b, ApJ, 417, 450

Keeton C. R., 2001a

Keeton C. R., 2001b, arXiv e-prints, pp astro–ph/0102341

Kilbinger M., 2015, Reports on Progress in Physics, 78, 086901

Kimble R. A., MacKenty J. W., O'Connell R. W., Townsend J. A., 2008, in Space Telescopes and Instrumentation 2008: Optical, Infrared, and Millimeter. p. 70101E, doi:10.1117/12.789581

Klypin A., Yepes G., Gottlober S., Prada F., Hess S., 2016, Mon. Not. Roy. Astron. Soc., 457, 4340

Kneib J.-P., Natarajan P., 2011a, A&Ar, 19, 47

Kneib J.-P., Natarajan P., 2011b, A&Ar, 19, 47

Kneib J.-P., Ellis R. S., Smail I., Couch W. J., Sharples R. M., 1996, ApJ, 471, 643

Knudsen K. K., Richard J., Kneib J.-P., Jauzac M., Clément B., Drouart G., Egami E., Lindroos L., 2016, MNRAS, 462, L6

Kriss G. A., Cioffi D. F., Canizares C. R., 1983, ApJ, 272, 439

Krolewski A., et al., 2018, Astrophys. J., 861, 60

Kronberger T., Kapferer W., Ferrari C., Unterguggenberger S., Schindler S., 2008, A&A, 481, 337

Kubo J. M., Stebbins A., Annis J., Dell'Antonio I. P., Lin H., Khiabanian H., Frieman J. A., 2007, ApJ, 671, 1466

Kuutma T., Tamm A., Tempel E., 2017, A&A, 600, L6

Laureijs R., et al., 2011, arXiv e-prints, p. arXiv:1110.3193

Leauthaud A., et al., 2007, ApJs, 172, 219

Leauthaud A., et al., 2012, ApJ, 744, 159

Leonard A., Pires S., Starck J.-L., 2012, MNRAS, 423, 3405

Leonard A., Lanusse F., Starck J.-L., 2015, MNRAS, 449, 1146

Limousin M., Kneib J.-P., Natarajan P., 2005, MNRAS, 356, 309

Liu C., Hao L., Wang H., Yang X., 2019, Astrophys. J., 878, 69

Ludlow A. D., et al., 2013, Mon. Not. Roy. Astron. Soc., 432, 1103

Luppino G. A., Gioia I. M., Hammer F., Le Fèvre O., Annis J. A., 1999, A&As, 136, 117

Ma C.-J., Ebeling H., Donovan D., Barrett E., 2008, ApJ, 684, 160

MacKenzie T. P., et al., 2014, MNRAS, 445, 201

Mantz A. B., Allen S. W., Morris R. G., Rapetti D. A., Applegate D. E., Kelly P. L., von der Linden A., Schmidt R. W., 2014, Mon. Not. Roy. Astron. Soc., 440, 2077

Mao T.-X., Wang J., Frenk C. S., Gao L., Li R., Wang Q., Cao X., Li M., 2018, MNRAS, 478, L34

Markevitch M., Gonzalez A. H., Clowe D., Vikhlinin A., Forman W., Jones C., Murray S., Tucker W., 2004, ApJ, 606, 819

Martinet N., et al., 2016, A&A, 590, A69

Martizzi D., Vogelsberger M., Torrey P., Pillepich A., Hansen S. H., Marinacci F., Hernquist L., 2019

Massey R., et al., 2007a, Astrophys. J. Suppl., 172, 239

Massey R., et al., 2007b, MNRAS, 376, 13

Massey R., Kitching T., Richard J., 2010, Reports on Progress in Physics, 73, 086901

Massey R., et al., 2013, MNRAS, 429, 661

Massey R., et al., 2014, MNRAS, 439, 887

Massey R., et al., 2015, MNRAS, 449, 3393

Mazzotta P., Rasia E., Moscardini L., Tormen G., 2004, MNRAS, 354, 10

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

McClintock T., et al., 2019, Mon. Not. Roy. Astron. Soc., 482, 1352

Mead J. M. G., King L. J., McCarthy I. G., 2010, MNRAS, 401, 2257

Medezinski E., et al., 2007, ApJ, 663, 717

Medezinski E., Broadhurst T., Umetsu K., Oguri M., Rephaeli Y., Benítez N., 2010, MNRAS, 405, 257

Medezinski E., et al., 2017,] 10.1093/pasj/psx128

Medezinski E., et al., 2018, PASJ, 70, 30

Meneghetti M., Jain B., Bartelmann M., Dolag K., 2005, MNRAS, 362, 1301

Merten J., Cacciato M., Meneghetti M., Mignone C., Bartelmann M., 2009, Astron. Astrophys., 500, 681

Merten J., et al., 2011, MNRAS, 417, 333

Merten J., et al., 2015, ApJ, 806, 4

Metcalf R. B., et al., 2019, A&A, 625, A119

Miles N., Lim P., Bellini A., Grogin N., 2018, STScI ISR, 2018-05

Milgrom M., 1983, ApJ, 270, 365

Miyatake H., et al., 2019, Astrophys. J., 875, 63

Molnar S. M., Broadhurst T., 2018, Astrophys. J., 862, 112

Molnar S. M., Hughes J. P., Donahue M., Joy M., 2002, ApJl, 573, L91

Moran S. M., Loh B. L., Ellis R. S., Treu T., Bundy K., MacArthur L. A., 2007a, ApJ, 665, 1067

Moran S. M., Ellis R. S., Treu T., Smith G. P., Rich R. M., Smail I., 2007b, ApJ, 671, 1503

More S., Kravtsov A. V., Dalal N., Gottlöber S., 2011, ApJs, 195, 4

More S., Diemer B., Kravtsov A. V., 2015a, ApJ, 810, 36

More S., Diemer B., Kravtsov A., 2015b, Astrophys. J., 810, 36

More S., et al., 2016, Astrophys. J., 825, 39

Natarajan P., Kneib J.-P., 1997, MNRAS, 287, 833

Natarajan P., et al., 2017, Mon. Not. Roy. Astron. Soc., 468, 1962

Navarro J. F., Frenk C. S., White S. D. M., 1996, Astrophys. J., 462, 563

Navarro J. F., Frenk C. S., White S. D. M., 1997, Astrophys. J., 490, 493

Nelson K., Rudd D., Shaw L., Nagai D., 2011, Astrophysical Journal, 751

Newman A. B., Treu T., Ellis R. S., Sand D. J., Nipoti C., Richard J., Jullo E., 2013, ApJ, 765, 24

Newman A. B., Ellis R. S., Treu T., 2015, ApJ, 814, 26

Norberg P., et al., 2002, MNRAS, 332, 827

Nuza S. E., Kitaura F.-S., Heß S., Libeskind N. I., Müller V., 2014, MNRAS, 445, 988

Ogrean G. A., et al., 2015, ApJ, 812, 153

Oguri M., Takada M., Okabe N., Smith G. P., 2010, MNRAS, 405, 2215

Okabe N., Smith G. P., 2016, MNRAS, 461, 3794

Okumura T., Jing Y. P., Li C., 2009, Astrophys. J., 694, 214

Pandey B., Bharadwaj S., 2006, Mon. Not. Roy. Astron. Soc., 372, 827

Peacock J. A., 1999, Cosmological Physics

Peebles P. J. E., 1993, Principles of Physical Cosmology

Perlmutter S., et al., 1998, Nature, 391, 51

Pires S., Starck J. L., Amara A., Teyssier R., Réfrégier A., Fadili J., 2009, MNRAS, 395, 1265

Pires S., Starck J. L., Amara A., Réfrégier A., Teyssier R., 2010, in Alimi J.-M., Fuözfa A., eds, American Institute of Physics Conference Series Vol. 1241, American Institute of Physics Conference Series. pp 1118–1127 (`arXiv:0904.2995`), doi:10.1063/1.3462608

Planck Collaboration et al., 2016, A&A, 594, A13

Planck Collaboration et al., 2018, arXiv e-prints, p. arXiv:1807.06209

Poole G. B., Fardal M. A., Babul A., McCarthy I. G., Quinn T. R., Wadsley J., 2006, Mon. Not. Roy. Astron. Soc., 373, 881

Postman M., et al., 2012, ApJs, 199, 25

Pratt G. W., Arnaud M., Biviano A., Eckert D., Ettori S., Nagai D., Okabe N., Reiprich T. H., 2019, , 215, 25

Press W. H., Schechter P., 1974, ApJ, 187, 425

Raghunathan S., Holder G. P., Bartlett J. G., Patil S., Reichardt C. L., Whitehorn N., 2019, , 2019, 037

Redmond S., et al., 2018, in Takami H., Evans C. J., Simard L., eds, Ground-based and Airborne Instrumentation for Astronomy VII. SPIE, doi:10.1117/12.2307754, `https://doi.org/10.1117%2F12.2307754`

Rehmann R. L., et al., 2019, Mon. Not. Roy. Astron. Soc., 486, 77

Rhodes J., Refregier A., Groth E. J., 2000, ApJ, 536, 79

Rhodes J. D., et al., 2007, ApJs, 172, 203

Richard J., Kneib J.-P., Ebeling H., Stark D. P., Egami E., Fiedler A. K., 2011, MNRAS, 414, L31

Riess A. G., et al., 1998, The Astrophysical Journal, 116, 1009

Riess A. G., Casertano S., Yuan W., Macri L. M., Scolnic D., 2019, ApJ, 876, 85

Ritchie B. W., Thomas P. A., 2002, Mon. Not. Roy. Astron. Soc., 329, 675

Rix H.-W., et al., 2004, ApJs, 152, 163

Robertson A., Massey R., Eke V., 2017, MNRAS, 465, 569

Robertson A., Harvey D., Massey R., Eke V., McCarthy I. G., Jauzac M., Li B., Schaye J., 2019, MNRAS, 488, 3646

Romualdez L. J., et al., 2016, arXiv e-prints, p. arXiv:1608.02502

Romualdez L. J., et al., 2019, arXiv e-prints, p. arXiv:1911.11210

Rozo E., et al., 2010, ApJ, 708, 645

Salpeter E. E., 1955, ApJ, 121, 161

Saunders W., Rowan-Robinson M., Lawrence A., 1992, MNRAS, 258, 134

Sayers J., et al., 2019, ApJ, 880, 45

Schneider P., Bartelmann M., 1997, Mon. Not. Roy. Astron. Soc., 286, 696

Schneider P., Seitz C., 1995, AA, 294, 411

Schneider P., King L., Erben T., 2000, AA, 353, 41

Schrabback T., et al., 2018, Mon. Not. Roy. Astron. Soc., 474, 2635

Schwinn J., Jauzac M., Baugh C. M., Bartelmann M., Eckert D., Harvey D., Natarajan P., Massey R., 2017, MNRAS, 467, 2913

Sereno M., Umetsu K., 2011, MNRAS, 416, 3187

Sereno M., Covone G., Izzo L., Ettori S., Coupon J., Lieu M., 2017, MNRAS, 472, 1946

Shaw L., Weller J., Ostriker J. P., Bode P., 2006, Astrophys. J., 646, 815

Sheldon E., 2015, NGMIX: Gaussian mixture models for 2D images (ascl:1508.008)

Shin T.-h., Clampitt J., Jain B., Bernstein G., Neil A., Rozo E., Rykoff E., 2018, Mon. Not. Roy. Astron. Soc., 475, 2421

Shirasaki M., Yoshida N., Ikeda S., 2019, Phys. Rev. D, 100, 043527

Simionescu A., Werner N., Mantz A., Allen S. W., Urban O., 2017, MNRAS, 469, 1476

Smith R. K., Brickhouse N. S., Liedahl D. A., Raymond J. C., 2001, ApJl, 556, L91

Smith G. P., Kneib J.-P., Smail I., Mazzotta P., Ebeling H., Czoske O., 2005, MNRAS, 359, 417

Smith G. P., et al., 2010, MNRAS, 409, 169

Smoot G. F., et al., 1992, ApJl, 396, L1

Soto K. T., Lilly S. J., Bacon R., Richard J., Conseil S., 2016, MNRAS, 458, 3210

Soucail G., Foëx G., Pointecouteau E., Arnaud M., Limousin M., 2015, A&A, 581, A31

Spergel D., et al., 2013, arXiv e-prints, p. arXiv:1305.5422

Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, MNRAS, 328, 726

Springel V., et al., 2005, Nature, 435, 629

Starck J.-L., Pires S., Réfrégier A., 2006, A&A, 451, 1139

Steinhardt C. L., et al., 2020, ApJs, 247, 64

Suto D., Kitayama T., Nishimichi T., Sasaki S., Suto Y., 2016, Publ. Astron. Soc. Jap., 68, 97

Suto D., Peirani S., Dubois Y., Kitayama T., Nishimichi T., Sasaki S., Suto Y., 2017, Publ. Astron. Soc. Jap., 69, 14

Takata T., et al., 2003, PASJ, 55, 789

Tewes M., Kuntzer T., Nakajima R., Courbin F., Hildebrandt H., Schrabback T., 2019, A&A, 621, A36

Treu T., Ellis R. S., 2015, Contemporary Physics, 56, 17

Umetsu K., Diemer B., 2017, Astrophys. J., 836, 231

Umetsu K., et al., 2014a, ApJ, 795, 163

Umetsu K., et al., 2014b, Astrophys. J., 795, 163

Umetsu K., et al., 2015, Astrophys. J., 806, 207

Umetsu K., et al., 2018, Astrophys. J., 860, 104

Umetsu K., et al., 2019, arXiv e-prints, p. arXiv:1909.10524

Van Waerbeke L., et al., 2013, MNRAS, 433, 3373

Vijayaraghavan R., Ricker P. M., 2013, MNRAS, 435, 2713

Voit G. M., 2005, Reviews of Modern Physics, 77, 207

Walsh D., Carswell R. F., Weymann R. J., 1979, Nature, 279, 381

Warren M. S., Quinn P. J., Salmon J. K., Zurek W. H., 1992, ApJ, 399, 405

Weilbacher P., Streicher O., Urrutia T., Jarno A., Pécontal-Rousset A., Bacon R., Böhm P., 2012, Proceedings of SPIE - The International Society for Optical Engineering, 8451

Weilbacher P. M., Streicher O., Urrutia T., Pécontal-Rousset A., Jarno A., Bacon R., 2014, 485, 451

Weinberg D. H., Bullock J. S., Governato F., Kuzio de Naray R., Peter A. H. G., 2015, Proceedings of the National Academy of Science, 112, 12249

White S. D. M., Rees M. J., 1978, MNRAS, 183, 341

White M., Cohn J. D., Smit R., 2010, MNRAS, 408, 1818

Wong K. C., et al., 2019, arXiv e-prints, p. arXiv:1907.04869

Wuyts S., van Dokkum P. G., Kelson D. D., Franx M., Illingworth G. D., 2004, ApJ, 605, 677

Yess C., Shandarin S. F., 1996, ApJ, 465, 2

York D. G., et al., 2000, The Astrophysical Journal, 120, 1579

Yuan L., et al., 2019, MNRAS, 487, 1315

Zhao H., 1996, MNRAS, 278, 488

Zitrin A., Broadhurst T., Barkana R., Rephaeli Y., Benítez N., 2011, MNRAS, 410, 1939

Zwicky F., 1933, Helv. Phys. Acta, 6, 110

de Haan T., et al., 2016, Astrophys. J., 832, 95

van Uitert E., et al., 2017, Mon. Not. Roy. Astron. Soc., 467, 4131

von der Linden A., et al., 2014, MNRAS, 439, 2