

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/55951> holds various files of this Leiden University dissertation

Author: Herbonnet R.T.L.

Title: Unveiling dark structures with accurate weak lensing

Date: 2017-09-26

Unveiling dark structures with accurate weak lensing

Ricardo Tian Long Herbonnet

ISBN: 978-94-6233-711-4

Unveiling dark structures with accurate weak lensing

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op dinsdag 26 september 2017
klokke 10.00 uur

door

Ricardo Tian Long Herbonnet

geboren te Rotterdam, Nederland
in 1988

Promotiecommissie

Promotor: Prof. dr. Henk Hoekstra

Co-promotor: Prof. dr. Koenraad Kuijken

Overige leden: Prof. dr. Huub Röttgering

Prof. dr. Joop Schaye

Prof. dr. Catherine Heymans (University of Edinburgh)

Dr. Anja Von der Linden (Stony Brook University)

Cover: graphics designed by Marinus de Graaf and the author. The image on the front shows galaxy cluster Abell 133 and the image on the back shows galaxy cluster Abell 2319. Both clusters are part of the Multi Epoch Nearby Cluster Survey observations used for the studies described in Chapters 4 and 5.

Contents

1	Introduction	1
1.1	Our view of the Universe	1
1.2	Structure in the Universe	2
1.3	Gravitational lensing	5
1.4	Shape measurements	8
1.4.1	Shape measurement techniques	10
1.4.2	Image simulations	11
1.5	This thesis	13
2	Shear nulling after PSF Gaussianisation	17
2.1	Introduction	18
2.2	Theory	20
2.2.1	Principles of SNAPG	20
2.2.2	Lensing basics	20
2.2.3	Effect of pre-seeing shear on observed Gauss-weighted moments	22
2.2.4	Bias as a consequence of centroiding errors	24
2.2.5	SNAPG shear nulling estimator	26
2.2.6	PSF Gaussianisation	27
2.3	Image simulations	27
2.4	Test runs	30
2.4.1	High signal-to-noise tests of SNAPG	30
2.4.2	Tests of centroid bias correction	30
2.4.3	Full test of SNAPG	32
2.5	GREAT08	34
2.6	Discussion	36
2.6.1	SNAPG formalism	36
2.6.2	Noise bias	36
2.6.3	Galaxy resolution	36
2.6.4	PSF Gaussianisation	38
2.6.5	Shear precision	38
2.6.6	Variable shear	39
2.7	Summary	39
2.7.1	Acknowledgements	40
2.A	Convolution calculations	42
3	KiDS shear calibration	45
3.1	Introduction	46
3.2	The shear measurement method	48
3.2.1	<i>lensfit</i>	48
3.2.2	Self Calibration of Noise Bias	49
3.2.3	Weight bias correction	51
3.3	Image simulations	51
3.3.1	The simulation of galaxies	51
3.3.2	Simulation volume	52
3.3.3	Input object catalogue	53

3.3.4	Simulation setup	57
3.3.5	Comparison to data	60
3.4	KiDS Calibration Method	62
3.4.1	The evaluation of shear bias	62
3.4.2	Selection bias	64
3.4.3	Calibration selection bias	65
3.4.4	<i>lensfit</i> results	69
3.4.5	Multiplicative shear bias calibration	70
3.4.6	Additive shear bias calibration and PSF properties	73
3.5	Calibration by resampling the simulated catalogue	77
3.5.1	A resampling approach to calibration	77
3.5.2	Application to the multiplicative bias in KiDS data	78
3.5.3	Robustness of the tomographic calibration	80
3.6	Calibration sensitivity analyses	80
3.6.1	Sensitivity to the magnitude distribution	80
3.6.2	Sensitivity to the galaxy size distribution	81
3.6.3	Sensitivity to accuracy of the galaxy ellipticity distribution	82
3.7	Conclusions	86
3.A	Model bias	89
4	MENeaCS weak lensing masses	91
4.1	Introduction	92
4.2	Data and shape analysis	94
4.3	Photometric source redshift distribution	96
4.4	Contamination of the source population by cluster members	101
4.4.1	Magnification	101
4.4.2	Obscuration	101
4.4.3	Excess galaxy weight density	102
4.4.4	Boost correction	105
4.5	MENeaCS cluster masses	106
4.5.1	Singular isothermal sphere	107
4.5.2	Navarro-Frenk-White profile	108
4.5.3	Hydrostatic mass comparison	111
4.5.4	Systematic error budget	112
4.6	Conclusions	113
4.A	Blank field counts	120
5	Satellite galaxy-galaxy lensing in low-<i>z</i> clusters	125
5.1	Introduction	126
5.2	Weak galaxy-galaxy lensing	128
5.2.1	Statistical errors: data covariance	129
5.3	Data set	130
5.3.1	Cluster and lens galaxy samples	130
5.3.2	Source galaxy sample	131
5.3.3	Shape measurements	133
5.4	Source sample and shear calibration	134
5.4.1	Sensitivity to background subtraction	134
5.4.2	Additive shear bias	135
5.4.3	Contamination by cluster members	136
5.4.4	More details on obscuration by cluster members	140

5.4.5	Source redshift distribution	140
5.4.6	Resulting lensing signal	142
5.5	Satellite galaxy-galaxy lensing model	144
5.5.1	Host cluster contribution	144
5.5.2	Subhalo contribution	146
5.5.3	Fitting procedure	147
5.6	The subhalo-to-stellar mass relation	147
5.6.1	The SHSMR in the context of the total-to-stellar mass relation of central galaxies	149
5.6.2	Comparison to other subhalo measurements and predictions . .	151
5.7	Subhalo mass segregation	152
5.7.1	Comparison to theoretical predictions	154
5.7.2	Previous measurements of subhalo mass segregation	155
5.8	Conclusions	156
5.A	Lens-induced bias on the shape measurements	163
6	Nederlandse Samenvatting	165
6.1	De duistere kant van het Universum	165
6.2	Zwaartekrachtslenzen	166
6.3	Dit proefschrift	168
	Curriculum Vitae	171
	Acknowledgements	173

1

Introduction

1.1 Our view of the Universe

Curiosity for the unknown has been an essential trait for humanity, propelling it forward to discover ever larger parts of the place we live in. A phenomenal change in perspective on the Universe has come about in roughly the last hundred years. With improved technology luminous nebulae were observed in the sky, which were later revealed to be extragalactic objects (Hubble 1926) and in fact are galaxies much like our own Milky Way. Later on, Hubble (1929) showed that these galaxies in the local Universe are actually moving apart *at a constant speed* H_0 (the *Hubble constant*) according to what is now known as Hubble's law. This expansion of the Universe affects its contents, and for a Universe filled with matter (and radiation), expansion reduces the temperature, so that the early Universe would be small and have extremely high temperatures. Theoretically, elementary particles in the super-heated early Universe would be locked together in equilibrium reactions until the temperature had fallen enough to break equilibrium. At a certain point in time, the decrease in temperature would bind protons and electrons together in neutrons and photons suddenly had an unimpeded path through the Universe. This sudden burst of photons happened around 13.4 billion years ago at a temperature of around 3000 K. This wave of primordial photons was observed, accidentally at first, as a remarkably homogeneous black body spectrum at 2.73 K and is known as the cosmic microwave background (CMB). The lower temperature is evidence for the cooling of the Universe as it expands. An intriguing property of the CMB is its extreme homogeneity, given the large variation in density observed in the local Universe. This discovery hinted at some unknown force of gravity. Fritz Zwicky already coined the term 'dark matter' to describe some source of gravity keeping galaxies with extremely high rotational velocities together in clusters of galaxies, as their combined visible mass was far too low (Zwicky 1937). Later, Vera Rubin found that the visible light in galaxies could not provide enough mass to sustain the high stellar velocities (Rubin et al. 1980). Both studies provided evidence for the presence of invisible mass or a incomplete understanding of gravity at cosmological scales. A Universe filled with mass should have galaxies falling towards each other due to their gravitational attraction. It was therefore a huge surprise, worthy of a Nobel prize in 2011, when two teams, who looked at the fluxes of type Ia supernovae, which have known luminosities (after calibration), found that the distances were incompatible with a Universe dominated by matter (Riess et al. 1998; Perlmutter et al. 1999). Instead, they discovered that the Universe is not just expanding, but that the expansion is happening at an accelerated pace. These observations form the basis for the hot Big Bang model of the Universe, in which everything expanded from a single super-heated point in space.

The evolution of the Universe can be remarkably well described by a relatively sim-

ple model, known as the Λ CDM model. Although this model can describe a plethora of cosmological observations, the caveat is the unknown physical origins of two of its parameters. The Λ in the Λ CDM model refers to the cosmological constant. This constant is an additional parameter in Einstein's field equations for general relativity, which causes space-time to undergo accelerated expansion. A possible explanation in the framework of the standard model of particle physics is the energy of vacuum originating from the creation and annihilation of particles and their anti-particles. However, the predicted energy density of vacuum is off by many orders to explain the observed acceleration. Instead, the origin of the observed accelerated expansion of the Universe is being called *dark energy*, which could be the cosmological constant, or some other form of energy providing negative pressure on cosmological scales, and possibly evolving over time. The CDM in the Λ CDM model stands for cold dark matter. The mysterious gravitational force, seen by Zwicky and Rubin, is explained in the Λ CDM model by a form of matter, which does not interact through the electromagnetic force and has a temperature which is low enough so that it can cluster in large quantities, hence named cold dark matter. The abundance of atomic elements is precisely predicted by the Big Bang Nucleosynthesis and rule out any possible origin of dark matter as being a known particle, so a physical explanation is still missing. The remaining components of the Universe in the Λ CDM model are baryonic matter, radiation and the curvature of the Universe. The relative abundances of the components are $\sim 70\%$ dark energy, $\sim 25\%$ dark matter and only $\sim 5\%$ baryonic matter. The current abundance of radiation is tightly set by measurements of the CMB to be negligible. A period of rapid expansion in the first second of the Universe, known as inflation, which can explain, amongst other observables, the smooth distribution of matter in the CMB, predicts a Universe with very little curvature, consistent with, for instance, CMB measurements (Planck Collaboration et al. 2016a).

The dark, not well understood components of the Λ CDM model are also the dominant components of the Universe. The nature of dark matter and dark energy are some of the biggest open questions in cosmology and astrophysics. Precise measurements of the abundances of dark matter and dark energy would allow for tests of theoretical models. Currently, some of the tightest constraints come from measurements of the CMB. However, the CMB is one snapshot of the matter distribution in the very early Universe and dark energy is an effect which has only recently become dominant. Measuring the matter distributions at different epochs can put even tighter constraints on cosmological parameters and help to unravel the origins of dark matter and dark energy.

1.2 Structure in the Universe

Matter is not distributed uniformly throughout the Universe, instead there is large spatial variation. Local variation is apparent - the Earth is an overdensity of matter in its immediate surroundings - and on much larger scales stars and their planets are mainly contained in galaxies. But there is also structure on the largest scales. Galaxies tend to cluster together in groups and there is a large variety of environments. The bottom right panel of Figure 1.1 shows observations of galaxy positions. Galaxies are preferentially clustered along thin filaments, and the knots connecting the filaments contain many tens of galaxies, and in between there are large regions devoid of galaxies. This pattern is similar to the strands in a spider's web, and hence, the distribution of galaxies in the Universe is known as the Cosmic Web.

As mentioned before, from a cosmological perspective, the observable structure in the Universe posed a problem. The CMB shows a Universe where luminous matter was distributed very uniformly across the sky. Any gravitational collapse in the photon-baryon plasma would be washed out by Coulomb interactions and Compton scattering. The highest density regions seen in the CMB are overdense by only one part in one hundred thousand. In the time since the release of the CMB these regions could not have collapsed gravitationally under their own mass to form the myriad of high density structures seen today. The key to understanding structure formation is dark matter. Unlike the baryons-photon plasma, dark matter in the early Universe could already aggregate gravitationally into overdense regions. After the emission of the CMB, when baryons were free from the stifling interactions with photons, baryons fell into the dark matter potential wells. The early collapse of dark matter facilitated the growth of cosmological baryonic structures.

In Figure 1.1 snapshots of a numerical dark matter simulation at different redshifts are shown. The final snapshot (in the bottom left panel) shows a distribution of dark matter very similar to the Cosmic Web of galaxies. These dark matter structures grow in mass through gravitational accretion of material and mergers with other structures. Structures thus form hierarchically, such that small structures form first and the largest structures of the scale of galaxy clusters arise later. This history of mergers creates two classes of galaxies: *centrals* and *satellites*. Satellite galaxies were part of distinct structures which have merged with a larger structure. As a satellite galaxy falls into a cluster, it is stripped of its reservoirs of cool gas by tidal interactions and collision with the hot cluster gas. The deficiency of the cool gas hampers the ability of satellite galaxies to form new stars, leaving them with relatively old populations of stars, which give them a tell-tale reddish colour. On the other hand centrals reside in the centres of dark matter halos and have undergone different violent events. Through gravitational attraction they accrete material, growing in size and mass by cannibalising other galaxies. Violence is also suffered by the infalling halos around satellite galaxies, which are absorbed as a subhalo in the host halo. Numerical simulations suggest that during infall, subhaloes are stripped of dark matter by dynamical friction and tidal stripping (e.g. van den Bosch et al. 2016; Moliné et al. 2017). Measurements of the fate of infalling satellites can thus provide valuable information on the validity of the Λ CDM model.

Theoretical simulations of structure formation have also shown that the distribution of matter is very dependent on cosmology. Intuitively, this is to be expected. The growth of overdensities depends on the abundance of matter and is hampered by the (accelerated) expansion of the Universe. In addition, the growth is naturally dependent on the initial conditions of the density fluctuations. The initial state of fluctuations in the Λ CDM model is assumed to be isotropic and its power spectrum is described by a power law. The cosmological parameters for this power law are the slope n_s and an amplitude set by σ_8 : the root mean square amplitude of matter fluctuations within a sphere of radius 8 Mpc. Computing the growth of structures from these initial conditions, numerical simulations can predict the abundance of large scale structure for any given cosmology. Measurements of the abundance of large scale structure at different epochs can be compared to these simulations to constrain cosmological parameters. Experiments have put constraints on Ω_m and σ_8 , by measuring the abundance of the whole of large scale structure (Heymans et al. 2013; Kilbinger 2015; Jarvis et al. 2016) and the abundance of the largest structures (Henry 2004; Vikhlinin et al. 2009; Planck Collaboration et al. 2016b).

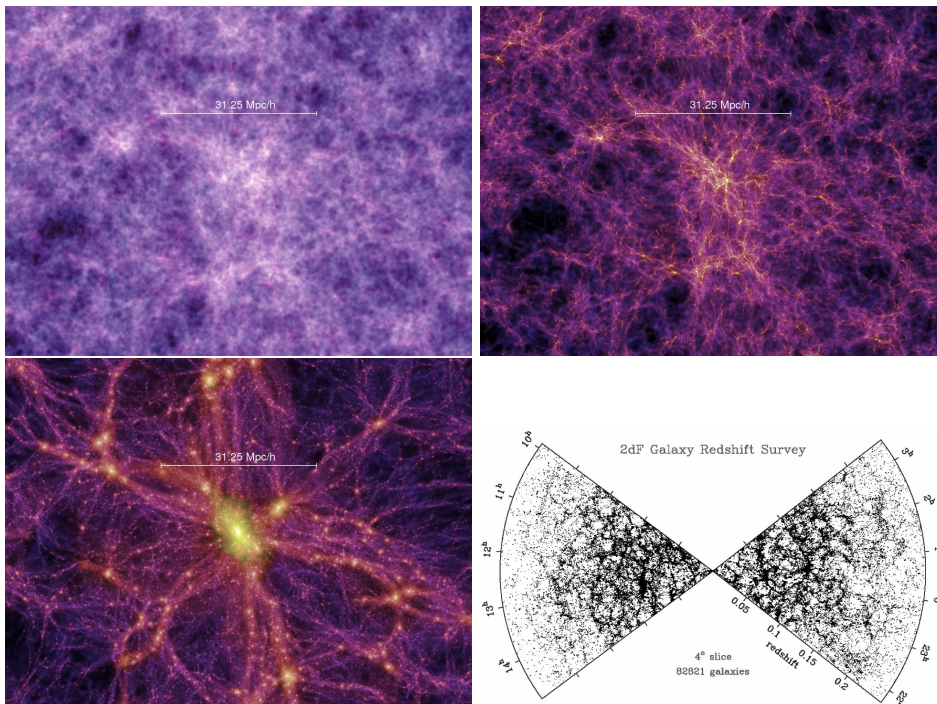


Figure 1.1: *Top left*: Snapshot at $z = 18.3$ of the distribution of cold dark matter particles in the Millennium simulation (Springel et al. 2005). *Top right*: Same area evolved to a redshift $z = 5.7$. *Bottom left*: Same area evolved to a redshift $z = 0$, showing a large variety of densities and strings of matter connecting the highest density (yellow) regions. *Bottom right*: Positions of real galaxies in the 2 Degree Field galaxy redshift survey (Peacock 2002). Due to the filamentary structure the distribution of galaxies is referred to as the Cosmic Web. The similar distribution of galaxies and dark matter is due to the early agglomeration of dark matter and baryons falling into the dark matter overdensities.

Dark matter makes up the majority of mass in the Universe and any experiments of structure formation must necessarily include measurements of dark matter halos. Baryons trace the dark matter potential, so baryonic observables can be used to probe dark matter distributions. For instance, in galaxy clusters the density and temperature of the hot gas or the velocities of member galaxies can be used to obtain a mass estimate for the clusters. However, these estimates rely on simplifying assumptions of hydrostatic or dynamical equilibrium, which cannot be guaranteed given the turbulent formation history of clusters. Moreover, the whole scala of baryonic physics is currently not fully understood and instead, a more direct estimator of dark matter distributions is needed to calibrate baryonic observables.

1.3 Gravitational lensing

The study of dark matter requires looking at the only (known) force through which dark matter interacts: gravity. Almost exactly one hundred years ago Albert Einstein postulated the theory of general relativity, which describes gravity as the curvature in space-time around a massive object. A light ray always follows a straight path through space-time, but as it passes through a curved space-time, its trajectory will change. Because of the analogy to optical lenses, this effect is known as gravitational lensing, where the massive object serves as the *gravitational lens* or *lens* for short.

The distortion of the path of light from a background source depends on the curvature of space time and the distances between source and lens, source and observer, and lens and observer. Usually the distances between source, lens and observer are much larger than the extent of the lens, so that the lens can be approximated as a thin plane in which light rays are instantaneously deflected. Figure 1.2 shows a sketch of such a configuration for a point mass lens. The figure is taken from Bartelmann & Schneider (2001), the standard text known to probably everyone who studies gravitational lensing, and here I will briefly discuss the basics of gravitational lensing and I refer to that text for more details. In Figure 1.2, a light ray emitted by a source at coordinate $\boldsymbol{\beta}$ is observed at a location $\boldsymbol{\theta}$ instead, due to the deflection at the lens plane. The deflection angle $\boldsymbol{\alpha}$ depends on the mass of the lens M and the impact parameter $\boldsymbol{\xi}$ via

$$\hat{\boldsymbol{\alpha}} = \frac{4GM}{c^2\boldsymbol{\xi}}, \quad (1.1)$$

where c is the speed of light in vacuum and G the Newtonian gravitational constant. As can be seen from the figure, the deflection angle itself is not measurable, instead we look at the angle $\boldsymbol{\beta}$ which is tied to the deflection angle according to the *lens equation*:

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta}), \quad (1.2)$$

where $\boldsymbol{\theta} = \boldsymbol{\xi}/D_d$ and $\boldsymbol{\alpha} = \hat{\boldsymbol{\alpha}}D_{ds}/D_d$. Here D_{ds}, D_d, D_s are the angular diameter distances between lens and source, observer and lens and observer and source, respectively, which can be obtained from the redshifts of lens and source. Equation 1.1 highlights the potential of gravitational lensing: a measurement of the deflection angle results in a mass estimate. However, the caveat is that gravitational lensing measures the projected 2D surface mass density along the line of sight $\Sigma(\boldsymbol{\xi})$, not the 3D mass distribution.

Gravitational lensing is produced by any massive lens with a bright background source. For example, evidence for Einstein's theory of general relativity was provided by Sir Arthur Eddington who looked at the displacement of stars behind the eclipsed

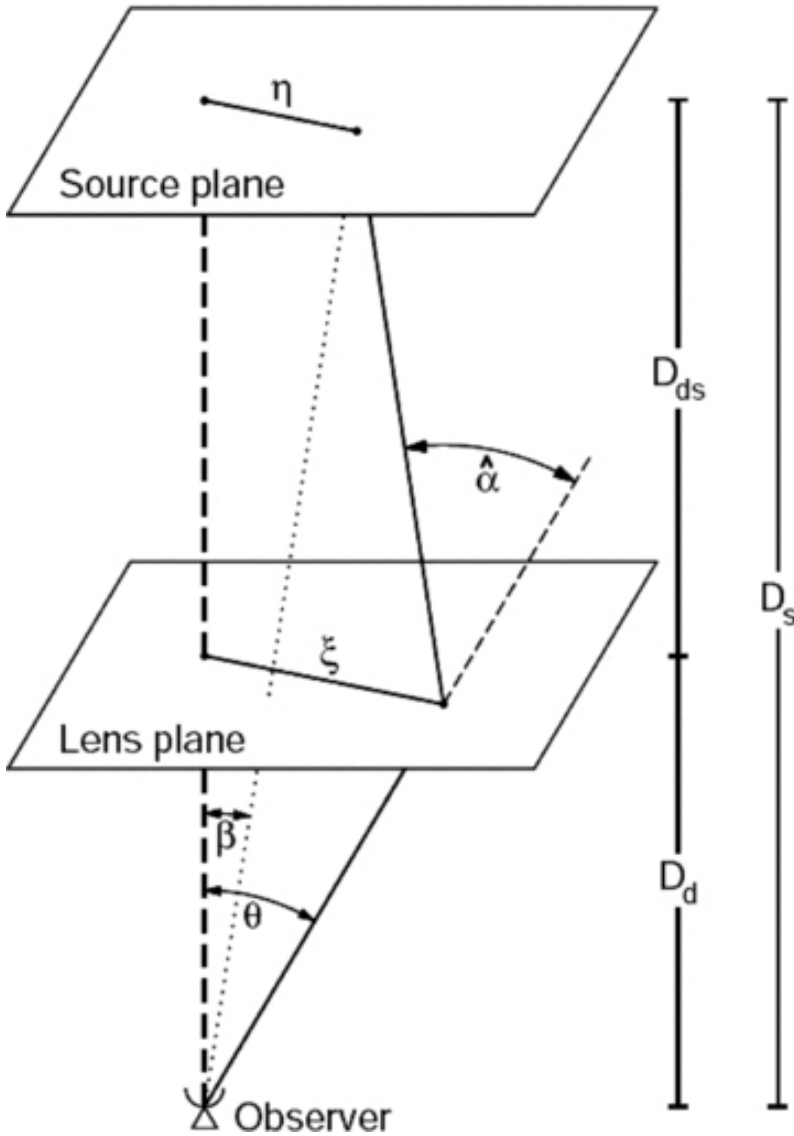


Figure 1.2: Graphical representation of the path of a light ray (shown as the solid line) from its source past a gravitational lens at the lens plane to the observer. The dashed line shows the direct line-of-sight between observer and source at location β , whereas the object is also observed at the position θ . The deflection angle α , and hence the alternate position θ , depends on the curvature induced by the massive lens and the distances between the source, lens and/or observer. These distances are usually so much larger than the extent of the curved space-time that the lens can be approximated by a single plane. Original figure in (Bartelmann & Schneider 2001).

sun in 1919. The subject of this thesis is gravitational lensing by the largest structures in the Universe, for which the light sources are distant galaxies. Galaxies are extended objects, so that the simple sketch in Figure 1.2 has to be expanded for multiple light rays. For an observed galaxy, the light profile can be described by

$$I(\boldsymbol{\theta}) = \hat{I}(\boldsymbol{\beta}(\boldsymbol{\theta})) = \hat{I}(\boldsymbol{\beta}(\boldsymbol{\theta}_0) + \mathbf{A}(\boldsymbol{\theta}_0)[\boldsymbol{\theta} - \boldsymbol{\theta}_0]), \quad (1.3)$$

where $I(\boldsymbol{\theta})$ gives the observed light intensity at angular position $\boldsymbol{\theta}$ and $\hat{I}(\boldsymbol{\beta}(\boldsymbol{\theta}))$ gives the light intensity at the source plane. As gravitational lensing conserves surface brightness, the two are equal. For the equation on the right hand side, we Taylor-expand the lens equation to first order around the location $\boldsymbol{\theta}_0$, assuming that the source is much smaller than the scales on which the lensing changes, which should be valid for large distances between observer, lens and source. The matrix \mathbf{A} is given by

$$\mathbf{A}(\boldsymbol{\theta}_0) = \left. \frac{\partial \boldsymbol{\beta}}{\partial \boldsymbol{\theta}} \right|_{\boldsymbol{\theta}_0} = \delta_{ij} - \frac{\partial^2 \phi}{\partial \theta_i \partial \theta_j} = \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix}. \quad (1.4)$$

The lensing potential ϕ is related to the surface mass density via Poisson's equation $\nabla^2 \phi(\boldsymbol{\theta}) = 2\kappa = \Sigma(\boldsymbol{\xi})/\Sigma_{\text{crit}}$, where κ is known as the *convergence* and the critical surface mass density is a geometrical factor defined as

$$\Sigma_{\text{crit}} = \frac{c^2}{4\pi G} \frac{D_s}{D_{\text{ds}} D_{\text{d}}}. \quad (1.5)$$

In Equation 1.4 we have also defined the complex *shear* γ which is related to the lensing potential via

$$\gamma = \gamma_1 + i\gamma_2 = \frac{1}{2} \left(\frac{\partial^2 \phi}{\partial \theta_1^2} - \frac{\partial^2 \phi}{\partial \theta_2^2} \right) + i \frac{\partial^2 \phi}{\partial \theta_1 \partial \theta_2}. \quad (1.6)$$

The effect of gravitational lensing on an image of a background galaxy is to magnify the image and the tidal gravitational field stretches the galaxy's observed shape. Magnification has been used to search for the most distant galaxies in the early Universe (e.g. Zitrin et al. 2014). The gravitational shear induced by a lens can be visually appreciated as spectacular luminous arcs seen in massive galaxy clusters. These examples occur only rarely when a bright background object is directly on the line of sight of a very massive matter overdensity. A more frequent form of gravitational lensing occurs when the source galaxies are not directly in the line of sight and their images are only slightly distorted. This regime is known as weak lensing and is the main subject of this thesis. The common occurrence of weak lensing makes it a powerful tool for observational cosmology, which is reflected in the large number of ongoing and upcoming weak lensing experiments. Surveys, such as the Kilo Degree Survey (de Jong et al. 2013), the Dark Energy Survey (The Dark Energy Survey Collaboration 2005), and the Hyper Suprime Cam survey (Miyazaki et al. 2012), are currently observing over thousand square degrees of the sky and in the future hemisphere-sized observations are planned with the *Euclid* satellite mission (Laureijs et al. 2011), the Large Synoptic Survey Telescope (Ivezic et al. 2008), and the Wide Field InfraRed Survey Telescope (Spergel et al. 2015). All these surveys will perform a cosmic shear analysis: a measurement of the shear-shear correlations between galaxies, which is a tracer for the distribution of dark matter structures. Comparison with theoretical cosmological models can provide constraints on the cosmological parameters governing the abundance and clustering of matter (see Kilbinger 2015 for a review). A tomographic analysis, in which galaxies

are divided into redshift bins, can also probe the time evolution of large scale structure and hence elucidate on dark energy. Tomographic cosmic shear has been hailed as one of the most powerful techniques for precise cosmological measurements.

1.4 Shape measurements

Weak gravitational lensing introduces small distortions in the observed shapes of distant galaxies, so the measurements of galaxy shapes are integral to weak lensing experiments. Projected light intensity profiles of galaxies can be very irregular and no single traditional shape can describe all galaxies, but we can generally consider galaxies as ellipses on the sky. The shape of an ellipse is fully described by a ratio between the semi-minor- and semi-major axes and the position angle, or equivalently by the two components of the ellipticity. The shear is a dimensionless spin-2 quantity, as is the ellipticity, so the ellipticity provides a natural observable for weak lensing. A galaxy with an intrinsic ellipticity will be observed after being gravitationally lensed to have an ellipticity (Seitz & Schneider 1997; Bartelmann & Schneider 2001)

$$\epsilon = \frac{\epsilon_s + g}{1 + g^* \epsilon_s} \quad \text{for } |g| \leq 1 \quad (1.7)$$

and

$$\chi = \frac{\chi_s + 2g + g^2 \chi_s^*}{1 + |g|^2 + 2\text{Re}[g \chi_s^*]}, \quad (1.8)$$

where ϵ and χ are the third flattening and third eccentricity, respectively. For an ellipse, these two definitions would be

$$\epsilon = \frac{1 - q}{1 + q} e^{2i\zeta} \quad (1.9)$$

and

$$\chi = \frac{1 - q^2}{1 + q^2} e^{2i\zeta} \quad (1.10)$$

for an axis ratio q and a position angle ζ . The subscript s for both definitions denotes the intrinsic shape of the galaxy at the source plane before it was sheared and an asterisk denotes a complex conjugate. The reduced shear g is the quantity measured in practice and it is related to the shear and convergence via $g = (1 - \gamma)/\kappa$.

The shape of a galaxy can be computed from the second order moments of the galaxy surface brightness:

$$\epsilon = \frac{Q_{20} - Q_{02} + 2iQ_{11}}{Q_{20} + Q_{02} + 2\sqrt{Q_{20}Q_{02} - Q_{11}^2}} \quad (1.11)$$

and

$$\chi = \frac{Q_{20} - Q_{02} + 2iQ_{11}}{Q_{20} + Q_{02}}. \quad (1.12)$$

Here the moments of image brightness are defined as

$$Q_{ij} = \int d^2x I(x, y) (x - x_c)^i (y - y_c)^j, \quad (1.13)$$

where x and y are coordinates in the pixel image and the moments are evaluated around the galaxy centroid (x_c, y_c) . The centroid can be estimated by finding the location where the flux Q_{00} of the galaxy is maximal, which is found by locating the coordinate (x_c, y_c) such that the first order moments $Q_{10} = Q_{01} = 0$. The moments of the image brightness provide an comprehensive set of quantities in which the pixel information in $I(x, y)$ is compressed. Combining many different orders can recover the original galaxy profile, although for weak lensing, typically, only the first few orders of moments are used.

Hypothetically, if all galaxies were circular, any anisotropy in the shape would be due to gravitational lensing. Unfortunately, galaxies have intrinsic ellipticities, which are usually much larger than the shear in the weak regime of gravitational lensing. It is therefore impossible to estimate the shear from a single galaxy. But according to the cosmological principle, galaxies should not have a preferential orientation, and galaxies in an ensemble should be round on average. This assumption of random projected shapes breaks down if there are intrinsic alignments between galaxies (see Joachimi et al. 2015 for a review). Mathematically we can rewrite Equations 1.7 and 1.8 for an ensemble average, assuming a small reduced shear $g \ll 1$ appropriate for weak lensing, to

$$\langle \epsilon \rangle \approx \langle \epsilon_s \rangle + g \approx g \quad (1.14)$$

and

$$\langle \chi \rangle \approx \langle \chi_s \rangle + 2g \approx 2g. \quad (1.15)$$

Here we assume that the intrinsic source ellipticities average to zero, so that the average observed value is a direct estimator of the gravitational reduced shear. The precision of a measurement of the shear is thus given by the number of galaxies in the ensemble. For ongoing cosmic shear surveys the precision is around a percent level, whereas future missions will have a precision of approximately one part-per-thousand.

The accuracy with which the shape of a galaxy can be measured directly affects the accuracy of the shear measurement. Although the task of measuring a shape sounds trivial, it is complicated by additional distortions to a galaxy's shape other than the gravitational shear. The excellent precision of upcoming cosmic shear surveys also puts unprecedented requirements on the accuracy of shape measurement methods. Here I review some of the main issues affecting accurate shape measurements on an image of a galaxy, although the full array of errors sources is much larger.

The effects undergone by light emitted from a distant galaxy are schematically shown in Figure 1.3 in chronological order from left to right. The leftmost panel shows the intrinsic light profile of some distant galaxy, which is sheared by gravitational lensing in the second panel. The shape of a galaxy becomes blurred as light rays travel through a turbulent atmosphere and telescope optics, which can be mathematically represented by a convolution with a point spread function (PSF). The PSF will have some shape, which is not identical to the shape of the sheared galaxy, so the convolution alters the observed shape. Although the middle panel sketches PSF convolution as a benign effect, the shape of a galaxy can be severely affected by the shape of the PSF, especially if the galaxy is small compared to the PSF. As weak lensing targets galaxies behind matter overdensities, source galaxies are often distant and thus small. The issue of the PSF is one of the main limitations in the recovery of accurate shapes and therefore, to remove the atmospheric contribution to the PSF, one of the future weak lensing experiments will be operated from a satellite. Next, the light rays hit the CCD cameras, which pixelises the galaxy image (fourth panel from the left in

The Forward Process.

Galaxies: Intrinsic galaxy shapes to measured image:

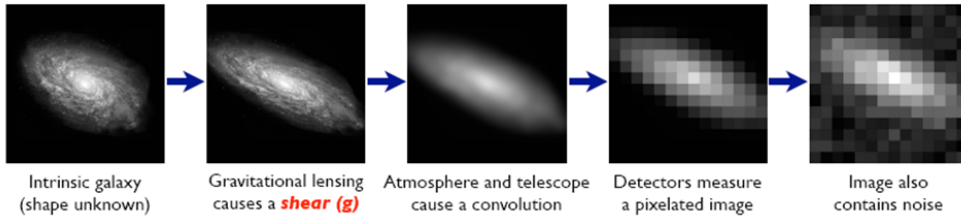


Figure 1.3: Sketch of the difficulties in shape measurement process. From left to right, the different processes affecting the light from a distant galaxy is shown in each panel, as the light rays move forward. Note that for visual effect the gravitational shear in the second panel is exaggerated by a factor ~ 10 to what is typical for weak gravitationally lensed galaxies. Original figure in (Bridle et al. 2009).

Figure 1.3). Pixelisation becomes troublesome for galaxies similar in size to the pixel scale. However, pixels are usually much smaller than the size of the PSF, and hence this issue has not received much attention. Finally, as is sketched for a very high signal-to-noise case in the rightmost panel, there is noise on the observed image from stray light, read out electronics, sky background and Poisson noise due to the finite amount of photons hitting the detector. Noise introduces an uncertainty in the observed light profile and this in turn introduces a bias in the measured shape of the galaxy: as can be seen from Equations 1.11 and 1.12 the ellipticity of a galaxy is a ratio of the observed light profile, and this non-linear dependence on the noisy data introduces a bias. Furthermore, the centroid is also measured from noisy data and as it appears non-linearly in Equation 1.13, it also adds to the *noise bias*.

1.4.1 Shape measurement techniques

A large amount of effort has been expended to find a method that can reliably measure galaxy shapes in the presence of observational nuisances. Early efforts (e.g. Kaiser et al. 1995) focused on using the moments of image brightness to estimate the ellipticity and are called *moment based methods*. The moments are measured with a weight function W which suppresses the noise at large distances from the galaxy centroid, which would otherwise dominate the integral

$$Q_{ij}^w = \int d^2x I^{\text{blur}}(x, y) W(x - x_c, y - y_c) (x - x_c)^i (y - y_c)^j. \quad (1.16)$$

The choice of the weight function is arbitrary, as long as it reduces the effect of noise on large scales. The optimal choice would be the PSF convolved galaxy image I^{blur} , but as it is not available in practice, a Gaussian is a usual choice. The use of a weight function biases the ellipticity measurement, because Equations 1.7 and 1.8 are no longer formally correct. In addition, the galaxy image has been convolved by the PSF

$$I^{\text{blur}}(x, y) = \int d^2x' I(x', y') P(x - x', y - y') \quad (1.17)$$

and the galaxy image has to be deconvolved. The PSF profile $P(x, y)$ can be obtained in practice by using stars in the observations as point sources on the sky which are

only affected by the atmosphere and telescope optics. The KSB method (Kaiser et al. 1995; Luppino & Kaiser 1997; Hoekstra et al. 1998) uses higher order (than two) moments to correct the observed third eccentricity χ for the use of the weight function and approximates the deconvolution using simplifying assumptions of the shape of the PSF. Later versions of moment-based methods have improved the algorithm by using elliptical weight functions matched to the shape of the galaxy and more accurate deconvolutions of the PSF profile (Melchior et al. 2011; Okura & Futamase 2011).

The approximate PSF deconvolution in early shape measurement methods could be overcome by forward modelling the galaxy image. In this approach, a model galaxy is generated and convolved with the PSF model. This convolved galaxy model is then fit to the observed galaxy by adjusting its properties, such as size, flux and ellipticity. The intrinsic ellipticity of the best fit galaxy model can then be used to estimate the shear. For obvious reason, this class of methods is referred to as *model fitting methods* in the literature, and two distinct galaxy models have been used: a linear combination of *shapelets* (e.g. Refregier & Bacon 2003; Massey & Refregier 2005; Kuijken 2006) and a linear combination of parametric profiles (e.g. Kuijken 1999; Miller et al. 2007; Zuntz et al. 2013; Jarvis et al. 2016). Shapelets are a set of basis functions which can describe a plethora of astronomical objects and have a well defined convolution operator. The downside of shapelets, and the reason that shapelet-based methods are not employed in ongoing large weak lensing experiments, is that the number of shapelets needs to be truncated because higher orders are more prone to noise, and the truncation leads to a biased shape measurement. Alternatively, parametric profiles provide a simpler fitting model. Sérsic profiles (Sérsic 1963) are a family of parametric radial profiles of the form $\ln [I(x, y)] \sim -(x^2 + y^2)^{1/2n}$ often used to describe galaxy light profiles. Some descriptions of galaxies used in the literature are linear combinations of Gaussians ($n = 0.5$) or the sum of an exponential disk ($n = 1$) and a De Vaucouleurs bulge ($n = 4$). Although the PSF convolution is accurately handled by this type of shape measurement method, a large assumption is made on the light profile of galaxies. The potential danger of using Sérsic profiles is that they might not capture the full morphological complexity of real galaxies, leading to a possible *model bias* in the ellipticity estimate (Voigt & Bridle 2010; Bernstein 2010).

Recently, the weak lensing community has focused more on noise bias and several new methods have been developed to reduce the effect. Noise bias is caused by a non-linear dependence of the shear estimator on the noisy image data (Hirata et al. 2004; Melchior & Viola 2012; Viola et al. 2014). For any method, noise is an addition to the galaxy image and is subject to the same mathematical formalism. Some authors have used this to calculate the effect of noise on the resulting shear estimate and implicitly correct the estimate for noise bias with notable success (Refregier et al. 2012; Okura & Futamase 2013). Alternatively, Bernstein & Armstrong (2014) have developed a Bayesian method which circumvents noise bias by removing non-linear dependence on the noisy image and the method shows good promise for the future (Bernstein et al. 2016).

1.4.2 Image simulations

A quantitative statement on the performance of shape measurement methods is necessary before they can be reliably applied on data. Given the size and depth of the survey and the desired constraints on cosmological parameters, there is a maximum

allowed bias. The biases in a shape measurement method are generally expressed by

$$g_i^{\text{meas}} = g_i^{\text{true}}(1 + m_i) + c_i, \quad (1.18)$$

where m and c are the multiplicative and additive biases, respectively, in the measured shear g_i^{meas} compared to the true shear g_i^{true} , for the two components of the shear. A multiplicative bias can arise due to noise or PSF convolution and an additive bias from any coherent sources of anisotropy, such as an imperfect correction of an elliptical PSF. Because the full process undergone by light rays is well understood, it can be simulated and these image simulations are currently the only reliable way to quantify the accuracy of shape measurement methods. Emulating telescope images will subject the method to the same difficulties faced in real observations, some of which can not be modelled otherwise, such as for instance the blending of the light of neighbouring galaxies, the impact of the detection algorithm, or stars misclassified as source galaxies.

The weak lensing community has a long history of community wide image simulations to improve overall understanding of systematic effects. The different performance of different implementations of the same method prompted the team behind the ‘shear testing programme’ (STEP; Heymans et al. 2006) to create large suites of simulated observations with known input shears. Different groups then ran their method on these simulated images to compute the shear, which when compared to the input shear gives the bias in their method. This exercise clearly showed that different choices could influence the performance. The first STEP was followed by a second programme (Massey et al. 2007) in which they searched for the perfect unbiased method. Instead, no unbiased method was found and the community turned to a new programme, which aimed to characterise the sources of bias inherent to methods. The ‘gravitational lensing accuracy testing’ (GREAT) challenges (Bridle et al. 2010; Kitching et al. 2012; Mandelbaum et al. 2015) were set up using much simpler image simulations to address individual sources of bias. These simulations contain only postage stamps of isolated galaxies and had different branches with different realistic complexities, such as constant versus varying shear profiles or parametric galaxy models versus actual galaxy images.

All of these public suites of image simulations have been of tremendous value to the weak lensing community. Each has pushed the understanding of sources of systematic error further and provides a benchmark on which to test new methods. The latest challenge has also provided the community with the well-tested software package `GalSim` with which to produce image simulations (Rowe et al. 2015). However, by systematically going through realistic features of telescope observations these simulations have shown the importance of having realistic image simulations with which to calibrate the observations. The input of the simulations has a strong effect on the bias measured from the simulations (Hoekstra et al. 2015; Kannawadi et al. 2015; Hoekstra et al. 2016) and realistic input is thus imperative for the calibration of shape measurement methods with image simulations. Different weak lensing experiments will have varying observational conditions, survey strategies and camera characteristics, which all affect the bias in the measurement. The requirement of percent level precision in the shear estimates posed by ongoing surveys already limits the use of general simulations for individual surveys. Instead, each cosmic shear survey requires its own dedicated set of image simulations (Miller et al. 2013; Jarvis et al. 2016).

1.5 This thesis

Weak gravitational lensing has the potential to provide excellent cosmological constraints from the evolution of large scale structure. However, the accuracy of weak lensing measurements is severely degraded because observational effects distort the shapes of galaxies, mimicking a shear signal. The first half of this thesis is concerned with these systematic sources of errors and how they can be mitigated.

In **Chapter 2** we develop a new shape measurement method which deals analytically with PSF convolution and noise in the image. We present the theoretical framework and test the method on simple image simulations to quantify the accuracy. We show that our method is capable of reaching subpercent accuracy even for small and noisy galaxies, which, taken at face value, is sufficient for ongoing cosmic shear surveys. However, more testing on more realistic image simulations is needed to characterise the performance for a whole host of observational effects.

At the beginning of 2016 the Kilo Degree Survey (KiDS) had observed roughly 450 square degrees of the sky. This provided us with the largest area to date for a cosmic shear analysis, but also put unprecedented requirements on the systematic uncertainties. **Chapter 3** describes the performance testing of the shape measurement algorithm used in the cosmic shear analysis. In this massive endeavour, I was responsible for the creation of a large suite of dedicated image simulations specifically designed to match the KiDS data as closely as possible. Some discrepancies between simulations and observations remained and so I performed extensive sensitivity analyses to ensure that the discrepancy did not affect the shear estimate beyond the precision afforded by the simulations. After all these tests we could confidently claim a residual multiplicative shear bias of 0.01 ± 0.01 and a negligibly small additive bias. This work has been an indispensable part of the analysis of the KiDS data and provided a calibration with enough precision for the KiDS cosmic shear analysis (Hildebrandt et al. 2017), and the calibration has since been used by every paper using the KiDS data.

The second half of this thesis focuses on weak lensing measurements using observations of galaxy clusters. For the analysis of Hoekstra et al. (2015), I created large suites of image simulations in an exercise similar to the GREAT challenges. We analysed simulations with ever increasing complexity to systematically account for various sources of error in the shear measurement pipeline, and eventually calibrated the algorithm for these biases. The increased accuracy of the shear estimates helped to obtain improved mass estimates for the sample of galaxy clusters studied in Hoekstra et al. (2015) compared to earlier work by Hoekstra et al. (2012). For this thesis I used this improved pipeline to study another large sample of clusters observed as part of the Multi Epoch Nearby Cluster Survey. Other factors affecting the accuracy of the weak lensing mass estimates are the determination of the critical surface density from the redshift distribution of source galaxies and the purity of the sample of source galaxies. I used auxiliary deep data containing reliable photometric redshift estimates to derive a source redshift distribution. With image simulations, similar to those used in **Chapter 3**, I studied the incompleteness of the population of background source galaxies due to obscuration by cluster members. Incorporating this incompleteness, I compute the purity of the source galaxy sample and statistically correct our weak lensing signal.

In **Chapter 4** we use our pipeline to measure the weak lensing masses of a large set of galaxy clusters. These total masses are combined with the sample of Hoekstra et al. (2015) and then used to determine a scaling relation with the mass estimate based on measurements of the hot cluster gas. We find that a mass dependent scaling relation is

favoured by our data over a constant bias in the gas-based mass measurements. Given the large statistical power of our full cluster sample, our findings may help to resolve the tension found between the cosmological parameters estimated from the primary CMB measurements and those estimated using the abundance of galaxy clusters.

In **Chapter 5** we apply our pipeline to the dark matter halos around satellite galaxies in the same sample of galaxy clusters. Shape measurements of galaxies are affected by light of nearby galaxies, which is a major concern in the crowded cluster environment. I calibrated the shape measurement algorithm for this effect with dedicated image simulations and determined the minimum radius from the galaxy centre for accurate weak lensing measurements. We constrain the relation between subhalo mass and stellar mass and find it to be consistent with expectations. There is no sign of significant mass segregation in our data, contrary to what has been found by other works.

Bibliography

- Bartelmann M., Schneider P., 2001, *Phys. Rep.*, 340, 291
- Bernstein G. M., 2010, *MNRAS*, 406, 2793
- Bernstein G. M., Armstrong R., 2014, *MNRAS*, 438, 1880
- Bernstein G. M., Armstrong R., Krawiec C., March M. C., 2016, *MNRAS*, 459, 4467
- Bridle S., et al., 2009, *Annals of Applied Statistics*, 3, 6
- Bridle S., et al., 2010, *MNRAS*, 405, 2044
- Henry J. P., 2004, *ApJ*, 609, 603
- Heymans C., et al., 2006, *MNRAS*, 368, 1323
- Heymans C., et al., 2013, *MNRAS*, 432, 2433
- Hildebrandt H., et al., 2017, *MNRAS*, 465, 1454
- Hirata C. M., et al., 2004, *MNRAS*, 353, 529
- Hoekstra H., Franx M., Kuijken K., Squires G., 1998, *ApJ*, 504, 636
- Hoekstra H., Mahdavi A., Babul A., Bildfell C., 2012, *MNRAS*, 427, 1298
- Hoekstra H., Herbonnet R., Muzzin A., Babul A., Mahdavi A., Viola M., Cacciato M., 2015, *MNRAS*, 449, 685
- Hoekstra H., Viola M., Herbonnet R., 2016, preprint, ([arXiv:1609.03281](https://arxiv.org/abs/1609.03281))
- Hubble E. P., 1926, *ApJ*, 64
- Hubble E., 1929, *Proceedings of the National Academy of Science*, 15, 168
- Ivezic Z., et al., 2008, preprint, ([arXiv:0805.2366](https://arxiv.org/abs/0805.2366))
- Jarvis M., et al., 2016, *MNRAS*, 460, 2245
- Joachimi B., et al., 2015, *Space Sci. Rev.*, 193, 1

- Kaiser N., Squires G., Broadhurst T., 1995, *ApJ*, 449, 460
- Kannawadi A., Mandelbaum R., Lackner C., 2015, *MNRAS*, 449, 3597
- Kilbinger M., 2015, *Reports on Progress in Physics*, 78, 086901
- Kitching T. D., et al., 2012, *MNRAS*, 423, 3163
- Kuijken K., 1999, *A&A*, 352, 355
- Kuijken K., 2006, *A&A*, 456, 827
- Laureijs R., et al., 2011, preprint, ([arXiv:1110.3193](https://arxiv.org/abs/1110.3193))
- Luppino G. A., Kaiser N., 1997, *ApJ*, 475, 20
- Mandelbaum R., et al., 2015, *MNRAS*, 450, 2963
- Massey R., Refregier A., 2005, *MNRAS*, 363, 197
- Massey R., et al., 2007, *MNRAS*, 376, 13
- Melchior P., Viola M., 2012, *MNRAS*, 424, 2757
- Melchior P., Viola M., Schäfer B. M., Bartelmann M., 2011, *MNRAS*, 412, 1552
- Miller L., Kitching T. D., Heymans C., Heavens A. F., van Waerbeke L., 2007, *MNRAS*, 382, 315
- Miller L., et al., 2013, *MNRAS*, 429, 2858
- Miyazaki S., et al., 2012, in *Ground-based and Airborne Instrumentation for Astronomy IV*. p. 84460Z, doi:10.1117/12.926844
- Moliné Á., Sánchez-Conde M. A., Palomares-Ruiz S., Prada F., 2017, *MNRAS*, 466, 4974
- Okura Y., Futamase T., 2011, *ApJ*, 730, 9
- Okura Y., Futamase T., 2013, *ApJ*, 771, 37
- Peacock J. A., 2002, in Metcalfe N., Shanks T., eds, *Astronomical Society of the Pacific Conference Series Vol. 283, A New Era in Cosmology*. p. 19 ([arXiv:astro-ph/0204239](https://arxiv.org/abs/astro-ph/0204239))
- Perlmutter S., et al., 1999, *ApJ*, 517, 565
- Planck Collaboration et al., 2016a, *A&A*, 594, A13
- Planck Collaboration et al., 2016b, *A&A*, 594, A24
- Refregier A., Bacon D., 2003, *MNRAS*, 338, 48
- Refregier A., Kacprzak T., Amara A., Bridle S., Rowe B., 2012, *MNRAS*, 425, 1951
- Riess A. G., et al., 1998, *AJ*, 116, 1009
- Rowe B. T. P., et al., 2015, *Astronomy and Computing*, 10, 121

- Rubin V. C., Ford Jr. W. K., Thonnard N., 1980, *ApJ*, 238, 471
- Seitz C., Schneider P., 1997, *A&A*, 318, 687
- Sérsic J. L., 1963, *Boletin de la Asociacion Argentina de Astronomia La Plata Argentina*, 6, 41
- Spergel D., et al., 2015, preprint, ([arXiv:1503.03757](https://arxiv.org/abs/1503.03757))
- Springel V., et al., 2005, *Nature*, 435, 629
- The Dark Energy Survey Collaboration 2005, *ArXiv Astrophysics e-prints*,
- Vikhlinin A., et al., 2009, *ApJ*, 692, 1060
- Viola M., Kitching T. D., Joachimi B., 2014, *MNRAS*, 439, 1909
- Voigt L. M., Bridle S. L., 2010, *MNRAS*, 404, 458
- Zitrin A., et al., 2014, *ApJL*, 793, L12
- Zuntz J., Kacprzak T., Voigt L., Hirsch M., Rowe B., Bridle S., 2013, *MNRAS*, 434, 1604
- Zwicky F., 1937, *ApJ*, 86, 217
- de Jong J. T. A., Verdoes Kleijn G. A., Kuijken K. H., Valentijn E. A., 2013, *Experimental Astronomy*, 35, 25
- van den Bosch F. C., Jiang F., Campbell D., Behroozi P., 2016, *MNRAS*, 455, 158

Moment-based weak lensing measurements with subpercent noise bias

Current optical imaging surveys for cosmology cover large areas of sky. Exploiting the statistical power of these surveys for weak lensing measurements requires shape measurement methods with subpercent systematic errors.

We introduce a new weak lensing shear measurement algorithm, shear nulling after PSF Gaussianisation (SNAPG), designed to avoid the noise biases that affect most other methods.

SNAPG operates on images that have been convolved with a kernel that renders the point spread function (PSF) a circular Gaussian, and uses weighted second moments of the sources. The response of such second moments to a shear of the pre-seeing galaxy image can be predicted analytically, allowing us to construct a shear nulling scheme that finds the shear parameters for which the observed galaxies are consistent with an unsheared, isotropically oriented population of sources. The inverse of this nulling shear is then an estimate of the gravitational lensing shear.

We identify the uncertainty of the estimated centre of each galaxy as the source of noise bias, and incorporate an approximate estimate of the centroid covariance into the scheme. We test the method on extensive suites of simulated galaxies of increasing complexity, and find that it is capable of shear measurements with multiplicative bias below 0.5 percent.

2.1 Introduction

The effect that masses can act as lenses and bend the path of light rays is called gravitational lensing. In the weak lensing regime first considered by Tyson et al. (1990) we statistically measure the slight distortion of the shapes of background galaxies by foreground lenses, called the shear. The subtle effects of weak gravitational lensing on galaxy shapes are an immensely powerful tool in observational astronomy. Amongst other applications, weak lensing has been an invaluable tool for cosmology through measurements of shear-shear correlations, called cosmic shear, which are connected to the dark matter power spectrum. After its first detection 15 years ago (Bacon et al. 2000; Van Waerbeke et al. 2001; Wittman et al. 2000; Kaiser et al. 2000) cosmic shear has been extensively used in cosmological studies (e.g. Kilbinger 2015; Hildebrandt et al. 2016; Jarvis et al. 2016).

Currently, large ($>1000 \text{ deg}^2$) cosmic shear surveys are ongoing, such as the Kilo Degree Survey (de Jong et al. 2013), the Dark Energy Survey (The Dark Energy Survey Collaboration 2005), and Hyper Suprime-Cam (Miyazaki et al. 2012); more hemisphere-sized missions are planned, such as LSST (Ivezic et al. 2008), *WFIRST* (Spergel et al. 2015), and *Euclid* (Laureijs et al. 2011). These surveys will observe unprecedented numbers of galaxies, pushing down statistical errors, and hence requiring percent (for ongoing missions) to subpercent level accuracies (for future missions) on the measured galaxy shapes.

In order to conduct weak lensing studies a crucial point is to measure the shapes of faint background galaxies with high accuracy as well as high precision in the face of inevitable noise, finite image resolution, and pixel effects. The first weak lensing techniques used the moments of the galaxy's image to estimate its shape and are known as moment-based methods (e.g. Kaiser 1992; Kaiser et al. 1995 (hereafter KSB); Rhodes et al. 2000). These techniques need to use a weighting function with which to cut off the moment integrals so that the moments are not dominated by noise. Having to correct for the effect of the weight function and the PSF convolution are the main challenges for this class of techniques. The widely used KSB method employs an approximate deconvolution scheme, which assumes that the PSF is nearly Gaussian. Newer moment-based methods have improved upon the PSF correction (Melchior et al. 2011), and there have been methods that change the PSF to make the measurement more exact (as explained in Hirata & Seljak 2003 and used by Mandelbaum et al. 2013; Okura & Futamase 2015, 2016).

An alternative class of techniques relies on models of galaxies which are convolved with a PSF and then fit to the galaxy image and are hence known as model-fitting methods (e.g. Kuijken 1999; Miller et al. 2013; Zuntz et al. 2013). These techniques have the benefit of an accurate treatment of the PSF, but in return require realistic models of galaxies. The model of a galaxy is usually a parametric model (e.g. a linear combination of Sersic profiles) and if it does not resemble the intrinsic galaxy, the results can be biased (Bernstein 2010; Voigt & Bridle 2010). A similar class of techniques, known as shapelets methods (Bernstein & Jarvis 2002; Refregier & Bacon 2003; Kuijken 2006), use a set of basis functions which can, in theory, model any galaxy morphology by invoking ever higher order functions. However, in practice the order has to be truncated as the higher functions are dominated by noise, again leading to an unrepresentative galaxy model. In addition, noise in the galaxy image biases all shape measurement methods due to the non-linear dependence of the galaxy's ellipticity (the usual description of its shape) on the surface brightness (Refregier et al. 2012; Melchior & Viola 2012; Viola et al. 2014).

In order to quantify these uncertainties and to find ways of calibrating the different techniques, the weak lensing community started shape measurement challenges in which teams competed by using their methods to obtain the most unbiased shear estimate. This started with a general census and benchmark tests in the STEP challenges (Heymans et al. 2006; Massey et al. 2007) and continued with GREAT challenges (Bridle et al. 2010; Kitching et al. 2012; Mandelbaum et al. 2015), which focused on the understanding of different sources of bias. After the most recent GREAT3 challenge it appears that the development in shape measurement algorithms is slowly reaching the goals set by ongoing cosmic shear surveys.

The recent improvement in accuracy was mainly due to the advanced understanding of noise bias. Several authors have introduced correction schemes into their shape measurement methods which are able to remove a large portion of the noise bias (Mandelbaum et al. 2015). An alternative route is to avoid biased shear estimators by using estimators with a linear response to the pixel data instead of traditional non-linear variables, such as the ellipticity. Several authors have used the second moments of the galaxy's image brightness to estimate the shear (Zhang & Komatsu 2011; Bernstein & Armstrong 2014; Viola et al. 2014). Recently, Bernstein et al. (2016) have reported that their Bayesian method based on moments is able to reach subpercent accuracy even with low signal-to-noise (S/N) galaxies. However, the drawback of any Bayesian analysis is the requirement of accurate priors, for which external deep observations would be required. This requirement also means that there is no shear estimate for single galaxies, as then knowledge of the intrinsic galaxy profile would be needed, but only a shear estimate for an ensemble of galaxies.

In this paper we propose a novel shape measurement method which may help to reach the ambitious goals of future cosmic shear experiments. Shear nulling after PSF Gaussianisation (SNAPG) is a moment-based method based on a circular Gaussian PSF and weight function, and requires the images to be preprocessed with a PSF Gaussianisation routine. For such galaxies we have an analytic relation between the moments of the galaxy and the shear. Shearing a population of galaxies introduces anisotropy to their ellipticity distribution. Using the analytic expressions, SNAPG reintroduces isotropy to this population by applying a nulling shear to the weighted second moments. The inverse of the nulling shear is then the shear estimate. Such a nulling technique was first advocated by Bernstein & Jarvis (2002). We propose an analytic correction to mitigate the bias due to centroid errors (Bernstein & Jarvis 2002), which is directly computed from the galaxy image.

SNAPG is similar to the Bernstein et al. (2016) method, but instead of a Bayesian framework it uses a nulling technique to extract the shear from the second moments of a population of galaxies. It does not require a prior on the intrinsic moments of the galaxy population, but instead relies on the more general requirement that galaxy ellipticities are isotropic. Our novel method thus only produces a shear value for an ensemble of galaxies, but has the benefit that no auxiliary data is needed.

In Sect. 2.2 we introduce the SNAPG concept and a correction for the bias due to centroid errors. Section 2.3 describes the image simulations we use to test SNAPG, and in Sect. 2.4 and Sect. 2.5 we present the results of the test runs. This is followed by a detailed discussion in Sect. 2.6, and a summary in Sect. 2.7.

2.2 Theory

2.2.1 Principles of SNAPG

Our novel method combines elements from a number of shear measurement methods. It follows KSB and Luppino & Kaiser (1997) in its use of Gaussian-weighted second moments, and uses a nulling technique to estimate the shear (Bernstein & Jarvis 2002). We explain the basics of moment-based methods, such as KSB and Luppino & Kaiser (1997) in Sec. 2.2.2.

Because the ellipticities used in KSB are non-linear functions of the pixel values, pixel noise makes them biased estimators. In SNAPG we work with the second moments of galaxies instead, which even in the presence of pixel noise are unbiased estimators as long as the pixel noise in the image is unbiased (as has been previously explored by Zhang & Komatsu 2011 and Viola et al. 2014). For mathematical tractability we require that the PSF in the images is Gaussian and circular; this allows us to work out analytically how the weighted moments respond to any pre-existing shear (Sec. 2.2.3). A similar exercise was done by Rhodes et al. (2000), but here we do not make any simplifying assumptions and find an expression which is valid for all values of the gravitational shear. Note that we do not try to find the intrinsic unweighted moments as in Luppino & Kaiser (1997); instead, we are only interested in the response of the weighted moments to a shear, similar to Bernstein et al. (2016).

The centroid of a source needs to be chosen before the second moments can be calculated. As this position is determined from noisy data, there is a noise dependent shift in the centroid. We show how to incorporate the uncertainty on the centroid of the galaxy into the shear estimator in the approximation where this uncertainty is distributed as a bivariate Gaussian in Sect. 2.2.4.

The response of the weighted second moments to shear can be used to find the inverse shear which counteracts the gravitational lensing shear. Hence, given the true shear, we can use the inverse shear to compute the second moments of the galaxy before it was lensed. As the true gravitational lensing shear is unknown, we cannot use each galaxy as an independent shear estimator. Rather, we use an ensemble of sheared galaxies as a probe of systematic alignments and calculate the nulling shear that needs to be applied to this ensemble to render their intrinsic ellipticity distribution isotropic. The nulling shear is then the opposite to the true shear affecting these galaxies (Sec. 2.2.5). Our approach differs from the approach of Viola et al. (2014) and Zhang & Komatsu (2011), who average the numerator and denominator of the ellipticity separately to avoid introducing biases. In SNAPG only the numerator of the ellipticity is used as a measure of the isotropy of the ellipticity distribution, and its response to shear calculated in order to null the signal.

Because real PSFs are not circular Gaussians, SNAPG can only be applied to images that have been convolved with a suitable Gaussianisation kernel (Sec. 2.2.6). Such a convolution is a linear operation on the pixels, so does not introduce noise bias in the second moments. However, it does correlate the pixel noise, the effect of which can be tracked and corrected for.

2.2.2 Lensing basics

Here we introduce the basic expressions regarding general shear estimation via the ellipticity of a galaxy as we refer to them often throughout this section. For a more detailed weak lensing review see Bartelmann & Schneider (2001).

A gravitational potential changes the path of light rays moving through it, thereby changing the observed direction of incoming light rays. For extended luminous objects different light rays can be deflected differently and thus we will observe a distorted image of a distant object. To the first order this distortion consists of a stretch (shear) and a magnification (convergence). The deflection angle of light rays from the source depends on the gradient of a suitably defined lensing potential, Ψ . The relation between the position of the source $\boldsymbol{\beta}$ and the position of the observed image $\boldsymbol{\theta}$ is known as the lens equation

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \nabla\Psi(\boldsymbol{\theta}). \quad (2.1)$$

Given that the deflection angles in weak lensing are small, the distortion can be expressed in terms of a Jacobian matrix

$$\begin{aligned} \mathbf{A} = \frac{\partial\beta_i}{\partial\theta_j} &= \left(\delta_{ij} - \frac{\partial^2\Psi(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j} \right) = \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix} \\ &\equiv \frac{1}{1 - \kappa} \begin{pmatrix} 1 - g_1 & -g_2 \\ -g_2 & 1 + g_1 \end{pmatrix}. \end{aligned} \quad (2.2)$$

The parameters

$$\begin{aligned} \kappa &= \frac{1}{2} \left(\frac{\partial^2\Psi(\boldsymbol{\theta})}{\partial\theta_1^2} + \frac{\partial^2\Psi(\boldsymbol{\theta})}{\partial\theta_2^2} \right), \\ \gamma_1 &= \frac{1}{2} \left(\frac{\partial^2\Psi(\boldsymbol{\theta})}{\partial\theta_1^2} - \frac{\partial^2\Psi(\boldsymbol{\theta})}{\partial\theta_2^2} \right), \\ \gamma_2 &= \frac{\partial^2\Psi(\boldsymbol{\theta})}{\partial\theta_1\partial\theta_2} \end{aligned} \quad (2.3)$$

are the gravitational lensing convergence κ and the two components of the shear γ_1 , γ_2 , respectively. Without information on the intrinsic size of the lensed source, only the reduced shears g_1, g_2 can be measured.

The shear affects a galaxy's polarisation according to

$$\chi^i = \frac{\chi - 2g + g^2\chi^*}{1 + |g|^2 - 2\Re(g\chi^*)}, \quad (2.4)$$

where χ and χ^i are the observed and the intrinsic, unlensed polarisation. As the intrinsic shape of a galaxy cannot be measured and the weak lensing shear is very small, the shear has to be statistically obtained from a large number of galaxies experiencing the same distortion. Assuming that galaxies are randomly oriented, the intrinsic polarisations should average out, $\langle\chi^i\rangle = 0$.

Moment-based methods construct the polarisation of an object from the second moments of image brightness Q_{ij}

$$\chi = \frac{Q_{11} - Q_{22} + 2iQ_{12}}{Q_{11} + Q_{22}}. \quad (2.5)$$

These moments are defined as the noiseless unweighted moments on the intrinsic galaxy image $I^i(\mathbf{x})$,

$$Q_{ij}^i = \int d\mathbf{x} I^i(\mathbf{x}) x_i x_j W(\mathbf{x}), \quad (2.6)$$

with the weight function $W(\mathbf{x}) = 1$. However, in practice a galaxy is observed convolved with a PSF $P(\mathbf{x})$,

$$I^o(\mathbf{x}) = \int d\mathbf{x}' I^i(\mathbf{x}') P(\mathbf{x} - \mathbf{x}'), \quad (2.7)$$

and the weight function $W(\mathbf{x})$ that goes to zero at large \mathbf{x} is required for the moments not to be dominated by the noise on the image. The aim of moment-based methods is then to estimate the intrinsic polarisations by correcting for the weight function and PSF.

2.2.3 Effect of pre-seeing shear on observed Gauss-weighted moments

In the case when the PSF is Gaussian, we can reconstruct what the second moments would have been if the galaxy had been sheared.

The weighted second moments of the observed image are

$$\begin{aligned} Q_{ij}^o &= \int d\mathbf{x} I^o(\mathbf{x}) x_i x_j W(\mathbf{x}) \\ &= \int \int d\mathbf{x} d\mathbf{x}' I^i(\mathbf{x}') P(\mathbf{x} - \mathbf{x}') x_i x_j W(\mathbf{x}), \end{aligned} \quad (2.8)$$

with $W(\mathbf{x})$ a weight function that depends only on $|\mathbf{x}|$. The order of integration can be swapped and Eq. 2.8 rewritten as

$$Q_{ij}^o = \int d\mathbf{x}' I^i(\mathbf{x}') \left[\int d\mathbf{x} P(\mathbf{x} - \mathbf{x}') x_i x_j W(\mathbf{x}) \right], \quad (2.9)$$

and so relate the weighted second moments directly to the intrinsic galaxy shape as an integral weighted by the expression in square brackets (which depends only on the weight function and PSF).

A gravitationally lensed source has a distorted image: the intrinsic image $I^i(\mathbf{x})$ is transformed to

$$I^A(\mathbf{x}) = I^i(\mathbf{A}\mathbf{x}) \quad (2.10)$$

by the distortion matrix \mathbf{A} .

In order to measure the gravitational shear, we need to know the weighted second moments that we would observe if the galaxy had been distorted by a distortion matrix \mathbf{A} before PSF convolution; these can be written as

$$\begin{aligned} Q_{ij}^A &= \int d\mathbf{x}'' I^A(\mathbf{x}'') \left[\int d\mathbf{x} P(\mathbf{x} - \mathbf{x}'') x_i x_j W(\mathbf{x}) \right] \\ &= \int \frac{d\mathbf{x}'}{|\det \mathbf{A}|} I^i(\mathbf{x}') \\ &\quad \times \left[\int d\mathbf{x} P(\mathbf{x} - \mathbf{A}^{-1}\mathbf{x}') x_i x_j W(\mathbf{x}) \right] \end{aligned} \quad (2.11)$$

by means of Eq. 2.10 and the transformation $\mathbf{x}' = \mathbf{A}\mathbf{x}''$. We now show how the moments Q_{ij}^A of the sheared source can be derived from the observed, PSF-convolved image $I^o(\mathbf{x})$, by constructing a new weight function $W_{ij}^A(\mathbf{x})$ which satisfies, for arbitrary \mathbf{x}' ,

$$\begin{aligned} \int \frac{d\mathbf{x}}{|\det \mathbf{A}|} P(\mathbf{x} - \mathbf{A}^{-1}\mathbf{x}') x_i x_j W(\mathbf{x}) \\ = \int d\mathbf{x} P(\mathbf{x} - \mathbf{x}') W_{ij}^A(\mathbf{x}). \end{aligned} \quad (2.12)$$

It is easy to see from Eq. 2.12 that integrating the observed (PSF-convolved but un-sheared) image times the weight function W_{ij}^A will give the moments Q_{ij}^A . Equation 2.12 shows that W_{ij}^A can be constructed from the original weight function W and the PSF P by the following sequence of operations:

1. Convolution $x_i x_j W$ with the PSF;
2. Distorting the result of the previous step with distortion matrix \mathbf{A} and divide by $|\det \mathbf{A}|$;
3. Deconvolving the result of the previous step by the PSF.

This recipe is valid as long as the deconvolution in the final step is well defined.

We do not attempt to solve the general problem, but concentrate on the simpler case where both the PSF P and the weight function W are round Gaussians:

$$P(\mathbf{x}) = \frac{1}{2\pi p^2} e^{-|\mathbf{x}|^2/2p^2} \quad (2.13)$$

and

$$W(\mathbf{x}) = e^{-|\mathbf{x}|^2/2w^2}. \quad (2.14)$$

In Appendix 2.A we derive an expression (Eq. 2.34) for the convolution of G_1 with $x_i x_j G_2$, where G_1 and G_2 are Gaussians of arbitrary covariance matrices \mathbf{P} and \mathbf{V} respectively. By substituting $\mathbf{V} = w^2 \mathbf{1}$, $\mathbf{P} = p^2 \mathbf{1}$, and $\mathbf{y} = \mathbf{A}^{-1} \mathbf{x}'$ in Eq. 2.34, we can write the left-hand side of Eq. 2.12 as

$$\begin{aligned} & \int \frac{d\mathbf{x}}{|\det \mathbf{A}|} P(\mathbf{x} - \mathbf{A}^{-1} \mathbf{x}') x_i x_j W(\mathbf{x}) \\ &= \frac{w^4}{(w^2 + p^2)^2} \frac{e^{-\frac{1}{2} |\mathbf{A}^{-1} \mathbf{x}'|^2 / (w^2 + p^2)}}{|\det \mathbf{A}|} \\ & \times \left[p^2 \delta_{ij} + \frac{w^2}{w^2 + p^2} (\mathbf{A}^{-1} \mathbf{x}')_i (\mathbf{A}^{-1} \mathbf{x}')_j \right]. \end{aligned} \quad (2.15)$$

The final step is now to deconvolve this expression by the PSF in order to obtain an expression for the weight function W_{ij}^A that satisfies Eq. 2.12. We first calculate the result of convolving with a general Gaussian PSF of covariance matrix \mathbf{P} ; the deconvolution we seek is then obtained by setting $\mathbf{P} = -p^2 \mathbf{1}$ (note the sign). The first term (involving δ_{ij}) is straightforward: convolving two Gaussians results in a new Gaussian with covariance matrix equal to the sum. The second term can be calculated using the result of Eq. 2.34 in Appendix 2.A by setting the matrix \mathbf{V} defined there to $\mathbf{V} = (w^2 + p^2) \mathbf{A}^2$.

After some work we find

$$\begin{aligned} W_{ij}^A(\mathbf{x}) &= w^4 \frac{e^{-\frac{1}{2} \mathbf{x}^T \mathbf{B}^{-1} \mathbf{x}}}{|\det \mathbf{B}|^{\frac{1}{2}}} \\ & \times \left[\delta_{ij} + w^2 \left((\mathbf{A} \mathbf{B}^{-1} \mathbf{x})_i (\mathbf{A} \mathbf{B}^{-1} \mathbf{x})_j - (\mathbf{A} \mathbf{B}^{-1} \mathbf{A})_{ij} \right) \right] \end{aligned} \quad (2.16)$$

with

$$\mathbf{B} = (w^2 + p^2) \mathbf{A}^2 - p^2 \mathbf{1}. \quad (2.17)$$

The weight function W_{ij}^A is only useful in practice if it tends to zero at large $|\mathbf{x}|$. This is the case as long as the distortions are small enough so that both eigenvalues of \mathbf{B} are positive, which is true when

$$\kappa + \gamma < 1 - \frac{p}{\sqrt{w^2 + p^2}}. \quad (2.18)$$

As long as the weight function is wider than the PSF ($w > p$, a reasonable choice if one wants to avoid unnecessarily noisy measurements), this means a useful W_{ij}^A can be constructed for $\kappa + \gamma$ at least up to 0.3. We show the form of the weight function W_{ij}^A for a grid of $(g_1, g_2, \kappa = 0)$ in Fig. 2.1.

2.2.4 Bias as a consequence of centroiding errors

Applying the filter W_{ij}^A in Eq. 2.16 to an image yields unbiased estimates for the post-seeing weighted second moments of the source about $\mathbf{x} = \mathbf{0}$ as long as the noise on each pixel is unbiased. However, in reality the true centre of the source is unknown, and must be estimated from the image itself. The associated scatter in the centroid biases the second moments. In this section we quantify that bias.

Suppose that \mathbf{x}_c is a noisy estimate of the centroid of the observed image I^o . Then, using this centroid our estimate for Q_{ij}^A is

$$\tilde{Q}_{ij}^A = \int d\mathbf{x} I^o(\mathbf{x}) W_{ij}^A(\mathbf{x} - \mathbf{x}_c). \quad (2.19)$$

If the error distribution of the centroids is $f(\mathbf{x}_c)$ then the expectation value of \tilde{Q}_{ij}^A is

$$\begin{aligned} \langle \tilde{Q}_{ij}^A \rangle &= \int d\mathbf{x}_c f(\mathbf{x}_c) \tilde{Q}_{ij}^A \\ &= \int d\mathbf{x} I^o(\mathbf{x}) \int d\mathbf{x}_c f(\mathbf{x}_c) W_{ij}^A(\mathbf{x} - \mathbf{x}_c), \end{aligned} \quad (2.20)$$

i.e. the weight function that determines \tilde{Q}_{ij}^A is the original weight function W_{ij}^A convolved with the centroid error distribution f . Hence, conversely, an unbiased estimate of Q_{ij}^A is obtained by using a weight function \hat{W}_{ij}^A obtained by deconvolving W_{ij}^A by the centroid error distribution f . We assume that f is Gaussian, of covariance \mathbf{C} . Remembering that expression 2.16 for W_{ij}^A was itself obtained by deconvolving Eq. 2.15 by the PSF, we see that \hat{W}_{ij}^A is the deconvolution of Eq. 2.15 by $P \otimes f$, i.e. by a Gaussian of covariance matrix $p^2 \mathbf{1} + \mathbf{C}$. As noted under Eq. 2.17, this deconvolution is simply accomplished by using Eq. 2.16 with a modified \mathbf{B} matrix

$$\hat{\mathbf{B}} = (w^2 + p^2) \mathbf{A}^2 - p^2 \mathbf{1} - \mathbf{C}. \quad (2.21)$$

It remains for us to quantify the covariance matrix \mathbf{C} of the centroid error for a given source. This will depend on the recipe used to determine the centre.

We centre each source by finding the peak of the correlation of its (noisy) image $I^n(\mathbf{x})$ with a suitable centring kernel f , equivalent to finding the optimum positional match between I^n and f . The centroid \mathbf{c} found this way satisfies

$$\begin{aligned} 0 &= \frac{\partial}{\partial c_i} \int I^n(\mathbf{x}) f(\mathbf{x} - \mathbf{c}) d\mathbf{x} \\ &= \int I^n(\mathbf{x}) (f_{,i}(\mathbf{x}) - c_{j,i} f_{,j}(\mathbf{x}) + \dots) d\mathbf{x} \end{aligned} \quad (2.22)$$

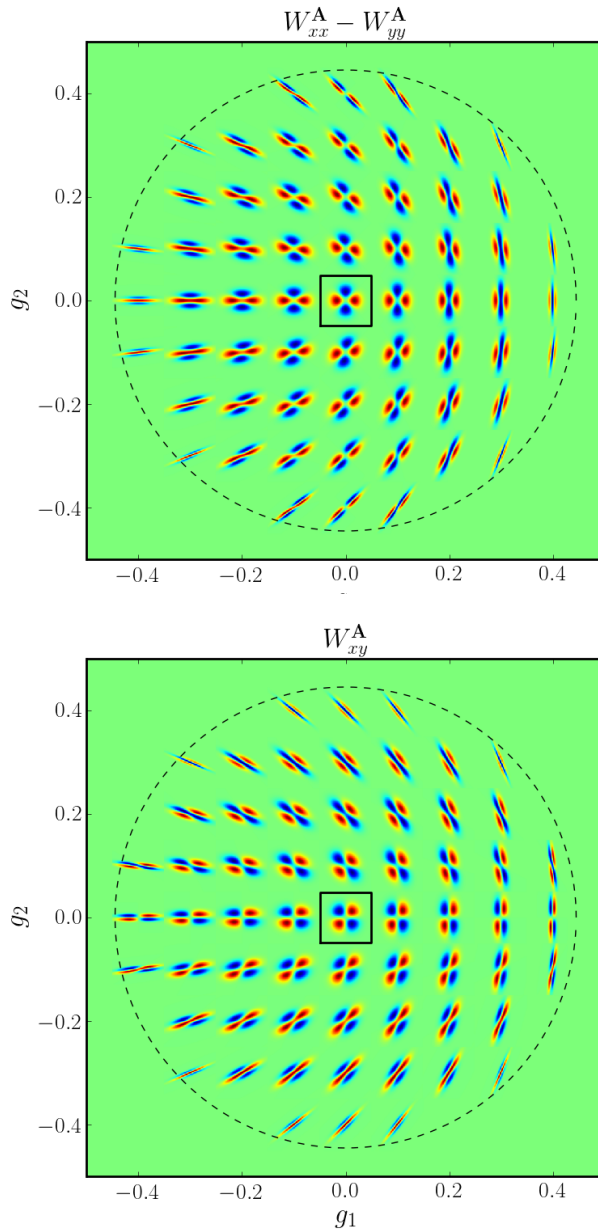


Figure 2.1: Example shear filter functions W_{ij}^A , for filter and PSF size $w = 3$, $p = 2$, ($\kappa = 0$). Integrating an image multiplied with the filter function for a particular shear (g_1, g_2) yields the weighted second moments the image would have had if it had been sheared by that amount before seeing convolution. The central box is 25 units on a side, and the dashed line indicates the maximum shear value that can be applied for the given w and p (see Eq. 2.18). Red is positive, blue is negative. Left: Filters corresponding to $Q_{xx} - Q_{yy}$. Right: Filters corresponding to Q_{xy} .

where we used a Taylor expansion on f about $\mathbf{c} = \mathbf{0}$, assumed to be the true centre of our source. To derive the noise properties of \mathbf{c} we first separate $I^m(\mathbf{x})$ into the noise-free observed image $I^o(\mathbf{x})$ and a noise field $\Delta(\mathbf{x})$, and obtain the first-order relation between Δ and \mathbf{c} :

$$c_j \int I^o(\mathbf{x}) f_{,ij} d\mathbf{x} = \int \Delta(\mathbf{x}) f_{,i}(\mathbf{x}) d\mathbf{x}. \quad (2.23)$$

To calculate the covariance $\overline{c_k c_l}$ we define F_{ij} as the integral on the left-hand side of Eq. 2.23, and we assume the background-limited case in which pixel noise is stationary, of constant covariance matrix $N(\mathbf{x} - \mathbf{x}')$ across a source image. Squaring Eq. 2.23 and averaging over all possible noise realisations then yields

$$\overline{c_k c_l} F_{ik} F_{jl} = \int \int f_{,i}(\mathbf{x}) f_{,j}(\mathbf{x}') N(\mathbf{x} - \mathbf{x}') d\mathbf{x} d\mathbf{x}' \equiv H_{ij}, \quad (2.24)$$

and hence the covariance matrix \mathbf{C} needed in Eq. 2.21 is given by

$$\mathbf{C} = \mathbf{F}^{-1} \mathbf{H} \mathbf{F}^{-1}. \quad (2.25)$$

Here, \mathbf{H} depends only on the kernel function f and the pixel noise properties N , and \mathbf{F} can be estimated from the noisy image I^m . If the kernel f is circular and the noise covariance matrix isotropic $N(\mathbf{x} - \mathbf{x}') = \sigma_n^2 \delta(\mathbf{x} - \mathbf{x}')$, where σ_n is the root mean square of the noise background, then \mathbf{H} becomes a scalar.

A convenient choice is a Gaussian

$$f(\mathbf{x}) = e^{-|\mathbf{x}|^2/2a^2} \quad (2.26)$$

for which

$$H_{ij} = \frac{\pi \sigma_n^2 \delta_{ij}}{2} \quad (2.27)$$

and

$$f_{,ij}(\mathbf{x}) = \left(\frac{x_i x_j}{a^4} - \frac{\delta_{ij}}{a^2} \right) e^{-|\mathbf{x}|^2/2a^2}. \quad (2.28)$$

2.2.5 SNAPG shear nulling estimator

In the previous sections we constructed the filter $W_{ij}^{\mathbf{A}}$ which, when applied to an observed Gaussian-PSF smeared image, yields the Gauss-weighted second moments $Q_{ij}^{\mathbf{A}}$ that would have been observed (with the same PSF) had the galaxy been distorted by distortion matrix \mathbf{A} . We have also quantified the noise bias on $Q_{ij}^{\mathbf{A}}$ due to centroiding errors, and constructed a modified filter $\hat{W}_{ij}^{\mathbf{A}}$ that compensates for it. In what follows we will drop the ‘‘hat’’ notation and assume that the centroid error correction is applied.

The weight function $W_{ij}^{\mathbf{A}}$ can be used to construct a shear estimator. If a galaxy is sheared by some known distortion matrix \mathbf{A} , then we can use the inverse of \mathbf{A} to find the intrinsic second moments of the galaxy. For a large ensemble of galaxies their combined intrinsic ellipticities (or equivalently their Stokes parameters) average out to zero. Then the search is for the distortion matrix \mathbf{A} which can null the Stokes parameters ($Q_{11}^{\mathbf{A}} - Q_{22}^{\mathbf{A}}, 2Q_{12}^{\mathbf{A}}$) of a sheared population of galaxies. The inverse of that distortion matrix is a good estimator of the shear those galaxies experience. To efficiently search for the distortion matrix we use a nulling scheme similar to one already used in shape measurements (Bernstein & Jarvis 2002). In practice, a trial distortion matrix \mathbf{A} is

chosen and the corresponding weight function W_{ij}^A is computed (see Fig. 2.1), with which the Stokes parameters for the ensemble of galaxies are calculated. Based on the (an)isotropy of the Stokes parameters a new distortion matrix is chosen, and the previous steps are repeated to reassess the isotropy. This procedure converges in roughly four trials, after which the inverse of the distortion matrix is taken as the shear estimate.

Because galaxies have a wide range of brightness, the Stokes parameters of a galaxy population have a large variance. This translates into a large variance in the nulling shear and increasing precision would require large numbers of galaxies. Alternatively, the moments could be weighted by flux or size to reduce the variance, but this would introduce a bias in the shear. In our current tests such a weight is not required, but we discuss possible solutions for future work in Sec. 2.6.5.

2.2.6 PSF Gaussianisation

As indicated in the beginning of Sec. 2.2.3, SNAPG relies on the assumption of a circular Gaussian PSF. Such a PSF is never present in observational data and thus we need to transform the actual observed PSF into the required PSF. We employ a Gaussianisation process which creates a circular Gaussian PSF by convolving the observed PSF with an appropriate kernel.

Gaussianisation starts by creating a shapelets model of the PSF. Shapelets are a set of basis functions of Gauss-Hermite polynomials, which can be linearly combined to model astronomical objects (Refregier 2003). Convolution in shapelet space is a straightforward procedure, making shapelets an ideal basis for the Gaussianisation process. We use the shapelet implementation of Kuijken (2006) to create a shapelet model of the PSF. In practice, bright stars can be used to obtain a model of the PSF. A best fit circular Gaussian of the shapelets model of the PSF is determined. Then a convolution kernel is found that convolves the PSF into the best fit Gaussian. The resulting kernel is applied to the whole image to create galaxies with circular Gaussian PSFs. See Kuijken et al. (2015) for more detail on the process of PSF Gaussianisation.

It is worth noting that this procedure is different from the one presented by Hirata & Seljak (2003). They assume a Gaussian form for the intrinsic shape of the galaxy when calculating the corrections for PSF non-Gaussianity, whereas our procedure is valid for any galaxy morphology. However, it does rely on well-sampled data and was designed with only ground-based PSFs in mind. It is unclear how the procedure would perform for diffraction-limited space telescopes.

The convolution mixes information from neighbouring pixels and hence introduces a correlation between the noise on different pixel values. The resulting noise covariance matrix $N(x-x')$ is given by the original image's pixel variance, multiplied by the auto-correlation function of the convolution kernel. It is important to propagate this noise covariance into the centroid error estimate (Eq. 2.24).

2.3 Image simulations

To test the performance of SNAPG we create simulated images of galaxies with known applied shear. Following the image simulations of the GREAT challenges, we create a grid of isolated galaxies on postage stamps. This approach gives us a clean test of the performance of SNAPG without introducing errors related to blended galaxy isophotes (see Hoekstra et al. 2015 for a discussion on how blends affect shear measurements).

Table 2.1: Overview and specifications of all simulated images used to test the performance of SNAPG.

Set	PSF	Galaxy type	
Well resolved	Gaussian	Exponential	
Barely resolved	Gaussian	Exponential	
GREAT08 RNK	Gaussianised Moffat	Exponential or de Vaucouleurs	
GREAT08 LNK	Gaussianised Moffat	Exponential or de Vaucouleurs	

Set	PSF HLR	Galaxy HLR	S/N
Well resolved	1.76 pixels	2.5 pixels	$\sim 5 - 100$
Barely resolved	1.76 pixels	1.5 pixels	$\sim 5 - 100$
GREAT08 RNK	1.72 pixels	2.1 or 10 pixels	~ 200
GREAT08 LNK	1.72 pixels	2.1 or 10 pixels	~ 20

The images of the GREAT challenges do not have circular Gaussian PSFs, so for a clean test of the SNAPG framework we use `GalSim` (Rowe et al. 2015) to create our own image simulations with perfect circular Gaussian PSFs.

Our simulated galaxy images are a grid of 100 x 100 galaxies, all with exponential profiles of the same size. The grid of postage stamps have a single galaxy randomly offset from the centre of the stamp. The postage stamp is large enough to avoid any bias due to truncation of the surface brightness profile. The half light radius (HLR) of the galaxies is 2.5 pixels. The flux of all galaxies is the same, so that when noise is added all galaxies will have the same S/N . The modulus of the ellipticity of a galaxy is randomly drawn from a Rayleigh distribution of width 0.25, cut off at 0.6 to avoid artificial truncation by the edge of the postage stamp. The position angle of the galaxy is taken from a random uniform distribution between 0 and 180 degrees. The galaxy models are convolved with a Gaussian PSF with a half light radius of 1.76 pixels. The size of the galaxies is larger than the size of the PSF, so we call this set of images the *well resolved* sample. Each image has a constant shear applied to all 10000 galaxies, where the shear is taken from a grid of $(-0.04, -0.035, -0.03, \dots, 0.04)$ for each shear component separately, resulting in 289 different g_1, g_2 pairs.

We also create a similar set of images where the galaxy half light radius is set to 1.5 pixels. The half light radius of the Gaussian PSF is 1.76 pixels, so that the PSF is larger than the galaxy. We call this set of images the *barely resolved* sample. Fluxes are fixed and the ellipticity is sampled in the same way as described above. Here too, constant shears are applied to all galaxies on an image, and the shear is taken from the same grid.

These two suites of image simulations contain a total of $5.78 \cdot 10^6$ galaxies. These galaxy images do not contain any noise, instead Gaussian noise is added as required for each test. Each set represents a different target for shape measurement. The well resolved images present our fiducial dataset as the galaxy shapes are not badly affected by the PSF and provide us with a benchmark test of the performance of SNAPG. The barely resolved images present a challenging sample, as galaxy shapes are heavily influenced by pixelisation and severely blurred by the PSF. These galaxies are a difficult target for most shape measurement methods and are sometimes cut from the sample owing to the uncertainty in the galaxy shapes. However, faint small galaxies are abundant in observations and their removal presents a serious loss of statistical power. Having a shear measurement technique able to reliably measure such objects

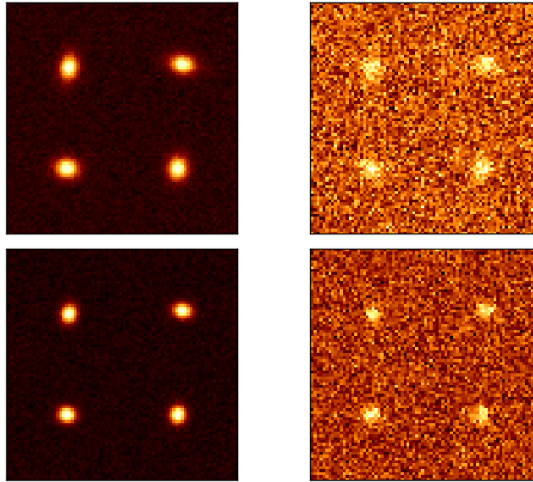


Figure 2.2: Examples of the simulated galaxies at different noise levels to help visualise varying S/N levels, and well resolved in contrast to barely resolved galaxies. *Top:* Cut-out of the well resolved sample of images for $S/N \approx 100$ (left) and $S/N \approx 10$ (right). *Bottom:* Cut-out of the barely resolved sample of images for $S/N \approx 100$ (left) and $S/N \approx 10$ (right). The shape of the barely resolved galaxies is rounded by the PSF and for low S/N both well resolved and barely resolved galaxies the shapes are very much affected by noise.

will be a huge advantage for future weak lensing experiments.

As a visual aid to interpreting the different sets of simulated galaxy images we show some of our mock galaxy images in Fig. 2.2. The upper images show the fiducial well resolved galaxies and the lower images show galaxies from the challenging set of barely resolved galaxies. The images have Gaussian noise added so that the mean $S/N \approx 100$ (left panel) and $S/N \approx 10$ (right panel), where S/N is defined as the $\text{FLUX}/\text{FLUXERR}$ measured by SExtractor on default settings (Bertin & Arnouts 1996). A summary of the image properties can be found in Table 2.1.

The images of the GREAT08 challenge (Bridle et al. 2010) provide us with a test of the PSF Gaussianisation. In addition, we can compare the performance of SNAPG to other tested methods. We use the 15 LowNoise_Known (LNK) and 300 RealNoise_Known (RNK) sets of images from the challenge, where each image has 10000 isolated galaxies in postage stamps of 40 pixels across. All 10000 galaxies in an image have the same shear applied to provide ample statistics. The galaxies are either an exponential or a de Vaucouleurs profile with a fiducial $S/N = 200$ for LNK and $S/N \approx 20$ for RNK. The sizes of galaxies are set so that the PSF convolved galaxy size is 1.4 times larger than the PSF size. We use the PSF Gaussianisation algorithm explained in Sec. 2.2.6 to outfit the GREAT08 images with a circular Gaussian PSF. The Gaussianisation algorithm is applied to PSF `set0001`, which is a Moffat profile of full width half maximum 2.85 truncated at 5.7 pixels and ellipticity components $e_1 = -0.019$ and $e_2 = -0.007$. The main properties of the GREAT08 images are summarised in Table 2.1.

2.4 Test runs

The new shear nulling method is coded in `python` and we apply it to the image simulations described in the previous section. The code returns the shear value g_i that nulls the average distortion of each 10,000-galaxy image. First we test the SNAPG formalism, then the centroid bias correction formalism, and finally a full implementation of SNAPG. Throughout this section we use $a=3$ and $w=3$ for the widths of the centroid and moment weight functions, respectively. As noted above, none of the images contains noise. Instead, we add noise for each test as storing each noise realisation presented storage problems. For every level of added noise we calculate the mean signal-to-noise ratio (S/N) using SExtractor with default settings and defined as `FLUX/FLUXERR`. Each measurement we present was obtained by using the full set of $2.89 \cdot 10^6$ galaxies in each set of simulations described in the previous section.

The performance of SNAPG is measured by performing a linear fit using the functional form $g_{i,\text{out}} = (1 + m_i)g_{i,\text{true}} + c_i$ (Heymans et al. 2006) for each shear component g_i . This procedure quantifies the shear bias as a multiplicative term m_i (e.g. arising from method assumptions or noise) and an additive term c_i (arising from imperfect corrections for the elliptical PSF). Because our simulations are ideal with a circular Gaussian PSF we do not expect any additive bias in these tests.

2.4.1 High signal-to-noise tests of SNAPG

We start by quantifying the performance of SNAPG on the fiducial set of images with well resolved galaxies for $S/N=100$ for the true centroid of the galaxy. In the left panel of Fig. 2.3 we show the measured residuals between the input shear and the measured one, and the true input versus the measured shear. We find $\langle m \rangle = (+1.6 \pm 1.9)10^{-4}$ and $\langle c \rangle = (-0.1 \pm 0.1)10^{-4}$ for the average of the two components of the shear. We also test the algorithm on the set of barely resolved galaxies and the result is plotted in the right panel. Again SNAPG retrieves the applied shear without detectable bias: $\langle m \rangle = (-0.9 \pm 2.3)10^{-4}$ and $\langle c \rangle = (-0.1 \pm 0.1)10^{-4}$.

As expected SNAPG returns unbiased shear estimates for our ideal images with circular Gaussian PSF with high S/N , showing that the pipeline works. These tests also show the potential of SNAPG as a shear measurement method, regardless of the size of the galaxy in relation to the size of the PSF. We note that at this high S/N the results remain unchanged if the centroid is measured from the data.

2.4.2 Tests of centroid bias correction

We expect a bias in the shear estimate to originate from the random error on the measured centroid due to image noise. In order to test the effectiveness of the centroid bias correction proposed in Sec. 2.2.4, we perform a test with truly random centroid values. The well resolved images are re-analysed with SNAPG, but the centroids are artificially offset from the true centroid by a random Gaussian value. The error on the centroid is taken from a normal distribution with a standard deviation of 0.5 pixels. Such a distribution would occur for our simulated images with a $S/N \approx 5.5$ for $a = 3$.

Besides introducing random centroid errors, we also add Gaussian noise to the images before measuring the shear. The addition of noise, which is uncorrelated to the centroid error, should not bias SNAPG as the moments are linear with respect to the noisy surface brightness. Hence, we analyse the images with SNAPG several times where each time Gaussian noise with a different root mean square is added to the

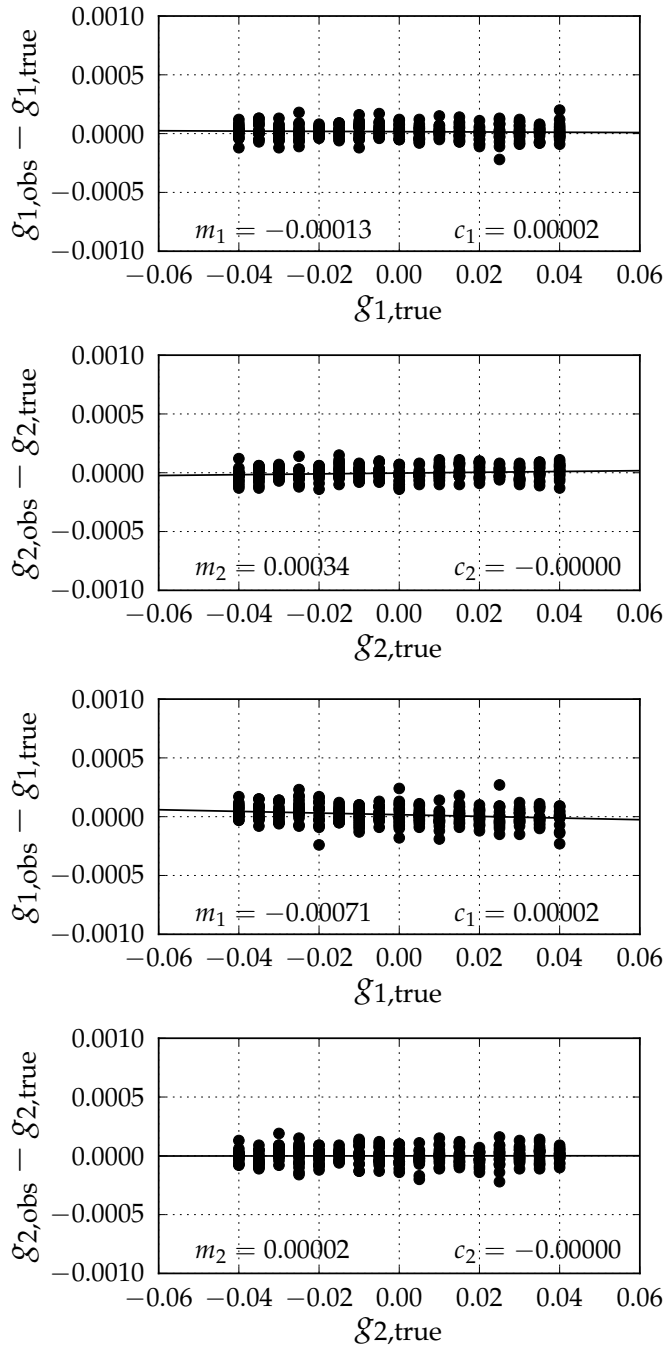


Figure 2.3: Shear estimates g_{out} measured by SNAPG on images with well resolved (left) and barely resolved (right) galaxies of $S/N \approx 100$ compared to the input shears g_{true} of the images. Each datapoint is the shear estimated from the 10000 isolated galaxies on an image. The text in the figures shows the multiplicative bias m and the additive bias c of the measurement obtained from the best linear fit shown in black. The potential of SNAPG as a shear measurement method is clear as the true shears can be recovered to an accuracy of less than one part per thousand.

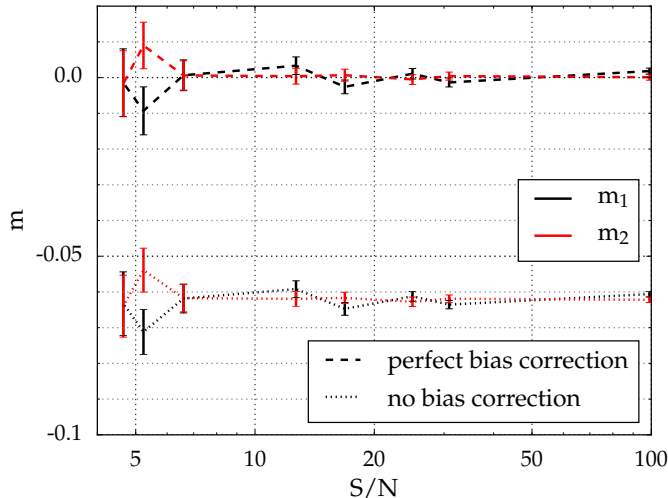


Figure 2.4: Galaxies in the well resolved image set are assigned random centroid errors, then Gaussian noise is added, after which the shear is measured with SNAPG. Measured multiplicative bias in the shear is plotted versus the mean S/N of all galaxies on the images, with m_1 in black and m_2 in red. Each datapoint is based on the shear estimates of all 2.89 million galaxies in the well resolved image set. SNAPG results with the exact centroid covariance matrix are shown as dashed lines and results without centroid error correction as dotted lines. The bias from misplaced centroids can be reliably removed by SNAPG, regardless of the noise added to the images.

images. The results of these tests are presented in Fig. 2.4 as the multiplicative bias m_1 in black and m_2 in red versus the mean S/N of all galaxies. The dotted lines show the measured bias when no correction is applied. The dashed lines show the bias when the covariance matrix is set to the correct centroid covariance $\mathbf{C} = 0.5^2 \mathbf{1}$. As we expected, the application of the correction reduces the multiplicative bias from centroid errors to subpercent levels regardless of the noise on the galaxy images. Higher levels of noise increase the variance of the measurements but do not lead to a bias. The mean corrected multiplicative bias over all S/N is $\langle m \rangle = -4.7 \cdot 10^{-4}$ and the measured additive bias is below 10^{-3} for all S/N .

2.4.3 Full test of SNAPG

We now add a centroid measurement algorithm to SNAPG and use it as input for SNAPG. The centroids of each galaxy are estimated by nulling the first moments of the galaxy and the centroid error covariance matrix is estimated using Eq. 2.25. Gaussian noise is added to the well resolved images and SNAPG is run to obtain a shear estimate. In contrast to the previous test, the noise is now directly related to the centroid error. Again we repeat this exercise for different noise levels and show the multiplicative bias as a function of the measured mean S/N in Fig. 2.5. The dotted line shows the bias in g_1 in black and g_2 in red without applying the centroid bias correction and the solid line shows the corrected bias. Centroid errors lead to a bias of several percent for $S/N \approx 10$, which is decreased to $\sim 1\%$ by the centroid bias correction. The measured additive bias is below 0.1% for all S/N . For very low S/N the estimate for the covariance matrix becomes dominated by noise and the formalism

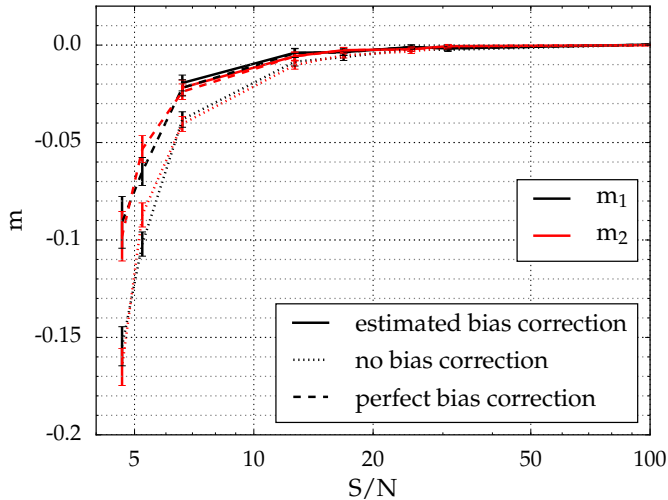


Figure 2.5: Similar set-up to that in Fig. 2.4. Dotted lines show the SNAPG results without using the centroid error correction and solid lines show the results for covariance matrices measured from the data. The centroid bias algorithm breaks down for very low S/N , so that the solid curves do not reach to $S/N < 6$. Dashed lines show the SNAPG results when the mean centroid variance in each image is taken to be the covariance matrix for that image. As the dashed and solid lines are almost indistinguishable, the measured covariance matrix is a good estimate of the true centroid covariance. However, it is unable to remove all bias caused by noise, leaving residual biases of percent level for very faint $S/N < 10$ galaxies.

breaks down.

The centroid error correction removes a large part of the noise bias, but does not remove the bias completely. This can occur if the measured covariance matrix is not a good representation of the true centroid variance. To check this hypothesis, we compute the true centroid variance and compare its performance to our previous results. First we measure the centroid from a noisy image and compare this to the true centroid to compute the centroid error $\Delta \mathbf{x}$. We then estimate the centroid variance as the average $C_{ij} = \langle \Delta x_i \Delta x_j \rangle$ over all 10000 galaxies in each image, square the centroid error, and compute the centroid variance as the squared centroid error averaged over all 10000 galaxies in an image $C_{ij} = \langle \Delta x_i \Delta x_j \rangle$. This true centroid error covariance is set into Eq. 2.21 as the covariance matrix for all 10000 galaxies in the image and the shear for the image is measured. We show these results as dashed lines in Fig. 2.5 and note that they are very similar to our previous results (solid lines). These findings indicate that the estimate of the centroid variance is good, but that there is unresolved noise bias in SNAPG.

We have also analysed the set of barely resolved images and plotted the multiplicative bias as a function of S/N in Fig. 2.6. These results sketch a similar picture: noise bias can be reduced by roughly half, but not completely removed. However, even with residual noise bias, SNAPG has only percent level biases for very faint, very small objects. This achievement is remarkable and highlights the potential of SNAPG.

We have corrected for the non-linearity due to the noisy estimates of the centroid and the second moments themselves have a linear relationship to the noise. Indeed it is shown in Fig. 2.4 that if the centroid bias is perfectly corrected for, the noise in

the second moments does not introduce a bias. However, an additional non-linearity in SNAPG remains: the correlation between the centroid error and the pixel noise. This correlation is not present in Fig. 2.4, but it is in Figs. 2.5 and 2.6. We now check whether this correlation is the origin of the residual bias after correction, by measuring the centroid and its variance from a different noise realisation than the second moments. We repeat this exercise again for different noise levels and show the results in Fig. 2.7. Even without the use of the correction (dotted lines) the bias is significantly decreased when compared to Fig. 2.5, highlighting the bias induced by the correlation. We see that the centroid bias can be corrected to subpercent levels even for galaxies with $S/N < 10$. Although the correction breaks down for galaxies that are too faint, the bias is consistent with zero down to the lowest S/N .

2.5 GREAT08

In the previous section we have shown that the shear estimated using SNAPG is accurate up to the percent level for images with circular Gaussian PSFs. Noise bias can be removed using the centroid correction described in Sec. 2.2.4 if different noise realisations of the same galaxy are available. However, it remains to be shown how SNAPG would perform on real observations and specifically on images without a circular Gaussian PSF. Therefore, we now apply the SNAPG algorithm to the more realistic Gaussianised images of the GREAT08 competition (see Sec. 2.3). These tests will show how well the PSF Gaussianisation algorithm performs. As before we use $a = 3$ and $w = 3$ for the sizes of the weight functions. For the covariance matrix, $N(\mathbf{x} - \mathbf{x}')$ is given by the auto-correlation function of the kernel which is used to Gaussianise the PSF in the GREAT08 images. This is convolved with the original covariance matrix $\sigma_r^2 \delta(\mathbf{x} - \mathbf{x}')$ and this convolution is used in Eq. 2.24 to compute the centroid covariance matrix.

We measure the shear with SNAPG on the 15 Gaussianised LNK images and the results are shown in the left panel of Fig. 2.8. There is a slight overestimation of the multiplicative bias of 1-2%, and there is a small additive bias inconsistent with zero $c_1 = (2.4 \pm 0.4)10^{-4}$ and $c_2 = (0.7 \pm 0.5)10^{-4}$. The sign of the multiplicative bias and the non-zero additive bias point towards a PSF, which is not a circular Gaussian. The results of SNAPG measurements on the 300 RNK images are shown in the right panel of Fig. 2.8. Again there is positive residual multiplicative bias $m = (+7.5 \pm 2.5)10^{-3}$, but a slightly lower value than the one we found for the LNK images. This is probably the combination of the Gaussianisation process and the imperfect centroid bias correction we found in the previous section. As we do not possess different noise realisations of the GREAT08 images, we cannot remove the residual noise bias. There is also a small, but statistically significant discrepancy between the bias in g_1 and g_2 which is not seen in other tests.

We investigate the percent level bias found in the GREAT08 in more detail by looking at the shear bias for various PSF profiles. We simulated two images of galaxies of opposite shears ($g_1 = \pm 0.03, g_2 = \mp 0.02$) with a non-Gaussian PSF. Six different Moffat profiles with $\beta = 2, 3, 4$, either circular or elliptical, with $\epsilon = +0.02, \epsilon = -0.01$ were used as PSFs. The PSF half light radius was 1.76 pixels and the galaxy half light radii were 2.5 pixels, so that these images resembled the *well resolved* images. At $S/N \sim 100$ the images underwent PSF Gaussianisation and afterwards the shear was estimated. We found that regardless of the original PSF, the bias in the shear is $\sim 2\%$, similar to the results from the GREAT08 images. At such a high S/N this bias is not

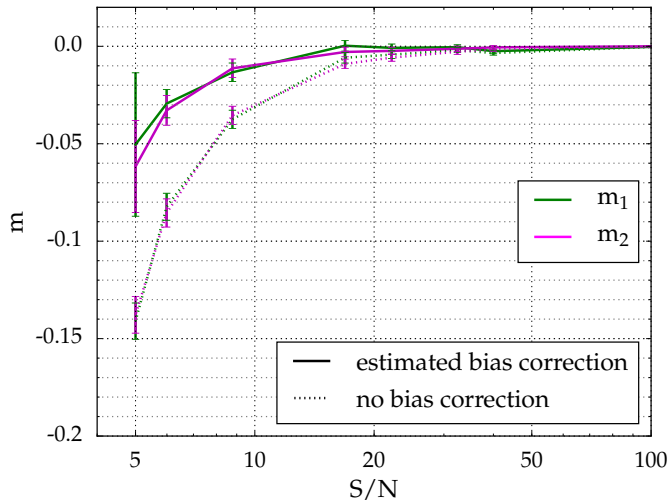


Figure 2.6: Same as Fig. 2.5, but now for the set of barely resolved galaxies. Again SNAPG shows that the PSF can be reliably accounted for, even for low S/N galaxies, as residual biases after noise bias correction are only several percentage points.

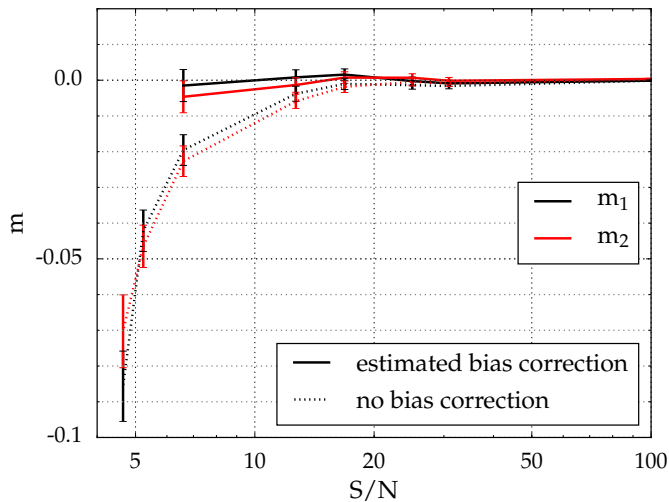


Figure 2.7: Same as Fig. 2.5 where the correlation between centroid error and the image noise is removed. Note the different y-axis scale compared to Fig.2.5. The centroid bias correction now accounts for all the bias to subpercent accuracy, until it fails for extremely low S/N .

due to a centroid error and therefore we suspect an imperfect Gaussianisation of the PSF to be the cause. It is unclear which aspect of the PSF Gaussianisation routine causes the bias in the shear estimate, although it does seem to be robust against variations in the PSF profile.

2.6 Discussion

2.6.1 SNAPG formalism

We have introduced the SNAPG formalism and tested its performance as a shear measurement method. For galaxy images convolved with a round Gaussian PSF the effect of shear on weighted second moments of image brightness can be analytically calculated. This analytical treatment is used to create a pipeline which finds the gravitational lensing shear by nulling the polarisations for an ensemble of galaxies. This procedure thus finds an estimate for the shear experienced by the galaxies. On test images with high S/N galaxies convolved with a circular Gaussian PSF, the method obtained shear estimates deviating from the input shears by only parts per thousand.

2.6.2 Noise bias

Like most shape measurement methods, SNAPG suffers a noise bias when applied to images of galaxies with low S/N . However, by using only linear combinations of second moments instead of ratios of moments such as the polarisation or ellipticity, much of the noise bias can be avoided. This strategy allows SNAPG to obtain only a percent level bias in images with a $S/N \approx 10$. Noise in the data introduces errors in the centroid estimates, which in turn biases the shear estimates. We compute an analytic treatment to correct the centroids and show that it can significantly improve the performance of SNAPG for low S/N galaxies. Remaining biases after correction for $S/N \approx 10$ are in the range of less than one percent.

The residual biases increase with decreasing S/N , which indicates that the centroid error correction does not account for the full effect of noise bias. We traced their origin to the correlation between the centroid errors and pixel noise in the second moments. By removing the correlation, we can greatly decrease the measured bias, and also correct for the remaining bias with our centroid bias correction to subpercent accuracy. For multi-band surveys a possible solution is to use different filters for the estimates of the centroid and the measurement of the moments. In this way, the correlation between the centroid and the image is removed and without this correlation SNAPG can produce almost unbiased results. The impact on the bias of such a scheme will have to be investigated as galaxy colours and colour gradients may become an issue. In addition, this introduces a correlation with the photometric redshift estimate, which might pose a problem for cosmic shear measurements.

2.6.3 Galaxy resolution

The shape of a galaxy similar in size to the PSF is heavily distorted by the PSF, making it difficult to estimate the intrinsic shape. However, the analytic treatment of the PSF in the SNAPG formalism ensures that shear estimation is possible even for barely resolved galaxies. For galaxies 0.84 times smaller than the PSF, the shear was retrieved to similar accuracy, as were resolved galaxies for $S/N \approx 7-10$ galaxies. By

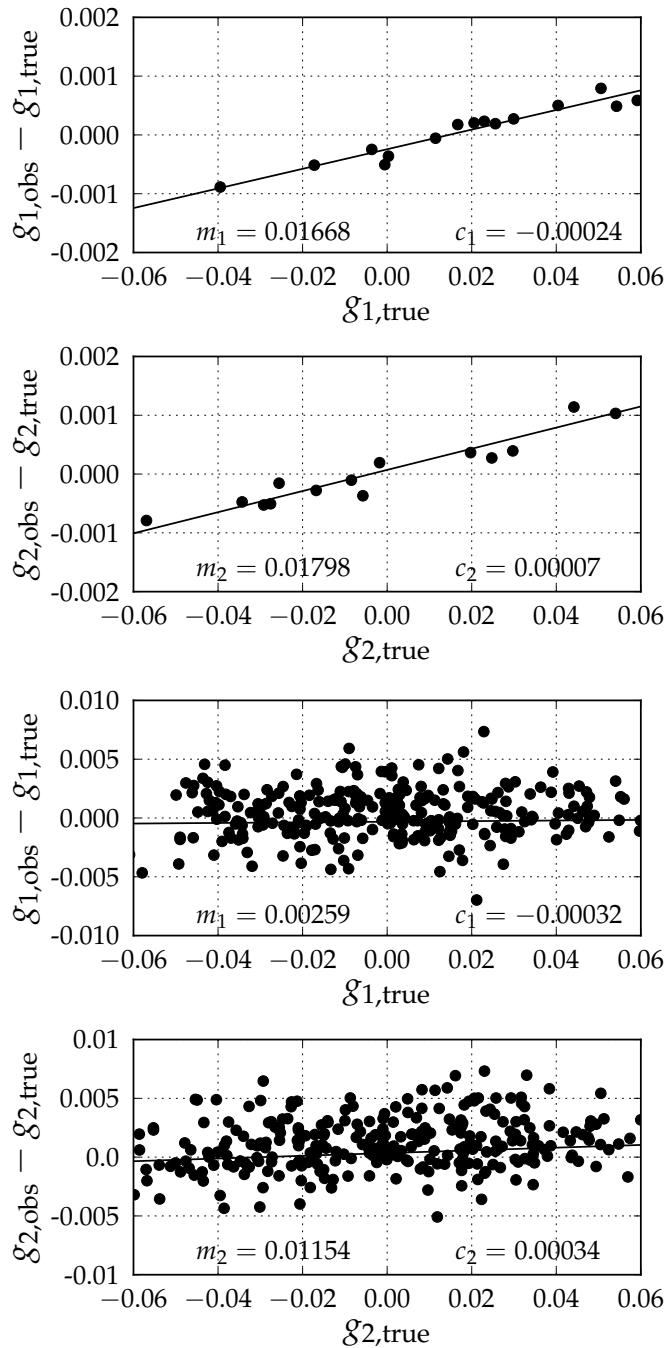


Figure 2.8: Same set-up as Fig. 2.3, here for the shear estimates g_{out} for the LowNoise_Known images of the GREAT08 challenge outfitted with a circular Gaussian PSF using SNAPG with a correction for correlated noise. The positive multiplicative bias for these high S/N galaxies shows that the PSF Gaussianisation routine did not produce a fully circular Gaussian PSF.

being able to measure unresolved galaxies reliably, SNAPG is able to use the large population of faint small galaxies to boost statistical power.

2.6.4 PSF Gaussianisation

We have run SNAPG on the images of the ‘LowNoise_Known’ and ‘RealNoise_Known’ branches of the GREAT08 challenge. To make them suitable for SNAPG, the GREAT08 images were first passed through the PSF Gaussianisation. We find a slight overestimation of the shear for the LNK images with $S/N = 200$, of the order of 1-2%. The PSF Gaussianisation introduces a correlation in the noise, which is analytically corrected for. SNAPG can retrieve the shear from the RNK images with $S/N = 20$ to an accuracy that is similar to that for the high S/N images. Further tests revealed that this percent level bias is probably inherent to the PSF Gaussianisation routine that we have used. For a variety of PSF profiles the multiplicative shear bias remained constant around 2%. The PSF Gaussianisation appears to be the limiting factor for SNAPG to obtain subpercent shear bias and detailed investigation into this routine is necessary before SNAPG can be reliably applied to observations.

We can compare the performance of SNAPG to the performance of the other methods tested in the GREAT08 challenge. This will only provide an indication as we did not run our pipeline on all datasets in the challenge and shear measurement methods have evolved since. However, a comparison to figures C3 and C4 in Bridle et al. (2010) shows that the 1-2% bias SNAPG has obtained is at least competitive with other shear measurement methods. A more quantitative comparison to other (recent) shape measurement methods will require testing on image simulations which incorporate realistic observational features. However, optimistically the performance we find for SNAPG is sufficient to meet the requirements of the largest cosmic shear survey to date (Hildebrandt et al. 2016) without any calibration being required.

2.6.5 Shear precision

So far we have been concerned only with the accuracy of SNAPG, but an equally valid demand is high precision. To estimate the scatter in the shear estimate we use the simulated images of galaxies observed with the Hubble Space Telescope (HST) included in the GalSim software. These galaxies were observed as part of the COSMOS survey (Koekemoer et al. 2007) and we used galaxies between magnitudes 20 and 24.5, similar to the depths of the Kilo Degree Survey and the Dark Energy Survey. These galaxies were rescaled to a pixel size of 0.214 arcseconds and convolved with a circular Gaussian PSF. We find that the scatter in the shear estimate for this set of galaxies is roughly $0.45/\sqrt{N_{gal}}$, where N_{gal} is the number of galaxies in the image. Thus the scatter in the SNAPG shear estimate for a fully realistic ensemble of galaxies is worse than an ellipticity based estimate; roughly 2-3 times more galaxies are needed by SNAPG for the same precision. This result is more optimistic than the increase by a factor of 10 found by Viola et al. (2014) in their analysis of a shear estimator based on Stokes parameters. Our use of a weight function reduces the variation of the moments, thereby shrinking the scatter in the Stokes parameter. In our tests we used identical weights for all sources, which naturally downweights large, bright galaxies, which would otherwise dominate the ensemble average of second moments. Ideally, in order to optimise the S/N of the individual moment measurement, the size of the weight function should match the observed size of the galaxy. However, fitting weight functions to individual galaxies is in itself a noisy process that may lead to a bias. We therefore advocate

using the same weight function size for all galaxies (since most will be only partially resolved, it is not difficult to find a size that is nearly optimal for most of the galaxies by picking a small multiple of the PSF size; see also Eq. 2.18).

A possible improvement is to assign each galaxy a weight to reduce the variance in the shear estimate. We find that for our sample of HST galaxies weighting by the inverse of the true flux can reduce the required number of galaxies by a factor of ~ 4 . This would bring the precision of SNAPG close to the precision of shear estimates based on galaxy ellipticities. In practice, estimating this weight factor from the galaxy fluxes measured in other images (e.g. adjacent photometric bands in a multi-band survey) uncorrelated with the lensing images will avoid introducing noise bias.

2.6.6 Variable shear

Observational weak lensing deals with varying shear fields, for instance in cosmic shear measurements or when measuring the mass of groups or clusters of galaxies. The traditional method is then to average the shear estimate for individual galaxies to obtain the lensing signal. This is not possible with SNAPG as it does not produce a shear estimate per source. In addition, SNAPG requires a large number of galaxies to obtain a precise shear estimate and satisfy the condition that the intrinsic ellipticities average to zero.

Instead of nulling a single shear value for an ensemble of galaxies, we therefore advocate nulling a parametrised model shear field for that ensemble. For example, to measure a galaxy-galaxy lensing signal, the model would include parameters that describe the average shear profile of galaxies and their scaling with pertinent galaxy properties. The model parameters would then be varied until the average shear in a number of annular bins around the lensing galaxies is nulled, analogous to a traditional tangential shear stacking analysis. As another example, for cosmic shear measurements, the amplitudes of independent Fourier modes in the shear field could be nulled.

Developing this procedure will be left to the future.

2.7 Summary

We have presented a new moment-based method that attempts to combine the best aspects of earlier approaches to the problem of high-accuracy, precise shear measurement from galaxy images. Moment-based methods generally approximate the deconvolution of the PSF, but do not require any information beyond the data and generally run very fast. Model fitting methods perform exact forward modelling, including convolution with the PSF, but are expensive to run because they need to search through a large parameter space, and may suffer model bias. The shear nulling after PSF Gaussianisation or SNAPG technique deals analytically with the PSF deconvolution and as a moment-based method only requires a few measurements on the data. In addition, SNAPG incorporates a correction scheme to mitigate the effects of noise bias, a major hurdle to all shape measurement techniques.

Idealised test images show that SNAPG can retrieve shears to percent level accuracy for galaxies with low signal-to-noise, even if they are smaller in size than the PSF. The main issue limiting this technique is the correlation between the noisy estimate of the centroid and the pixel noise, which may be mitigated by incorporating further data about the sources, such as images from neighbouring bands in a multi-wavelength

survey. In such a set-up, SNAPG can obtain shear estimates to subpercent accuracy for galaxies with a Gaussian PSF.

Application to real data requires PSF Gaussianisation and if this routine is imperfect it can introduce percent level biases. This level of accuracy is comparable to what is required of the shape measurement algorithms used for ongoing surveys. As such, we expect SNAPG to be a useful asset for current and future weak lensing experiments.

2.7.1 Acknowledgements

We would like to thank Massimo Viola, Henk Hoekstra, and Peter Schneider for useful discussions. We also thank the anonymous referee, whose suggestions helped to improve this paper. RH acknowledges support from the European Research Council FP7 grant number 279396. AB was supported for this research partly through a stipend from the International Max Planck Research School (IMPRS) for Astronomy and Astrophysics at the Universities of Bonn and Cologne and through funding from the Transregional Collaborative Research Centre ‘The dark Universe’ (TR 33) of the DFG. KK acknowledges support from an Alexander von Humboldt Foundation research award.

Bibliography

- Bacon, D. J., Refregier, A. R., & Ellis, R. S. 2000, *MNRAS*, 318, 625
- Bartelmann, M. & Schneider, P. 2001, *PhR*, 340, 291
- Bernstein, G. M. 2010, *MNRAS*, 406, 2793
- Bernstein, G. M. & Armstrong, R. 2014, *MNRAS*, 438, 1880
- Bernstein, G. M., Armstrong, R., Krawiec, C., & March, M. C. 2016, *MNRAS*, 459, 4467
- Bernstein, G. M. & Jarvis, M. 2002, *AJ*, 123, 583
- Bertin, E. & Arnouts, S. 1996, *A&AS*, 117, 393
- Bridle, S., Balan, S. T., Bethge, M., et al. 2010, *MNRAS*, 405, 2044
- de Jong, J. T. A., Verdoes Kleijn, G. A., Kuijken, K. H., & Valentijn, E. A. 2013, *Experimental Astronomy*, 35, 25
- Heymans, C., Van Waerbeke, L., Bacon, D., et al. 2006, *MNRAS*, 368, 1323
- Hildebrandt, H., Viola, M., Heymans, C., et al. 2016, *ArXiv e-prints* [[arXiv:1606.05338](https://arxiv.org/abs/1606.05338)]
- Hirata, C. & Seljak, U. 2003, *MNRAS*, 343, 459
- Hoekstra, H., Herbonnet, R., Muzzin, A., et al. 2015, *MNRAS*, 449, 685
- Ivezic, Z., Tyson, J. A., Abel, B., et al. 2008, *ArXiv e-prints* [[arXiv:0805.2366](https://arxiv.org/abs/0805.2366)]
- Jarvis, M., Sheldon, E., Zuntz, J., et al. 2016, *MNRAS*[[arXiv:1507.05603](https://arxiv.org/abs/1507.05603)]
- Kaiser, N. 1992, *ApJ*, 388, 272

- Kaiser, N., Squires, G., & Broadhurst, T. 1995, *ApJ*, 449, 460
- Kaiser, N., Wilson, G., & Luppino, G. A. 2000, arXiv:astro-ph/0003338 [[astro-ph/0003338](#)]
- Kilbinger, M. 2015, *Reports on Progress in Physics*, 78, 086901
- Kitching, T. D., Balan, S. T., Bridle, S., et al. 2012, *MNRAS*, 423, 3163
- Koekemoer, A. M., Aussel, H., Calzetti, D., et al. 2007, *ApJS*, 172, 196
- Kuijken, K. 1999, *A&A*, 352, 355
- Kuijken, K. 2006, *A&Ap*, 456, 827
- Kuijken, K., Heymans, C., Hildebrandt, H., et al. 2015, *MNRAS*, 454, 3500
- Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, *ArXiv e-prints* [[arXiv:1110.3193](#)]
- Luppino, G. A. & Kaiser, N. 1997, *ApJ*, 475, 20
- Mandelbaum, R., Rowe, B., Armstrong, R., et al. 2015, *MNRAS*, 450, 2963
- Mandelbaum, R., Slosar, A., Baldauf, T., et al. 2013, *MNRAS*, 432, 1544
- Massey, R., Heymans, C., Bergé, J., et al. 2007, *MNRAS*, 376, 13
- Melchior, P. & Viola, M. 2012, *MNRAS*, 424, 2757
- Melchior, P., Viola, M., Schäfer, B. M., & Bartelmann, M. 2011, *MNRAS*, 412, 1552
- Miller, L., Heymans, C., Kitching, T. D., et al. 2013, *MNRAS*, 429, 2858
- Miyazaki, S., Komiyama, Y., Nakaya, H., et al. 2012, in *SPIE*, Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV, 84460Z
- Okura, Y. & Futamase, T. 2015, *ArXiv e-prints* [[arXiv:1506.09156](#)]
- Okura, Y. & Futamase, T. 2016, *ArXiv e-prints* [[arXiv:1606.07837](#)]
- Refregier, A. 2003, *MNRAS*, 338, 35
- Refregier, A. & Bacon, D. 2003, *MNRAS*, 338, 48
- Refregier, A., Kacprzak, T., Amara, A., Bridle, S., & Rowe, B. 2012, *MNRAS*, 425, 1951
- Rhodes, J., Refregier, A., & Groth, E. J. 2000, *ApJ*, 536, 79
- Rowe, B. T. P., Jarvis, M., Mandelbaum, R., et al. 2015, *Astronomy and Computing*, 10, 121
- Spergel, D., Gehrels, N., Baltay, C., et al. 2015, *ArXiv e-prints* [[arXiv:1503.03757](#)]
- The Dark Energy Survey Collaboration. 2005, *ArXiv Astrophysics e-prints* [[astro-ph/0510346](#)]
- Tyson, J. A., Valdes, F., & Wenk, R. A. 1990, *ApJL*, 349, L1

Van Waerbeke, L., Mellier, Y., Radovich, M., et al. 2001, *A&A*, 374, 757

Viola, M., Kitching, T. D., & Joachimi, B. 2014, *MNRAS*, 439, 1909

Voigt, L. M. & Bridle, S. L. 2010, *MNRAS*, 404, 458

Wittman, D. M., Tyson, J. A., Kirkman, D., Dell'Antonio, I., & Bernstein, G. 2000, *Nature*, 405, 143

Zhang, J. & Komatsu, E. 2011, *MNRAS*, 414, 1047

Zuntz, J., Kacprzak, T., Voigt, L., et al. 2013, *MNRAS*, 434, 1604

2.A Convolution calculations

In this Appendix we calculate the result of convolving $[x_i x_j G_1]$ with a Gaussian point spread function G_2 , where G_1 and G_2 are Gaussians of arbitrary covariance matrix.

First we consider the product of a non-circular Gaussian of covariance matrix \mathbf{V} with an offset one of covariance \mathbf{P} and centre \mathbf{y} :

$$e^{-\frac{1}{2}(\mathbf{x}^T \mathbf{V}^{-1} \mathbf{x})} e^{-\frac{1}{2}(\mathbf{x}-\mathbf{y})^T \mathbf{P}^{-1}(\mathbf{x}-\mathbf{y})}. \quad (2.29)$$

The sum $\mathbf{x}^T \mathbf{V}^{-1} \mathbf{x} + (\mathbf{x} - \mathbf{y})^T \mathbf{P}^{-1}(\mathbf{x} - \mathbf{y})$ can be rearranged to yield

$$(\mathbf{x} - \mathbf{z})^T \mathbf{K}(\mathbf{x} - \mathbf{z}) - \mathbf{z}^T \mathbf{K} \mathbf{z} + \mathbf{y}^T \mathbf{P}^{-1} \mathbf{y} \quad (2.30)$$

with

$$\mathbf{K} = \mathbf{V}^{-1} + \mathbf{P}^{-1} \quad (2.31)$$

and

$$\begin{aligned} \mathbf{z} &= \mathbf{K}^{-1} \mathbf{P}^{-1} \mathbf{y} \\ &= (\mathbf{V}^{-1} + \mathbf{P}^{-1})^{-1} \mathbf{P}^{-1} \mathbf{y} = \mathbf{V}(\mathbf{V} + \mathbf{P})^{-1} \mathbf{y}. \end{aligned} \quad (2.32)$$

The terms in Eq. 2.30 not involving \mathbf{x} simplify to

$$\begin{aligned} -\mathbf{z}^T \mathbf{K} \mathbf{z} + \mathbf{y}^T \mathbf{P}^{-1} \mathbf{y} &= \\ &= -\mathbf{y}^T \mathbf{P}^{-1} (\mathbf{V}^{-1} + \mathbf{P}^{-1})^{-1} \mathbf{P}^{-1} \mathbf{y} \\ &+ \mathbf{y}^T \mathbf{P}^{-1} \mathbf{y} \\ &= \mathbf{y}^T (\mathbf{V} + \mathbf{P})^{-1} \mathbf{y}. \end{aligned} \quad (2.33)$$

Using this result we can calculate the convolution of $x_i x_j e^{-\frac{1}{2} \mathbf{x}^T \mathbf{V}^{-1} \mathbf{x}}$ with a normalised

Gaussian of covariance \mathbf{P} as

$$\begin{aligned}
& \int d\mathbf{x} x_i x_j e^{-\frac{1}{2}(\mathbf{x}^T \mathbf{V}^{-1} \mathbf{x})} \frac{e^{-\frac{1}{2}(\mathbf{x}-\mathbf{y})^T \mathbf{P}^{-1}(\mathbf{x}-\mathbf{y})}}{2\pi |\det \mathbf{P}|^{\frac{1}{2}}} \\
&= \int d\mathbf{x} \frac{x_i x_j}{2\pi |\det \mathbf{P}|^{\frac{1}{2}}} e^{-\frac{1}{2}(\mathbf{x}-\mathbf{z})^T \mathbf{K}(\mathbf{x}-\mathbf{z})} e^{-\frac{1}{2}\mathbf{y}^T (\mathbf{V}+\mathbf{P})^{-1} \mathbf{y}} \\
&= \left| \frac{\det \mathbf{K}^{-1}}{\det \mathbf{P}} \right|^{\frac{1}{2}} e^{-\frac{1}{2}\mathbf{y}^T (\mathbf{V}+\mathbf{P})^{-1} \mathbf{y}} \left(K_{ij}^{-1} + z_i z_j \right) \\
&= \left| \frac{\det \mathbf{V}}{\det (\mathbf{V} + \mathbf{P})} \right|^{\frac{1}{2}} e^{-\frac{1}{2}\mathbf{y}^T (\mathbf{V}+\mathbf{P})^{-1} \mathbf{y}} \times \\
&\quad \left[(\mathbf{V}(\mathbf{V} + \mathbf{P})^{-1} \mathbf{P})_{ij} + (\mathbf{V}(\mathbf{V} + \mathbf{P})^{-1} \mathbf{y})_i (\mathbf{V}(\mathbf{V} + \mathbf{P})^{-1} \mathbf{y})_j \right] \\
&= \left| \frac{\det \mathbf{V}}{\det \mathbf{B}} \right|^{\frac{1}{2}} e^{-\frac{1}{2}\mathbf{y}^T \mathbf{B}^{-1} \mathbf{y}} \\
&\quad \times \left[(\mathbf{V} - \mathbf{V} \mathbf{B}^{-1} \mathbf{V})_{ij} + (\mathbf{V} \mathbf{B}^{-1} \mathbf{y})_i (\mathbf{V} \mathbf{B}^{-1} \mathbf{y})_j \right]
\end{aligned} \tag{2.34}$$

where in the last line we have defined $\mathbf{B} = \mathbf{V} + \mathbf{P}$. We note that a deconvolution is simply accomplished by changing the sign of \mathbf{P} .

Calibration of weak-lensing shear in the Kilo-Degree Survey

We describe and test the pipeline used to measure the weak lensing shear signal from the Kilo Degree Survey (KiDS). It includes a novel method of ‘self-calibration’ that partially corrects for the effect of noise bias. We also discuss the ‘weight bias’ that may arise in optimally-weighted measurements, and present a scheme to mitigate that bias. To study the residual biases arising from both galaxy selection and shear measurement, and to derive an empirical correction to reduce the shear biases to $\lesssim 1\%$, we create a suite of simulated images whose properties are close to those of the KiDS survey observations. We find that the use of ‘self-calibration’ reduces the additive and multiplicative shear biases significantly, although further correction via a calibration scheme is required, which also corrects for a dependence of the bias on galaxy properties. We find that the calibration relation itself is biased by the use of noisy, measured galaxy properties, which may limit the final accuracy that can be achieved. We assess the accuracy of the calibration in the tomographic bins used for the KiDS cosmic shear analysis, testing in particular the effect of possible variations in the uncertain distributions of galaxy size, magnitude and ellipticity, and conclude that the calibration procedure is accurate at the level of multiplicative bias $\lesssim 1\%$ required for the KiDS cosmic shear analysis.

3.1 Introduction

The matter distribution in the Universe changes the geometry of spacetime, thus altering the paths of light rays. As this mimics the effects of a lens, with the gravitational potential taking the role of the index of refraction, this phenomenon is referred to as gravitational lensing. If the deflector is massive and the light rays pass sufficiently close, multiple images of the same source may be observed. More typically the source position only appears shifted by an unknown amount. The variation in the deflection across the image results, however, in a stretching (shear) and changes the observed size (magnification). This regime is commonly referred to as weak gravitational lensing (see e.g. Bartelmann & Schneider 2001, for an extensive introduction).

The original source properties are unknown, and thus the measurement of a single galaxy does not provide meaningful information. However, sources that are close on the sky have experienced similar deflections and consequently their observed orientations are correlated. The changes in the shapes of the observed galaxies are small, typically at the level of a few percent, much smaller than their intrinsic shapes. Hence, the weak lensing signal can only be determined statistically by averaging the shapes of many sources, under the assumption that there are no intrinsic correlations (but see e.g., Joachimi et al. 2015, for a review on intrinsic alignments).

The ellipticity correlations can be related directly to the statistics of matter density fluctuations (e.g. Blandford et al. 1991; Miralda-Escude 1991; Kaiser 1992) and can thus be used to infer the cosmological model. This application, commonly known as cosmic shear, is one of the most powerful ways to study the nature of dark energy and constrain modified gravity theories (see Kilbinger 2015, for a recent review). Since the first detections in 2000 (Bacon et al. 2000; Kaiser et al. 2000; Van Waerbeke et al. 2000; Wittman et al. 2000) the precision of the measurements has improved dramatically thanks to deep imaging surveys of ever larger areas (e.g. Hoekstra et al. 2006; Fu et al. 2008). Moreover, observations in multiple pass-bands allowed for the determination of photometric redshifts, which are essential to improve constraints on cosmological parameters (Schrabback et al. 2010; Heymans et al. 2013; Jee et al. 2015). The measurement of cosmic shear is also a major science driver for a number of ongoing large imaging surveys, such as the Kilo Degree Survey (KiDS; de Jong et al. 2015; Kuijken et al. 2015), the Dark Energy Survey (DES; Becker et al. 2015; Jarvis et al. 2015) and the Hyper-Suprime Cam Survey ¹.

The increase in precision afforded by these surveys needs to be matched by a corresponding improvement in the accuracy with which galaxy shapes can be measured. The main complications are (i) that the true galaxy image is convolved with a point spread function (PSF) due to atmospheric effects and telescope optics; (ii) the resulting image is pixelised by the detector; (iii) the images contain noise from various sources. Each effect introduces systematic changes in the galaxy shapes, or affects our ability to correct for it. Although shape measurement algorithms differ in their sensitivity to some of the systematics, because of differences in their implementation or the assumptions that are made, they are all affected by noise in the data.

Fortunately, it is well understood how the galaxy surface brightness is transformed into an image, and this process can be emulated. Creating mock images of telescope observations can thus be used to understand the impact of systematic effects and their propagation throughout the shear measurements. Moreover, by comparing the output shears to the input values the biases can be quantified. The biases themselves are

¹<http://www.naoj.org/Projects/HSC/surveyplan.html>

classified in additive and multiplicative bias. The former arises from an incomplete correction for the convolution by the (typically) anisotropic PSF, or by residual errors in the PSF model itself. The data themselves can be used to examine the presence of additive biases (see e.g. Heymans et al. 2012). Multiplicative bias, a change in the amplitude of the lensing signal, can only be reliably studied using simulated data. The Shear TEsting Programme (STEP; Heymans et al. 2006; Massey et al. 2007) represented the first community-wide effort to benchmark the performance of various weak lensing pipelines using simulated images. Although simplistic in many regards, the simulated data included some of the complexity of real data, such as blending of objects. To examine the differences between algorithms more systematically, the Gravitational LEnsing Accuracy Testing (GREAT; Bridle et al. 2010; Kitching et al. 2012; Mandelbaum et al. 2015) challenges focused on more idealised scenarios.

When applying an algorithm to actual data, evaluating the performance on realistic mock data is essential (Miller et al. 2013; Hoekstra et al. 2015). An essential step in this process is to ensure that the simulations are sufficiently realistic, such that the inferred bias is robust given the uncertainties of the input parameters. One approach is to match the observed properties of the simulated images to those of the real data by modifying the input distributions in case differences are found (e.g. Bruderer et al. 2015). Alternatively, the simulated output can be used to account for differences with the actual data by parameterising the bias as a function of observed galaxy properties. In Kuijken et al. (2015) and Jarvis et al. (2015) the shear biases for KiDS DR1/2 and DES, respectively, were corrected using a function of size and signal-to-noise ratio (hereafter SNR). Another option we explore is to re-weight the catalogue entries such that they match the observations.

In this paper we focus on *lensfit* (Miller et al. 2013), a likelihood based algorithm, which fits observed galaxy profiles with an elliptical surface brightness model that is convolved with a model of the PSF. This algorithm has been used to measure the lensing signal from CFHTLenS (Heymans et al. 2013) and RCSLenS (Hildebrandt et al. 2016a), as well as the initial release of KiDS (Kuijken et al. 2015). Like any other method, the *lensfit* measurements are biased if the SNR is low (this is commonly referred to as noise bias; e.g. Melchior & Viola 2012; Refregier et al. 2012; Miller et al. 2013). In the latest of these challenges, GREAT3 (Mandelbaum et al. 2015) an improved version of *lensfit* was introduced and tested: a new self-calibrating algorithm was added to alleviate the effect of noise bias. This improvement reduced the biases from tens of percents to a percent level. In this paper we expand on this formalism and apply the algorithm to simulated images that are designed to mimic KiDS data.

The third public data release of KiDS (KiDS-450 hereafter; Hildebrandt et al. 2016b) comprises 360.3 square degrees of unmasked area with an effective number density of 8.3 galaxies per square arcminute. Hildebrandt et al. (2016b) calculate that the required level of bias in shape measurements that can be tolerated given the precision afforded by KiDS-450 implies that the multiplicative bias needs to be determined to better than $\sim 1\%$. In spite of the fact that the performance of the self-calibrating version of *lensfit* is close to this requirement, a final adjustment is nonetheless required to reduce the bias further. Although this is only a small correction in absolute terms when compared to the improvement by self-calibration itself, we note that the actual implementation can be rather complex .

To reduce the biases in the shear determination for KiDS-450 to the required level of accuracy, we present SCHOol for KiDS, the Simulations Code for Heuristic Optimization of *lensfit* for the Kilo Degree Survey, which was used to obtain a shear

bias calibration for the latest KiDS-450 lensing catalogues obtained with a new version of *lensfit*. SCHOol was designed to carry out the following: i) testing of the newest version of the *lensfit* algorithm; ii) deriving bias calibration functions for the KiDS-450 data; iii) evaluating the robustness of the final calibration functions to the input of the calibration data. The main modifications to *lensfit* are presented in §3.2. The image simulations are described in detail in §3.3. These are used to quantify and account for the residual bias in the self-calibrating *lensfit* algorithms in §3.4. In §3.5 we examine how differences between the simulated and observed data can be accounted for using a resampling of the the simulated measurements. In §3.6.3 we examine the robustness of the results.

3.2 The shear measurement method

3.2.1 *lensfit*

The shear measurement method used in the analysis of KiDS data is *lensfit* (Miller et al. 2007; Kitching et al. 2008; Miller et al. 2013), which has also been used to measure the lensing signal from CFHTLenS (Heymans et al. 2013), RCSLenS (Hildebrandt et al. 2016a) and the initial release of KiDS (Kuijken et al. 2015). It is a likelihood based algorithm that fits observed galaxy profiles with a surface brightness model that is convolved with a model of the PSF. The PSF model is obtained from a fit to the pixel values of stars, normalised in flux, with a polynomial variation across individual CCD images and across the full field of each individual exposure. Galaxies are modelled as an exponential disk plus a bulge (Sérsic index $n = 4$) component. There are seven free parameters (flux, size, ellipticity, position and bulge-to-total flux ratio). To reduce the model complexity, the ratio of disk and bulge scale lengths is a fixed parameter and the ellipticities of the disk and bulge are set equal. The likelihood for each galaxy, as a function of these parameters, is obtained from a joint fit to each individual exposure, taking into account the local camera distortion. The measured ellipticity parameters are deduced from the likelihood-weighted mean parameter value, marginalised over the other parameters, adopting priors for their distribution. To determine the lensing signal, the ellipticities of the galaxy models are combined with a weight, which takes care of the uncertainty in the ellipticity measurement, to form an estimate of the shear from the weighted average. The complexity of the galaxy model has been designed to be sufficient to capture the dominant variation in galaxy surface brightness distributions visible in ground-based data, without unduly overfitting a model that is too complex to noisy data ($\text{SNR} \gtrsim 10$). In principle, we may be concerned that differences between the *lensfit* model and actual surface brightness distributions may introduce model bias (e.g. Kacprzak et al. 2014; Zuntz et al. 2013), however Miller et al. (2013) have argued that the possible model bias should be sub-dominant in ground-based data analyses, an argument that is supported by the performance of *lensfit* on simulated realistic galaxies in the GREAT3 challenge (Mandelbaum et al. 2015).

We investigate the possible amplitude of such model bias in Appendix 3.A and conclude that indeed the effect is expected to be small in the KiDS-450 analysis.

For the latest analysis of KiDS-450 data (Hildebrandt et al. 2016b) we use an updated version of *lensfit*, which is based largely on the methods adopted for CFHTLenS as described by Miller et al. (2013), but with some modifications and improvements to the algorithms. The most prominent changes are the self-calibration for noise bias and

the procedure to calibrate for weight bias, which are described in more detail below in §3.2.2 and §3.2.3, respectively. Moreover, the handling of neighbouring objects, and the sampling of the likelihood surface were improved.

In surveys at the depth of CFHTLenS or KiDS, it is essential to deal with contamination by closely neighbouring galaxies (or stars). The *lensfit* algorithm fits only individual galaxies, so contaminating stars or galaxies in the same postage stamp as the target galaxy are masked out during the fitting process. The masks are generated from an image segmentation and deblending algorithm, similar to that employed in SExtractor (Bertin & Arnouts 1996). However, the CFHTLenS version rejected target galaxies that were too close to its neighbours. For KiDS, a revised deblending algorithm was adopted that resulted in fewer rejections and thus a higher density of measured galaxies. The distance to the nearest neighbour was recorded in the catalogue output so that any bias as a function of neighbour distance could be identified and potentially rectified by selecting on that measure. The sampling of the likelihood surface was improved in both speed and accuracy, by first identifying the location of the maximum likelihood and only then applying the adaptive sampling strategy described by Miller et al. (2013). More accurate marginalisation over the galaxy size parameter was also implemented.

In the following analysis, the identical version of *lensfit*, with the same data handling setup, was used for the simulations as for the KiDS-450 data analysis of Hildebrandt et al. (2016b).

3.2.2 Self Calibration of Noise Bias

In common with other shear measurement methods, *lensfit* measurements of galaxy ellipticity are biased by the presence of pixel noise: even if the pixel noise is Gaussian or Poissonian in nature, the non-linear transformation to ellipticity causes a skewness of the likelihood and a bias in any single-point estimate of individual galaxy ellipticity that propagates into a bias on measured shear values in a survey (Refregier et al. 2012; Melchior & Viola 2012; Miller et al. 2013). The bias is a complex function of SNR, size, ellipticity and surface brightness distribution of the galaxies, but also depends on the point spread function (PSF) morphology. Given that we only have noisy estimates of galaxy properties, it is difficult to predict the bias with sufficient accuracy, and to date published shear surveys have used empirical methods to calibrate the bias, typically by creating simulations that match the properties of the survey, measuring the bias in the simulation as a function of observed (noisy) galaxy properties and applying a calibration relation derived from those measurements to the survey data (Miller et al. 2013; Kuijken et al. 2015; Jarvis et al. 2015; Hoekstra et al. 2015).

In the current analysis we first apply an approximate correction for noise bias that is derived from the measurements themselves, which we refer to as self-calibration. The method was first used for the “MaltaOx” submission in the GREAT3 challenge (Mandelbaum et al. 2015). When a galaxy is measured, a nominal model is obtained for that galaxy, whose parameters are obtained from a mean likelihood estimate. The idea of self-calibration is to create a simulated test galaxy with those parameters, remeasure the test galaxy using the same measurement pipeline, and measure the difference between the remeasured ellipticity and the known test model ellipticity. It is assumed that the measured difference is an estimate of the true bias in ellipticity for that galaxy, which may be subtracted from the data measurement. The estimate of a galaxy’s size is also simultaneously corrected with the ellipticity. Ideally, when the test galaxy is remeasured, we would like to add multiple realisations of pixel noise

and marginalise over the pixel noise: however such a procedure is computationally expensive, so in the current self-calibration algorithm we adopt an approximate method in which the noise-free test galaxy model is measured, but the likelihood is calculated as if noise were present. Mathematically we may represent the log likelihood of a measurement, $\log \mathcal{L}$ as

$$\begin{aligned}
\log \mathcal{L}(p) &= -\frac{1}{2}(\vec{D} - \vec{M}(p))^T C^{-1}(\vec{D} - \vec{M}(p)) \\
&= (\vec{M}_0 + \vec{N} - \vec{M}(p))^T C^{-1}(\vec{M}_0 + \vec{N} - \vec{M}(p)) \\
&= (\vec{M}_0 - \vec{M}(p))^T C^{-1}(\vec{M}_0 - \vec{M}(p)) \\
&\quad + 2(\vec{M}_0 - \vec{M}(p))^T C^{-1}\vec{N} \\
&\quad + \vec{N}^T C^{-1}\vec{N}
\end{aligned} \tag{3.1}$$

where we express the data as a vector \vec{D} , the model obtained with parameters p as $\vec{M}(p)$ and the pixel noise covariance matrix as C , and where we decompose the data into a true model \vec{M}_0 and a noise vector \vec{N} . Our self-calibration procedure corresponds to generating a test galaxy whose model \vec{M}_0 is described by the parameters measured from the data for that galaxy and where we only calculate the leading term in the likelihood, equation 3.1, for this test galaxy, ignoring terms involving \vec{N} , when estimating the bias. In the case where the noise is uncorrelated with the galaxy, corresponding to the background-limited case of a faint galaxy, the noise-model cross-term would disappear if we were to marginalise $\log \mathcal{L}$ over the noise, the final term would be a constant, and the leading term would provide a good estimate of the expected distribution. Unfortunately, when estimating the ellipticity, we are interested in the likelihood \mathcal{L} and not its logarithm, $\log \mathcal{L}$, and so ignoring the noise-model cross-term may lead to an error in the derived bias. However, we also make the approximation that the values of the model parameters measured from the data are close to the true galaxy parameters, which at low SNR may not be true. Hence our procedure can only be approximate.

However, self-calibration has the advantage that, unlike calibration from an external simulation, it does not rely on an assumed distribution of galaxy parameter values: the input model parameter values are taken from those measured on each individual galaxy in the data analysis. The method appears particularly useful in removing PSF-dependent additive bias, which is otherwise hard to mitigate using external simulations, which typically do not reproduce the PSF for each observed galaxy.

In making the self-calibration likelihood measurements, we are careful to ensure that the galaxy ellipticity and size parameters are sampled at the same values as in the data measurement for each galaxy, so that sampling variations do not cause an additional source of noise in the self-calibration. This procedure also makes self-calibration computationally fast, as the step of identifying which samples to use is not repeated.

The GREAT3 results (Mandelbaum et al. 2015) showed that the self-calibration correction does, on average, reduce the shear bias to the percent level and that the amplitude of the residual bias is almost independent of the morphology of the simulated galaxies. Importantly, the reduction in noise bias improves both the multiplicative and additive biases, and the self-calibration procedure therefore has been applied to the survey data measurements presented in Hildebrandt et al. (2016b). The residual bias, however, is still correlated with galaxy properties such as SNR and size. As the distributions of those properties are redshift- and magnitude-dependent, the residual

bias may be large enough to lead to a significant bias in tomographic shear analyses. We therefore seek to empirically calibrate the residual bias using conventional methods, employing realistic image simulations as described in §3.3.

3.2.3 Weight bias correction

In our standard analysis, we apply a weight to each galaxy that takes account of both the shape noise variance and the ellipticity measurement noise variance, following Miller et al. (2013). The ellipticity noise variance is measured from the ellipticity likelihood surface for each galaxy, after marginalisation over other parameters, with a correction for the finite support imposed by requiring ellipticity to be less than unity. This contrasts with approaches such as that of Jarvis et al. (2015), where an average correction as a function of galaxy parameters, such as flux signal-to-noise ratio, is derived and applied.

Our scheme should result in optimal SNR in the final shear measurements, but any bias in the weights would introduce a shear bias. Inspection of the distribution of weight values shows that indeed there are two sources of weight bias that arise. First, the measurement variance is a systematic function of the ellipticity of the galaxy, with a tendency for galaxies to have smaller measurement variance, and hence higher weight, at intermediate values of ellipticity, compared with either low or high ellipticity, for galaxies of comparable isophotal area and SNR. This results in a tendency to overestimate shear at intermediate and low values of SNR, to an extent that is sensitive to the distribution of galaxy ellipticities.

A second bias that arises is correlated with the PSF anisotropy. Galaxies of a given total flux that are aligned with the PSF tend to have a higher SNR than galaxies that are cross-aligned with the PSF, and also tend to have a smaller measurement variance. This orientation bias has the same origin as that discussed by Kaiser (2000) and Bernstein & Jarvis (2002) and results in a net anisotropy in the overall distribution of weights which, if uncorrected, would result in a net shear bias.

In the KiDS-450 analysis, we adopt an empirical correction for these effects by determining the mean measurement variance for the full sample of galaxies as a function of their 2D ellipticity, e_1 , e_2 , and as a function of their SNR and isophotal area. From that mean variance, a correction is derived that may be applied to the weights to ensure that, on average, the distribution of weights is neither a strong function of ellipticity nor of position angle. The anisotropic bias depends on the size and ellipticity of the PSF, so to accommodate variations in the PSF across the survey, galaxies from the entire completed survey are binned according to their PSF properties, and the weights correction is derived in each PSF bin (Hildebrandt et al. 2016b). In the simulations, we apply the equivalent weight bias correction to each of 13 sets of PSFs that are simulated (see §3.3.4).

3.3 Image simulations

3.3.1 The simulation of galaxies

The performance of shape measurement algorithms can only be evaluated using simulated images. To this end, a number of community-wide efforts have been undertaken to benchmark methods. The self-calibrating version of *lensfit* performed well on simulated images from GREAT3 (Mandelbaum et al. 2015), the latest of these challenges,

with an average shear bias of about a percent. Whilst useful to test new algorithms and to better understand common sources of bias in shape measurements, these general image simulations cannot be used to evaluate the actual performance. First of all, they ignore the effects neighbouring objects can have on the shape measurement, which was shown to be important by Hoekstra et al. (2015). Moreover, to calibrate the performance with high accuracy, the simulations should match the real data in terms of survey depth, number of exposures, noise level, telescope PSF and pixelisation.

To quantify and calibrate the shear biases of the self-calibrating version of *lensfit* for the new KiDS-450 dataset we created the SCHOoL for KiDS pipeline, Simulations Code for Heuristic Optimization of *lensfit* for the Kilo Degree Survey. We use it to generate a suite of image simulations that mimic the *r*-band KiDS observations that were used in Hildebrandt et al. (2016b) to measure the cosmic shear signal. As discussed below, we match the dither pattern, instrument footprint, average noise level, seeing and PSF properties. The simulated images are created using GALSIM (Rowe et al. 2015), a widely used galaxy simulation software tool developed for GREAT3. Note that we do not aim to test the PSF modelling (this was presented in Kuijken et al. 2015).

3.3.2 Simulation volume

The precision with which biases are measured can be improved by creating and analysing more simulated images. However, it is a waste of computational resources if the biases are already known sufficiently well compared to the statistical uncertainties of the cosmic shear signal. Moreover, as a result of simplifications in the simulated data, residual biases may remain. It is therefore useful to establish the level of accuracy that is required, given the KiDS-450 data set, and use these results to determine the simulation volume that is needed. Hildebrandt et al. (2016b) showed that the *lensfit* shear multiplicative bias has to be known with an accuracy of at least 1% for the error bars on cosmological parameters not to increase by more than 10% (see their Appendix A3). Hildebrandt et al. (2016b) do not set requirements on the knowledge of the additive bias from the simulations. In fact the residual additive bias is measured from the data themselves (Heymans et al. 2012) as there are a number of steps in the data acquisition, processing and analysis which are not simulated and might contribute to amplitude of the additive bias (e.g. cosmic rays, asteroids, binary stars, imperfect PSF modelling, non-linear response of CCD...). The observed level of residual bias may be used to determine the maximum scale where the cosmic shear signal is robust, in contrast to multiplicative shear bias, which affects all angular scales.

In our simulations we apply a shear with a modulus $|g| = 0.04$ to all galaxies. This is a compromise between the small shears we aim to recover reliably, whilst minimising the number of simulated images. For a fiducial intrinsic dispersion of ellipticities $\sigma_\epsilon = 0.25$, the minimum required number of galaxies to reach a precision of 0.01 on the multiplicative bias is then $N_{\text{gal}} = (\sigma_\epsilon / (0.01|g|))^2 \approx 3.9 \times 10^5$. This number should be considered the bare minimum, because in practice we wish to explore the amplitude of the bias as a function of galaxy and PSF properties.

The dominant source of uncertainty is the intrinsic dispersion of ellipticities. This source of noise can, however, be reduced in simulations using a shape noise cancellation scheme (Massey et al. 2007). This results in a significant reduction in the number of simulated galaxies, without affecting the precision with which the biases can be determined. Previous studies have done so by introducing a copy of each galaxy, rotated in position angle by 90° before applying a shear and convolution by the PSF, such that the mean of the intrinsic ellipticity ϵ^s satisfies $\langle \epsilon^s \rangle = 0$ (e.g. Massey et al.

2007; Hoekstra et al. 2015). Although this reduces the shape noise caused by galaxies, such a scheme does not guarantee that the mean of the *observed* ellipticity values $\langle \epsilon \rangle = g$. That condition is only satisfied by a population of galaxies that are uniformly distributed around circles of ϵ^s . Fortunately, even a small number of rotated copies of each galaxy suffices to meet this criterion to adequate accuracy.

In this work we create four copies of each galaxy, separated in intrinsic position angle by 45° . If we write the first copy as having intrinsic ellipticity ϵ^s , we may write the complex intrinsic ellipticity of each copy as $\epsilon_n^s = i^n \epsilon^s$ for each rotation, $n = 0 \dots 3$. The relation between the sheared ellipticity ϵ_n , the reduced shear g and ϵ_n^s , for each rotation, is

$$\epsilon_n = \frac{\epsilon_n^s + g}{1 + g^* \epsilon_n^s} = \frac{i^n \epsilon^s + g}{1 + g^* i^n \epsilon^s}, \quad (3.2)$$

where the asterisk denotes the complex conjugate. A shear estimate $\tilde{g} = \langle \epsilon_n \rangle$ then reduces to

$$\tilde{g} = \frac{g - g^{*3} (\epsilon^s)^4}{1 - (g^* \epsilon^s)^4}. \quad (3.3)$$

For the same fiducial values, $|\epsilon^s| \simeq 0.25$ and $|g| = 0.04$, this expression differs from g with a relative error of order $\Delta g/g \simeq |g|^2 |\epsilon^s|^4 \simeq 6 \times 10^{-6}$, compared with $\Delta g/g \simeq |\epsilon^s|^2 \simeq 0.06$ for the shape noise reduction achieved using only pairs of galaxies (Massey et al. 2007). The four-rotation method has significantly higher accuracy relative to the two-rotation method at the highest values of ϵ^s .

Using a larger number of rotated galaxies reduces the shear measurement error further, to $\Delta g/g \sim 10^{-13}$ for 8 duplicated galaxies. However, for a given simulation volume, this reduces the diversity in other galaxy properties. Moreover, pixel noise in the simulated images reduces the effectiveness of shape noise cancellation for galaxies with low SNR, which are the most numerous. Furthermore, not all rotated galaxy copies may be detected, thus breaking the assumed symmetry in the analytical estimate. The weighted dispersion of the mean input ellipticities of the set of four catalogues is 0.084, a factor about 3 reduction compared to the case without shape noise cancellation. This corresponds to a decrease of a factor about 9 in the number of simulated galaxies required to achieve a fixed uncertainty in shear bias measurement.

3.3.3 Input object catalogue

To measure meaningful shear biases from the simulated data it is essential that the properties of the simulated objects are sufficiently realistic. For instance, neighbouring galaxies affect shape measurements (Dawson et al. 2014), and therefore the correct number density of galaxies needs to be determined. Moreover, Hoekstra et al. (2015) highlighted the importance of simulating galaxies well beyond the detection limit of the survey in order to derive a robust shear calibration. Galaxies just below the detection limit can still blend with brighter galaxies, directly affecting the measurement of the object ellipticity, whereas even fainter galaxies affect the background and noise determination by acting as a source of correlated noise. Hence we include in our simulations galaxies as faint as 28th magnitude, which should be adequate given the depth of KiDS.

We place the objects at random positions, and thus ignore the additional complication from clustering. The fraction of blended objects in the simulations might

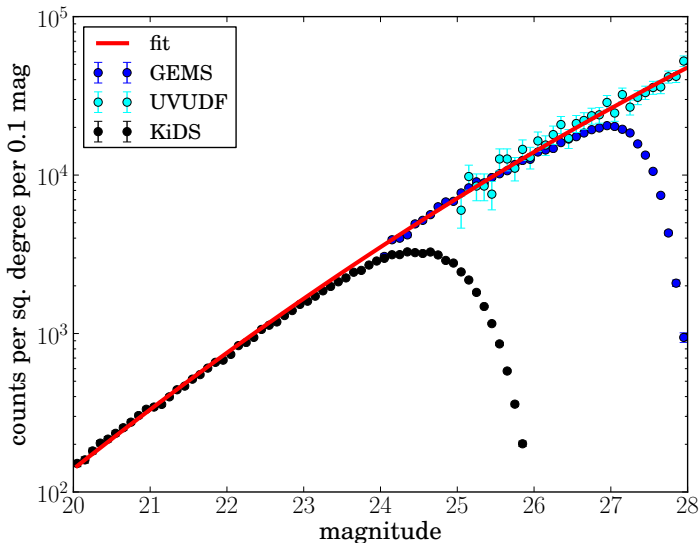


Figure 3.1: r -band magnitude histograms of KiDS-450 data (black), GEMS survey data (blue) and UVUDF survey (cyan), with uncertainties given by the Poisson errors of each point. The red line is the best fit through KiDS-450 $20 < m_r < 23$ points, GEMS $25 < m_r < 26$ points and UVUDF $26 < m_r < 29$ datapoints and is used as the input magnitude distribution of the simulations.

therefore be low compared to the true Universe. Alternatively, galaxies could be positioned in the simulations according to their positions in observations (e.g. Miller et al. 2013; Jarvis et al. 2015). This would naturally include realistic clustering, but cannot be used for the galaxies below the detection limit, and thus unusable for our deep magnitude distribution. However, we examined the impact of varying number density and found the changes in bias to be negligible for the KiDS-450 analysis (see §3.4.4 for details).

To create a realistic magnitude distribution that extends to 28th magnitude, we augment the measured KiDS-450 galaxy counts with measurements from deeper *Hubble Space Telescope* (HST) images. We use the HST/ACS $F606W$ counts from GEMS (Rix et al. 2004) and UVUDF (Rafelski et al. 2015), because this filter resembles the KiDS r filter fairly well. We remove objects classified as stars from all three data sets, and exclude masked areas in the KiDS-450 data. Fig. 3.1 shows the magnitude distributions of a subsample of KiDS-450 data (black), GEMS data (blue) and UVUDF data (cyan). The error bars show the Poisson errors of the data points.

We fit a second order polynomial to the logarithm of the number counts, using KiDS-450 data between $20 < m_r < 23$, GEMS data between $25 < m_r < 26$ and UVUDF data between $26 < m_r < 29$. The resulting magnitude distribution for the simulated galaxies is given by:

$$\log N(m_r) = -8.85 + 0.71m_r - 0.008m_r^2, \quad (3.4)$$

where $N(m_r)$ is the number of objects with r -band magnitude m_r per square degree. The fit is mostly constrained by the KiDS data, with the ancillary data driving the

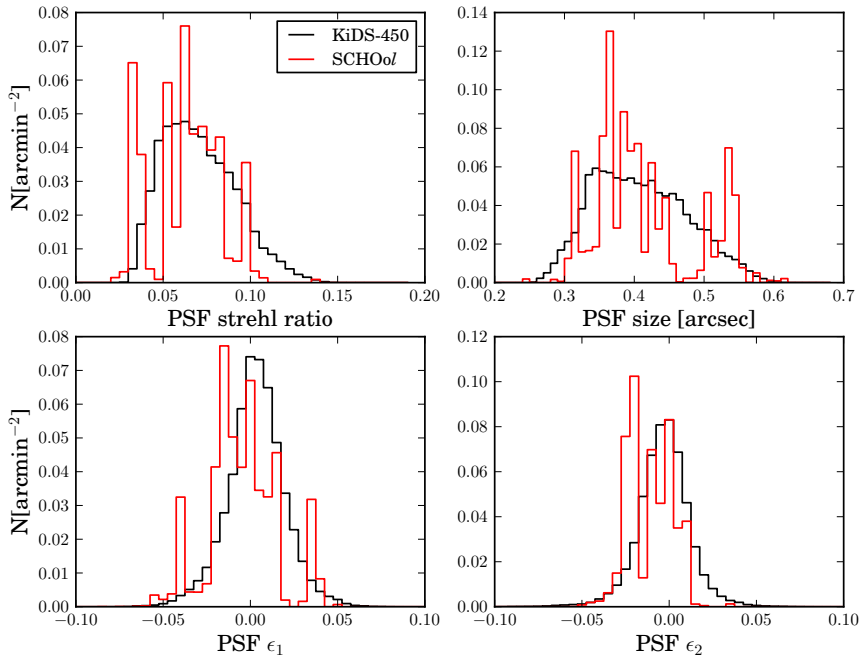


Figure 3.2: Distributions of PSF parameters in the simulations (red) and KiDS-450 (black) measured by *lensfit* using a 2.5 pixel weighting function. Shown are the distributions of measured pseudo-Strehl ratio, size and the two components of the ellipticity. The constant PSFs (for individual exposures) in the SCHOol images give rise to very peaky distributions, but overall the range in properties in the data are matched by the image simulations.

flattening of the curve at faint magnitudes. Magnitudes are converted to counts to be used by GALSIM using a magnitude zeropoint of 24.79, the median magnitude zeropoint in the KiDS-450 data.

Creating images of large numbers of faint galaxies with $m \geq 25$ by GALSIM would be rather time consuming. However, we are not interested in their individual properties, because they are too faint to enter the sample used for the lensing analysis. Instead we only need to ensure that their impact on shape measurements is captured, for which it is sufficient that their number densities and sizes are realistic. To improve the speed of the pipeline, we therefore create postage stamps for a representative sample of these faint galaxies, and use these to populate the simulations by randomly drawing from this sample, whilst ensuring that the magnitude distribution in equation 3.4 is obeyed. These faint galaxies also have lensing shear applied.

Realistic galaxy morphologies, in particular the distribution of surface brightness profiles, and consequently sizes and ellipticities, are another essential ingredient for image simulations. The intrinsic ellipticity distribution for galaxies is the same as in the CFHTLenS image simulations and the functional form is taken from Appendix B2 in Miller et al. (2013). It corresponds, as is the case for the size distribution, to the prior used by *lensfit* to measure galaxy shapes. We model the galaxies as the linear combination of a de Vaucouleur profile for the bulge and an exponential profile for the disk. The bulge flux to total flux ratio, B/T , is randomly sampled from a truncated Gaussian distribution between 0 and 1 with its maximum at 0 and a width of 0.1, the same as was used for the CFHTLenS simulations presented in Miller et al. (2013). Ten percent of all galaxies are set to be bulge-only galaxies with $B/T = 1$, and the rest have a disk with random values for the bulge fraction.

The sizes of the galaxies are defined in terms of the scale length of the exponential disk along the major axis, and are randomly drawn from the distribution

$$P(r) \propto r \exp(-(r/A)^{4/3}), \quad (3.5)$$

where A is related to the median of the distribution, r_{med} , by $A = r_{\text{med}}/1.13^2$ and where the relationship between r_{med} and magnitude is given by $r_{\text{med}} = \exp(-1.31 - 0.27(m_r - 23))$. This distribution is the same as given by Miller et al. (2013) but with the r_{med} relation shifted to be appropriate for observations in the KiDS r filter (see Kuijken et al. 2015). The distribution corresponds also to the *lensfit* prior used in the analysis of the KiDS observations. For the bulge-plus-disk galaxies simulated here, the half-light radius of the bulge component is set equal to the exponential scale length of the disk component (see Miller et al. 2013, for details). Galaxies are simulated using GALSIM, which defines the size as $r_{ab} = \sqrt{ab}$, where a and b are the semi-major and semi-minor axis of the object, respectively, so the sizes sampled from equation 3.5 were converted to r_{ab} prior to simulation.

We also include stars in the simulations, as they might contaminate the galaxy sample and blend with real galaxies (see Hoekstra et al. 2015, for a discussion of the effect of stars on shear measurements). The simulated stars are perfect representations of the PSF in the simulated exposure and we do not include realistic CCD features around bright stars, such as bleeding, stellar spikes or ghosts, as these effects are masked in the real data. The stellar r -band magnitude distribution is derived using the Besançon model³ (Robin et al. 2003; Czekaj et al. 2014) for a right ascension

²There was an error in Appendix B1 of Miller et al. (2013): the factor 1.13 shown here was also used for the CFHTLenS analysis, instead of the incorrectly reported value of 0.833.

³model.obs-besancon.fr

Table 3.1: Overview and specifications of all simulated images created with the SCHOoL pipeline

Total simulated area	416 square degrees
Tile	5 exposures of ~ 1 square degree dithered by 25 arcsec, 85 arcsec
Exposure	32 chips of $\sim 2000 \times 4000$ pixels with 70 pixel wide chip gaps in between
Applied shears	(0.0,0.04) (0.0283,0.0283) (0.04,0.0) (0.0283,-0.0283) (0.0,-0.04) (-0.0283,-0.0283) (-0.04,0.0) (-0.0283,0.0283)
Applied PSF	The same shear is applied to all galaxies in a tile 13 sets; each set contains 5 different PSF models of KiDS-450 observations Each PSF model is applied to all galaxies in an exposure
Shape noise reduction	Each tile is copied with galaxies rotated by 45, 90 and 135 degrees

$\alpha = 175^\circ$ and a declination $\delta = 0^\circ$, corresponding to one of the pointings in the KiDS-450 footprint. We note that the star density in that pointing is higher than average. This is not a concern for the bias calibration, as discussed in §3.4.4. We do not include very bright ($m_r < 20$) stars, because they would be masked in real observations and we exclude stars fainter than $m_r > 25$.

3.3.4 Simulation setup

As described in detail in de Jong et al. (2015) and Kuijken et al. (2015), *lensfit* measures galaxy shapes using the five r -band exposures that make up a tile covering roughly one square degree of the sky. The KiDS-450 data are analysed tile-by-tile, i.e. data from the overlap of tiles is ignored. It is thus sufficient to simulate individual tiles. Each VST/OmegaCam exposure is seen by a grid of 8×4 CCD chips, where each chip consists of 2040×4080 pixels that subtend $0''.214$. There are gaps of around 70 pixels between the chips and to fill the gaps the exposures are dithered. To capture the resulting variation in depth due to this dither pattern we simulate individual tiles of data, using the same dither pattern described in de Jong et al. (2015), which we incorporate by adding artificial astrometry. We also add a small random shift in pointing between the exposures, so that the same galaxy is mapped on a slightly different location in the pixel grid for each exposure. This extra shift is accounted for when stacking the exposures. Gaussian background noise is added to the simulated exposures, where the root mean square of the noise background $\sigma_{\text{bg}} = 17.03$ was determined as the median value from a sub-sample of 100 KiDS-450 tiles. When exposures are stacked, the noise level varies with position in the simulated tile as in the real data, owing to the chip gaps.

The simulated images for each exposure are created using GALSIM (Rowe et al. 2015) which renders the surface brightness profiles of stars and sheared galaxies using the input catalogues detailed in §3.3.3. The five exposures for each tile are created using the same input catalogue. The 32 individual chips in each of the five exposures are coadded using SWARP⁴ (Bertin 2010). Finally we run SExtractor (Bertin &

⁴Note that we do not use the resampling option of SWARP to reduce the processing time. This might introduce some incorrect sub-pixel matching of the pixels in the coadded image, but does not

Arnouts 1996) to detect objects in the coadded image. We use the same version of the software and configuration file as is used in the analysis of the KiDS-450 data (de Jong et al. 2015) to ensure homogeneity. Only the magnitude zeropoint is set to the value of 24.79 which was used to create the simulations.

Eight shear values are sampled isotropically from a circle of radius $|g| = 0.04$ and using evenly spaced position angles (see Table 3.1 for the exact values). We apply the same shear to each simulated galaxy in the five exposures in a simulated tile, using the GALSIM *Shear* function which preserves galaxy area, but vary the shear between tiles. The sheared galaxies are convolved with an elliptical Moffat PSF, whose parameters are representative of the ones measured in KiDS-DR1/2 (de Jong et al. 2015). To obtain the PSF parameters, we ran PSFEX (Bertin 2013) on KiDS-DR1/2 data. As the VST seeing conditions change over time, so that different exposures have different PSFs, we mimic this temporal variation of the PSF in the *SCHOOl* simulations. To this end we selected a series of PSF parameters corresponding to 5 subsequently observed dithered exposures of KiDS data. This gave us a set of Moffat parameters for the PSF in each of the 5 exposures of a tile. All galaxies in a simulated exposure were convolved with the same Moffat profile. All galaxies in the first simulated exposure thus have the PSF in the first exposure of the observed KiDS tile. The second simulated exposure has galaxies convolved with the observed PSF in the second exposure of the KiDS tile. And so on for all five exposures of the simulated tile. This ensures that the PSFs in the simulations are the same as in the KiDS observations. We used the PSF parameters from 13 KiDS tiles, so that we have in total 65 different PSFs in the simulations. This number of PSFs gave us enough statistical power to reach the required precision. The 13 tiles were chosen so that the distributions of PSF parameters in the simulations would match the distribution of the full KiDS data. The distributions of simulated PSF properties measured by *lensfit* on the *SCHOOl* images are shown in the red histograms in Fig. 3.2. We define the PSF size in terms of the weighted quadrupole moments P_{ij} of the surface brightness of the PSF:

$$r_{\text{PSF}}^2 := \sqrt{P_{20}P_{02} - P_{11}^2}, \quad (3.6)$$

where we measure the moments employing a Gaussian weighting function with a size of 2.5 pixels. The bottom panels show the two components of the weighted ϵ ellipticity. Comparison with the distributions measured in the KiDS-450 data (shown in black) shows that the simulations sample the range in PSF properties. The median full width to half maximum (FWHM) of $0''.64$ in our sample is very similar to the value of $0''.65$ from the full KiDS sample. However, the lack of spatial variation in the simulations produces very spiky distributions. This also leads to an over-representation of large and elliptical PSFs in the simulations.

In total we have simulated 416 deg^2 of KiDS observations, slightly more than the unmasked area of the KiDS-450 dataset. However, the use of shape noise reduction ensures that we have ample statistical power in the calibration, because the simulated data are equivalent to an area of $\sim 3750 \text{ deg}^2$ without the shape noise cancellation. A summary of the set of simulations created with the *SCHOOl* pipeline is provided in Table 3.1.

affect the *lensfit* measurements, which are made by jointly fitting to the original individual exposures.

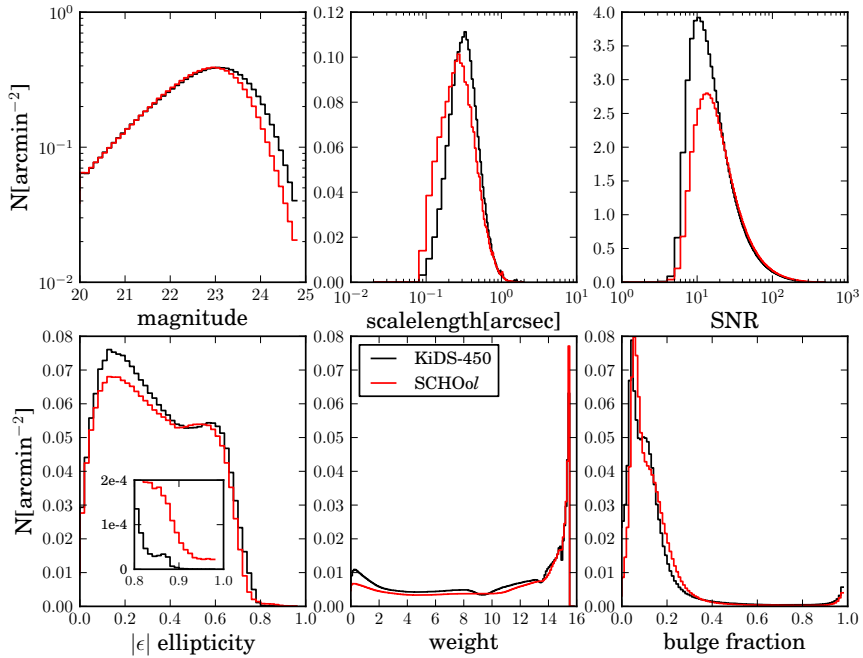


Figure 3.3: Comparison of KiDS-450 data (black) and SCHOoI simulations (red) for weighted normalised distributions of galaxy properties. From left to right, top to bottom: magnitude, size, SNR, modulus of the ellipticity $|\epsilon|$, *lensfit* weights, bulge fraction. The inset shows a zoom in of the ellipticity distributions for $\epsilon > 0.8$.

3.3.5 Comparison to data

Although our input catalogue is based on realistic prior distributions, it is important to verify whether the simulated data are a good representation of the observations. Differences with the actual KiDS-450 measurements may occur because of simplifying assumptions or errors in the prior distributions. For instance, in the simulations the PSF is constant over one square degree and the noise level does not vary. Therefore, the resulting *lensfit* measurements are not identical to those in KiDS-450 data and the average shear biases inferred from the simulations may differ from the actual shear biases in the data. Rather than adjusting the input catalogue such that the agreement with the data is improved (Bruderer et al. 2015), we instead aim to model the biases as a function of observed properties (see §3.4). This approach does not require perfect simulations, but does require that the simulations capture the variation in galaxy properties seen in the data. To examine whether this is indeed the case, we compare the measured galaxy properties in the simulations to those in the KiDS-450 data.

We run *lensfit* on the entire volume of the simulations, using the SExtractor detection catalogue as input. For each detected object *lensfit* returns a measurements of the ellipticities, weights as well as measurements of the galaxy properties such as SNR and size. A measurement of the observed magnitude is provided by SExtractor. In order for the comparison with the data to be meaningful the same cuts have to be applied to both datasets. In both cases we consider only measurements of galaxy shapes for objects fainter than $m_r = 20$. Moreover, to study selection biases (see §3.4.2) we create a catalogue that contains for each detected object its input properties and those measured by SExtractor and *lensfit*. This is done using a kD-tree based matching routine which combines each *lensfit* output catalogue with the input catalogue used to create the galaxy images.

For each object in a given *lensfit* catalogue we find its five nearest neighbours in the input catalogue, according to the L2-norm spatial separation. We discard all candidates with a separation larger than three pixels and select from the remainder the one with the smallest difference in measured magnitude and input magnitude as the final match. This last step introduces a sensible metric to discard by-chance close-neighbour pairs of physically different objects. This matching process removes spurious detections from the catalogue. This is not a problem for the bias characterisation, as *lensfit* would have assigned a vanishing weight to such spurious detections.

After the matching we apply a series of cuts to the data, starting with the removal of all objects with a vanishing *lensfit* weight to reduce the size of the analysis catalogues. This does not have any effect on the recovered shear since this is calculated as a weighted average of the measured ellipticities. This initial selection automatically removes the following:

1. Objects identified as point sources (`fitclass = 1`)
2. Objects that are unmeasurable, usually because they are too faint (`fitclass = -3`)
3. Objects whose marginalised centroid from the model fit is further from the SExtractor input centroid than the positional error tolerance set to 4 pixels (`fitclass = -7`).
4. Objects where insufficient data is found, for example an object at the edge of an image or defect (`fitclass = -1`)

Additionally, in order to match the cuts applied to the KiDS-450 data (see Appendix D in Hildebrandt et al. (2016b)), we also remove:

5. Objects with a reduced $\chi^2 > 1.4$ for their respective *lensfit* model, meaning that they are poorly fit by a bulge plus disk galaxy model (`fitclass = -4`).
6. Objects whose *lensfit* segmentation maps contain more than one catalogue object (`fitclass = -10`).⁵
7. Objects that are flagged as potentially blended, defined to have a neighbouring object with significant light extending within a contamination radius > 4.25 pixels of the SExtractor centroid.
8. Objects that have a measured size smaller than 0.5 pixels.

After these cuts, considering all image rotations, shear and PSF realisations, we obtain a sample of ~ 16 million galaxies which are used in the analysis. Fig. 3.3 shows the resulting weighted distributions of magnitude, scale length, modulus of ellipticity, bulge fraction, SNR and weight measured from KiDS-450 data (black) and the SCHOoI simulations (red).

The distributions of the *lensfit* measurement weight and bulge fraction are in good agreement with the data, although the measured bulge fractions are extremely noisy, and are eliminated from the shear measurement by a marginalisation step. However, the agreement in the simulated and observed distributions gives some reassurance that the simple parametric galaxy profiles are an adequate representation of the KiDS-450 data. The simulated galaxy counts are in good agreement with the observations for bright galaxies, but the magnitude and SNR distributions suggest that the simulations lack faint, low SNR objects. The paucity in the simulated catalogues might be attributed partly to the fixed noise level or the spatially constant PSF in the simulations, which is not fully representative of KiDS-450 observations, but also partly to a difference in intrinsic size distributions of the data and simulations, which may also be seen in Fig. 3.3.

The shear measurement bias that we seek to calibrate depends primarily on galaxy size and SNR (e.g. Miller et al. 2013), and differences in the distributions of these quantities between the data and the simulations mean that we cannot simply measure the total bias from the simulations and apply the result to the data. Furthermore, this consideration applies to the bias for any sub-selection of the data, such as the analysis of shear in tomographic bins of Hildebrandt et al. (2016b). Even if the data and simulations were a perfect match in Fig. 3.3, any dependence of bias on galaxy properties would mean that a ‘global’ bias for the simulations might not be appropriate to the galaxy selection in tomographic bins. Thus, in this paper we derive a shear calibration that includes a dependence on size and SNR, but also investigate the sensitivity of the final shear calibration to modifications of the assumed distributions, in §3.6.1 and §3.6.2.

⁵ In order to remove contamination from nearby objects, *lensfit* builds a dilated segmentation map that is used to mask out a target galaxy’s neighbours. It was found that a small fraction of targets had two input catalogue target galaxies within a single segmented region associated with the target, owing to differing deblending criteria being applied in the SExtractor catalogue generation stage from the *lensfit* image analysis. When measured, this leads to two catalogue objects being measured using the same set of pixels, and thus the inclusion of two correlated, high ellipticity values in the output. As these accounted for a very small fraction of the catalogue, these instances were flagged in the output and excluded from subsequent analysis.

The ellipticity distributions also differ, both at low and high ellipticity. Both the simulations and the KiDS-450 data contain very elliptical galaxies, as is clear from the inset in the lower left panel of Fig. 3.3, which shows the high ellipticity tail of the distribution. In the simulations these high ellipticities are caused by noise or blending with neighbours, as there are no galaxies with an intrinsic ellipticity $\epsilon > 0.804$. However, in the data this is not necessarily the case. Differences in the ellipticity distribution may lead to an incorrect estimate of the shear bias and this is especially worrying for highly elliptical objects (Melchior & Viola 2012; Viola et al. 2014). In §3.6.3 we investigate the (origin of the) discrepancy and also quantify the resulting uncertainty in shear bias that arises from the differences between the data and the simulations.

As noted above, the observed differences suggest that the simulations cannot be used directly to infer the shear biases, and in the remainder of this paper we explore calibration strategies that use observed properties to estimate the bias for a given selection of galaxies (Miller et al. 2013; Hoekstra et al. 2015). For this to work, it is important that the simulations at least cover the multi-dimensional space of relevant parameters. Moreover, differences in selection effects should be minimal. Before we explore these issues in more detail, we first examine the distributions of the two most relevant parameters, namely the SNR and the ratio of the PSF size and the galaxy size (e.g. Massey et al. 2013). The latter parameter, which we define as,

$$\mathcal{R} := \frac{r_{\text{PSF}}^2}{(r_{ab}^2 + r_{\text{PSF}}^2)}, \quad (3.7)$$

quantifies how the shape is affected by the convolution by the PSF. For the analysis, we adopt the r_{ab} size definition, because it has significantly lower correlation with the measured ellipticity in noisy data (cf. §3.4.3).

Fig. 3.4, shows the ratio between the number of simulated and real galaxies on a grid in SNR and \mathcal{R} defined using the KiDS-450 data. The size of each data point is proportional to the sum of the *lensfit* weight in each grid cell. The red stars indicate the region where the ratio is 0; i.e. the simulations do not contain objects with that SNR and resolution. The simulations are lacking very large objects (low \mathcal{R}) and with low SNR. Those objects contribute only 0.001 % of the total weight and hence the fact that they are not present in the simulations can be safely ignored.

3.4 KiDS Calibration Method

3.4.1 The evaluation of shear bias

As our image simulations are a good, but not perfect representation of the KiDS-450 data, and as in our data analyses (e.g. Hildebrandt et al. 2016b) we select sub-samples of galaxies with differing distributions of intrinsic properties, it would be incorrect to simply compute the average multiplicative and additive bias from the simulations and use the result as a scalar calibration of the KiDS-450 shear measurements. This is because previous analyses (e.g. Miller et al. 2013; Hoekstra et al. 2015), and analytical arguments (e.g. Massey et al. 2013) have demonstrated that the shear bias depends on galaxy and PSF properties. In particular, we expect the bias to be a function of the galaxy SNR and size, and to depend on the PSF size and ellipticity. Estimating those functional dependencies is crucial in order to derive a shear calibration that may be robustly applied to the data.

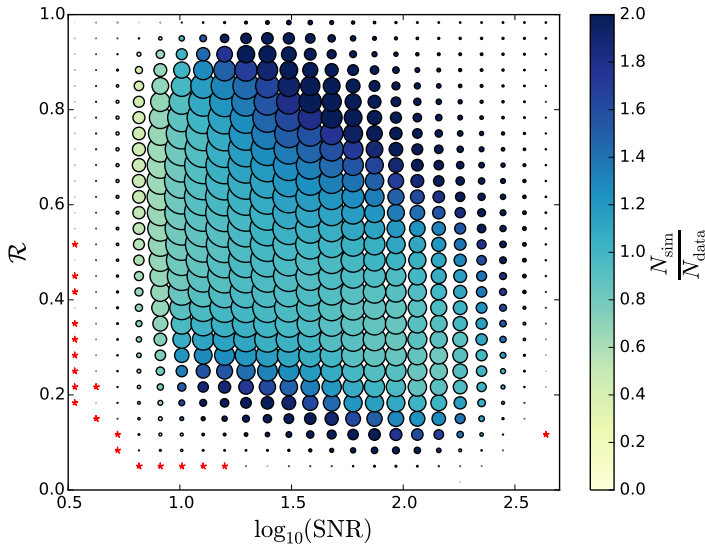


Figure 3.4: Ratio between the number of galaxies in the simulation and the data on a SNR and resolution grid defined using the real galaxies. The size of each data point is proportional to the total *lensfit* weight in each grid cell. The red stars indicate the grid points with a ratio of 0.

A practical procedure for estimating the bias and its dependences from the simulations is to bin the simulated data, and compute the multiplicative and additive shear bias in each bin. To do so, we use the *lensfit* measurements of the galaxy ellipticities ϵ_j in combination with the re-calibrated weights w_j (see §3.2.3) to compute the two components of the measured shear g_j :

$$g_j^{\text{meas}} = \frac{\sum_i w_i \epsilon_{ij}}{\sum_i w_i}. \quad (3.8)$$

Following Heymans et al. (2006) we quantify the shear bias in terms of a multiplicative term m and an additive term c :

$$g_j^{\text{meas}} = (1 + m_j) g_j^{\text{true}} + c_j, \quad (3.9)$$

where we consider the biases for each of the ellipticity components separately. In our analysis below, we designate m, c values for components evaluated in the original ‘sky’ co-ordinate frame by $m_{1,2}, c_{1,2}$. When investigating PSF-dependent anisotropy, we also investigate biases on components where the ellipticity and shear values have been first rotated to a co-ordinate frame that is aligned with the orientation of the major axis of each galaxy’s PSF (c.f. Mandelbaum et al. 2015). We designate the latter linear bias components as $m_{\parallel}, c_{\parallel}, m_{\times}, c_{\times}$ for the components parallel to and at 45° to the PSF orientation, respectively.

Several calibration binning schemes may be considered, such as fixed linear or logarithmic bin sizes, or a scheme that equalises the number of objects in each bin. In the following, we choose a binning scheme that equalises the total *lensfit* weight in each bin and assign the median as the centre of each bin for each respective data sample.

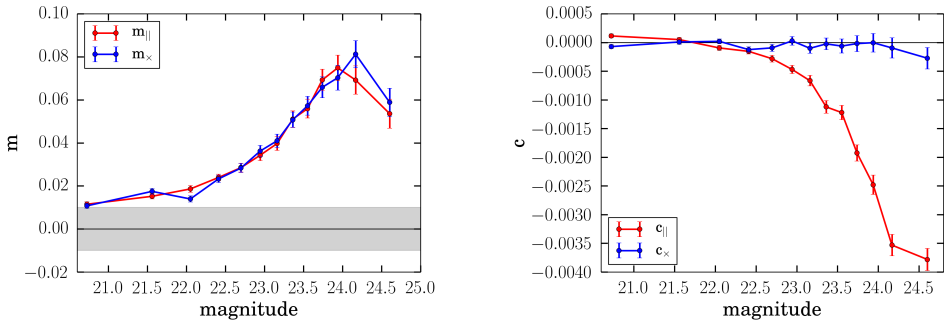


Figure 3.5: Multiplicative (left panel) and additive (right panel) selection bias, m and c , for the components aligned ($m_{\parallel}, c_{\parallel}$) or cross-aligned (m_{\times}, c_{\times}) with the PSF major axis orientation, as a function of galaxy magnitude, as discussed in §3.4.2. The grey band in the left panel indicates the requirement on the knowledge of the multiplicative bias set by Hildebrandt et al. (2016b) in the context of a cosmic shear analysis.

The multiplicative and additive biases for both shear components are then obtained by a linear regression with intersection of all measured average ellipticity values $\langle \epsilon \rangle_j$ against the true input reduced shear values g_j^{true} .

We use two different methods to assign errors to the respective biases in m and c in each bin. In the first method, the uncertainties are estimated from the scatter of the measurements around the best fit line. The other method is to bootstrap resample the sets of galaxies that share the same input shear values. The number of bootstrap realisations is chosen to be large enough for the resulting errors to stabilise. We find this to be the case after the creation of 20 bootstrap realisations.

3.4.2 Selection bias

Bias in the measurement of the shear arises from the combined processes of galaxy detection or selection (‘selection bias’) and the shear measurement itself (‘model bias’ and ‘noise bias’). In this section, we inspect the individual selection bias contributions. Selection biases may occur if the intrinsic ellipticity distribution of galaxies is anisotropic (Kaiser 2000; Bernstein & Jarvis 2002; Hirata & Seljak 2003), which may happen if galaxies are preferentially detected when they are aligned with the shear or the PSF, or if an anisotropic weighting function is employed in the measurement. Multiplicative shear bias may also arise if the distribution of ellipticities that are selected is systematically biased with respect to the underlying distribution. Such anisotropic or multiplicative selection effects may arise at two stages of the process. First, galaxies and stars are detected on stacked images using SExtractor. In principle, the dependence of the SNR on galaxy size, ellipticity, orientation and PSF properties may result in biases at this detection stage. Second, the *lensfit* shear measurement process may not be able to measure useful ellipticity values for some galaxies, leading to an additional contribution to selection bias.

We investigate these biases by inserting the ‘true’ sheared ellipticity value of each simulated galaxy into our shear measurement framework, characterising a linear relation between shear estimates formed from these quantities and the true shear. In this approach, there is no contribution to the bias estimate, or to its measurement

uncertainty, from noise bias. The only potential source of bias is sampling noise, but in our simulations ellipticity shape noise has largely been ‘cancelled’ (see §3.3.2), apart from the effect of galaxies that are not detected. In this test, we find a small bias, $m_{\parallel} \simeq m_x \simeq -0.005 \pm 0.001$, $c_{\parallel} \simeq 0.0002 \pm 0.00004$, $c_x \simeq 0.00005 \pm 0.00004$, as a result of the SExtractor stage. However, if we measure the shear bias after the *lensfit* stage by selecting those galaxies that are both detected by SExtractor and with shear measurement weight greater than zero, we do find a significant multiplicative bias, of 4.4 percent when averaged across the sample, with little difference between biases whether the true shear values are unweighted or weighted by the *lensfit* weight, for those galaxies with non-zero weight. As shown in Fig. 3.5 the bias is strongly magnitude-dependent, with a maximum bias around 8 percent. By rotating galaxy ellipticity and shear values to the coordinate frame aligned with the PSF major axis (the PSF orientation varies in our simulations), we may also look for additive selection bias that is correlated with the PSF: Fig. 3.5 also demonstrates the existence of such an additive selection bias, with a significant aligned c term (there is no significant bias detected in the cross-aligned c term).

The bias is caused by the inability to measure small galaxies: if an object has a *lensfit* star-galaxy discrimination classification that favours the object being a star over a galaxy (see Miller et al. 2013), it is classified as a star and given zero weight in the subsequent analysis. This step introduces a significant selection bias, because galaxies are more easily measured and distinguished from stars if they are more elliptical: thus galaxies whose intrinsic ellipticity is aligned with its shear value are more likely to be selected as measurable galaxies, than those whose intrinsic ellipticity and shear values are cross-aligned. This results in a significant bias in the average intrinsic ellipticity of the measured galaxies, and thus a significant shear bias.

This measurement selection bias should arise in both the data and the simulations, and thus our calibration derived from the simulations should remove the effect from the data. We note however that the selection bias is not small relative to our target accuracy (grey band in Fig. 3.5), and is comparable to the noise bias that has received more attention in the literature. We expect the selection bias to have some sensitivity to the distributions of size and ellipticity and thus not to be precisely reproduced in our fiducial simulations: as previously mentioned, in §3.5 we resample the simulations to match the observed distributions in the KiDS tomographic bins, and in §3.6.2 we further test the effect of modifying the size distribution. We also consider the possible contribution of object selection bias to the PSF leakage in §3.4.6.

3.4.3 Calibration selection bias

In a conventional approach to shear calibration, the objective is to establish a shear calibration relation, whose parameters are observed quantities, which may be applied to the survey data. Ideally, to ensure that unbiased measurements of the cosmology are obtained, after shear calibration has been applied, we should aim for a lack of residual dependence on true, *intrinsic* galaxy properties (such as size or flux) in the simulations, even though the calibration relation must be derived from observed quantities. The absence of such dependencies would imply that the results are not sensitive to changes in the input distributions.

However, if we attempt to deduce a shear calibration that depends on observed quantities, the correlations between observed quantities may cause calibration relations themselves to be biased, and may even mislead the investigator into believing that their shear measurement is biased when it is not. In this section, we discuss biases in

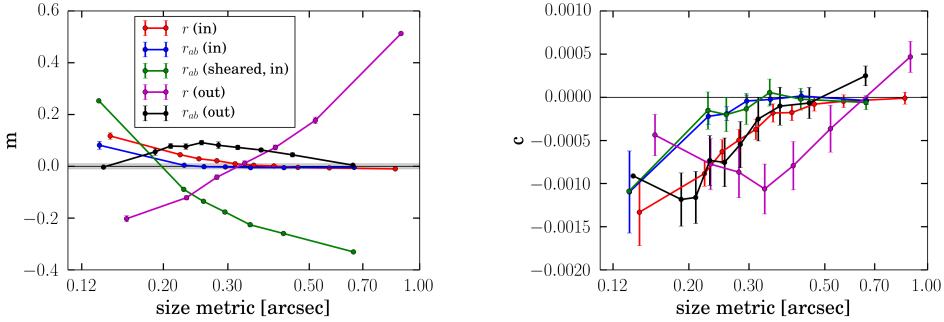


Figure 3.6: The apparent multiplicative (left panel) and additive (right panel) calibration selection bias, m and c , deduced from the analysis of true, noise-free, sheared galaxy ellipticity values, as a function of galaxy size. Relations are shown for five definitions of galaxy size: (red) size r measured from true input major axis values; (magenta) size r measured from noisy output major axis values; (blue) r_{ab} size, measured from true input, unsheared major and minor axis values; (green) r_{ab} size, measured from true input, sheared major and minor values; (black) r_{ab} size, measured from noisy output major and minor values. The additive bias c is shown for the component aligned with the PSF major axis. See §3.4.3.

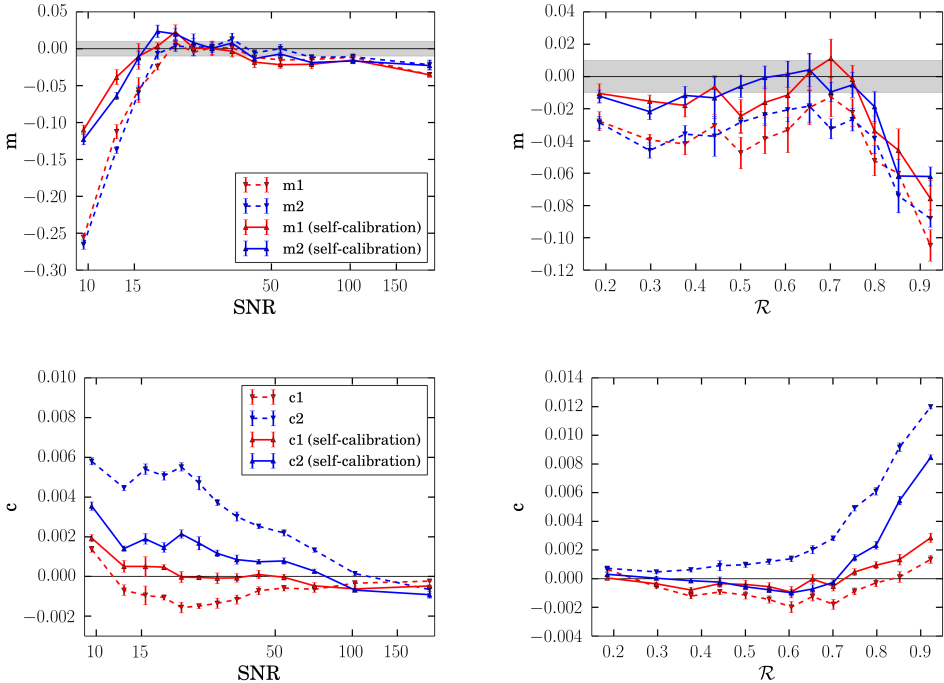


Figure 3.7: The multiplicative shear bias m (top) and additive shear bias c (bottom) as a function of measured galaxy properties. The left panels show the bias with and without *lensfit* self-calibration as a function of measured SNR. The right panels show the same measurements as a function of \mathcal{R} . The grey band in the top panels indicates the requirement on the knowledge of the multiplicative bias set by Hildebrandt et al. (2016b) in the context of a cosmic shear analysis.

calibration relations that arise artificially as a result of correlations between size and ellipticity, and thus shear, when following a calibration approach such as that adopted for CFHTLenS (Miller et al. 2013) or Dark Energy Survey (Jarvis et al. 2015). We distinguish this ‘calibration selection bias’ from the ‘galaxy selection bias’ discussed above, in §3.4.2.

First, we consider the choice of size parameter. The definition of galaxy size measured by *lensfit* is the scale length, r , along the galaxy’s major axis: for disk galaxies, where the ellipticity arises from the inclination of the disk to the line-of-sight, this choice of size measure is the most invariant with the galaxy’s ellipticity. However, at low SNR, pixel noise leads to a strong statistical correlation of the major axis size with the ellipticity. The distribution of observed ellipticity directly affects the inferred shear in a population, and thus a calibration relation that depends on major axis size causes large, apparent size-dependent biases that in fact arise from the choice of observable.

This difficulty may be mitigated by adopting instead r_{ab} , the geometric mean of the major and minor axis scale lengths. In noisy data r_{ab} has significantly lower correlation with the measured ellipticity, but a bias on calibration relations still exists. This selection bias is illustrated in Fig. 3.6. Here, we follow §3.4.2 and again calculate the apparent shear bias that is deduced from using the true, noise-free sheared galaxy ellipticity values. It is important to realise that the biases seen here do not arise from any process in the noisy measurement of shear, other than through the correlation between the size parameter and shear. The blue and red lines show the bias on the input (true) galaxy size, for the r_{ab} and major axis r size definitions respectively: it is this bias that we wish to minimise in order to achieve cosmological results that are unbiased. It may be seen that the r_{ab} measure yields a somewhat lower apparent bias, compared with r , which is a reflection of how the small, unmeasurable galaxies enter each plotted bin. As a comparison, the green curve shows the results for the r_{ab} input size definition, but where now the sheared major and minor axis values have been used to calculate r_{ab} : a very large bias results.

However, any calibration relation that we adopt must instead be a function of the noisy, measured galaxy size, rather than the true size, which is unknown in real data. In Fig. 3.6 (magenta line), we also show that the correlation with the noisy, measured r parameter has a bias that vastly exceeds the input size bias, and which is strongly dependent on the size value. The r_{ab} size definition (black line in Fig. 3.6) is better behaved in this regard, although the bias observed using output size still does not reflect the bias on the input size. On the other hand, the r size definition should be less correlated with ellipticity in the true, astrophysical joint distribution. Hence, we continue to parameterise the *lensfit* models in terms of r , and marginalise over r when estimating galaxy ellipticity as described in §3.2, but we adopt r_{ab} as the size parameter in our calibration relation. We then test how well the bias as a function of input parameters is corrected.

An alternative strategy that would mitigate the selection effects shown in Fig. 3.6 is to subtract the true, intrinsic ellipticity value from every galaxy, before forming any shear estimates: this accurately compensates for the calibration selection bias. This was the procedure adopted for the CFHTLenS shear calibration (Miller et al. 2013), but it has the severe disadvantage that it also removes both the primary selection bias described in §3.4.2 and the weight bias described in §3.2.3. As these are percent-level effects, we must include them in our KiDS calibration, and accordingly do not use this strategy here. We note in passing that neglect of these biases in CFHTLenS may have resulted in larger amplitude shear values (and hence a larger value of the σ_8

Table 3.2: The total multiplicative and additive shear bias, both with (‘self-cal’) or without (‘no-cal’) the *lensfit* self-calibration having been applied. Biases are quoted for components measured either in the co-ordinate system of the sky simulations (upper Table section), or where shear and ellipticity components have been rotated to be aligned, $m_{\parallel}, c_{\parallel}$, or cross-aligned, m_{\times}, c_{\times} , with the PSF orientation (lower Table section).

sky-frame analysis	m_1 [10^{-2}]	$\Delta m_1(\text{regr})/(\text{BS})$ [10^{-2}]	m_2 [10^{-2}]	Δm_2 [10^{-2}]
no-cal	-4.09	0.33/0.25	-3.84	0.21/0.22
self-cal	-1.90	0.33/0.25	-1.68	0.19/0.22
self-cal, no stars	-1.40	0.30/0.29	-1.22	0.18/0.19
self-cal, low density, no stars	-1.39	0.19/0.21	-0.93	0.18/0.26
sky-frame analysis	c_1 [10^{-3}]	Δc_1 [10^{-3}]	c_2 [10^{-3}]	Δc_2 [10^{-3}]
no-cal	-0.73	0.09/0.07	3.32	0.06/0.05
self-cal	0.12	0.05/0.05	1.10	0.05/0.05
self-cal, no stars	0.15	0.09/0.08	1.26	0.05/0.05
self-cal, low density, no stars	0.09	0.05/0.06	0.80	0.05/0.06
PSF-frame analysis	m_{\parallel} [10^{-2}]	$\Delta m_{\parallel}(\text{regr})/(\text{BS})$ [10^{-2}]	m_{\times} [10^{-2}]	Δm_{\times} [10^{-2}]
no-cal	-3.96	0.22/0.43	-3.97	0.20/0.42
self-cal	-1.78	0.18/0.21	-1.79	0.18/0.27
PSF-frame analysis	Δc_{\parallel} [10^{-3}]	c_{\times} [10^{-3}]	Δc_{\times} [10^{-3}]	[10^{-3}]
no-cal	-2.51	0.06/0.10	-0.84	0.06/0.09
self-cal	-0.55	0.05/0.07	-0.15	0.05/0.09

cosmological parameter), by a few percent, than reported by Heymans et al. (2013) and other related cosmology analysis papers.

Finally, we note that Clampitt et al. (2016) found significant size-dependent shear bias in their null test of *Dark Energy Survey* galaxy-galaxy lensing: this bias may have been the result of the selection-induced size bias we have discussed here, and in general, tests of the dependence of shear on measured galaxy size should be avoided as a null test.

In the following sections, we investigate the full bias introduced by the noisy measurement process: this bias includes the object selection bias discussed in §3.4.2 and we should be mindful of the artificial biases of this section when investigating the size dependence and when deriving a calibration relation: biases as a function of galaxy size measured in noisy simulations may have a significant contribution from the calibration selection bias. Provided the simulated galaxy distributions match well the data distributions, any derived calibration relation should correctly include such effects and should result in correctly calibrated data, but it makes sense to minimise the effect of the choice of size definition by calibrating using r_{ab} rather than r , as this should minimise the sensitivity to any mismatch between data and simulations.

3.4.4 *lensfit* results

We start the analysis of the noisy measurement biases by quantifying the impact of the *lensfit* self-calibration (see §3.2.2) on the recovered shear biases. This is done by simply removing the self-calibration corrections (which are reported in the catalogue) from the measured galaxy ellipticities before computing the shear. Without the self-calibration we find that the average multiplicative bias for the full galaxy sample is $\sim 4\%$ in both components. This number reduces to $\sim 2\%$ in each component once we use the *lensfit* self-calibration. We report the exact values, together with their errors, in Table 3.2. Even more dramatic is the reduction of the additive bias when we use the self-calibrated version of *lensfit*: it reduces by a factor five in c_1 and by a factor of three in c_2 . This is extremely encouraging, in particular for cosmic shear analysis, where a large additive bias hampers the ability to measure the cosmological signal at large angular separations (e.g. Heymans et al. 2013; Hildebrandt et al. 2016b).

We also explore the impact of misclassified stars on the average bias in the simulations. In fact, *lensfit* occasionally classifies true stars as galaxies and assigns them a non vanishing weight. As stars are not sheared, the net effect is a reduction of the measured shear and hence a multiplicative bias. By measuring the shear bias either including or excluding these misclassified stars, we quantify the effect of star misclassification on the multiplicative bias as approximately 5×10^{-3} . In the following analysis we keep misclassified stars in the catalogue used to estimate the shear bias. We also ran a set of simulations where the density was lowered by 50 % to explore the effect of galaxy number density on the recovered biases. We found the multiplicative bias to differ by only 2×10^{-3} , suggesting that at the current level of accuracy, simulating the correct number density of galaxies is not crucial for shear calibration, which in turn also implies that galaxy clustering should not impact the shear bias at the KiDS-450 measurement accuracy.

Despite the significant improvements of the self-calibrating *lensfit*, residual shear bias remains, arising from both selection bias and from residual uncorrected noise bias, and we now investigate how the total bias budget is distributed over bins of key input and observed quantities. As discussed above, we expect the shear bias to depend predominantly on the galaxy SNR and on the ratio of the PSF size and galaxy relative size \mathcal{R} , defined by equation 3.7 (Massey et al. 2013). This is confirmed by Fig. 3.7, which shows the multiplicative and additive bias from the simulated data as a function of *lensfit* model SNR and \mathcal{R} with, and without, self-calibration. We notice that at low SNR (and faint magnitude) the self-calibration reduces the multiplicative bias by more than a factor of 2; similar improvements are seen as a function of \mathcal{R} . However, even with self-calibration, the residual multiplicative bias can still be substantially above the 5% level for very faint (low SNR) and very small (large \mathcal{R}) objects. This emphasises the need for an additional, post measurement bias calibration based on the results of the image simulations.

When the self-calibration corrections are included, the residual bias almost vanishes, within its errors, for c_1 but remains significant for c_2 . Motivated by the difference in the two components and in order to explore whether the residual additive bias depends on PSF properties, we perform the same analysis in the PSF frame, by rotating all ellipticity and shear values into a frame where the two axes of the PSF align with the coordinate frame. Once we repeat the bias analysis in the PSF frame, we find that the additive bias is now consistent with zero in the cross-aligned component and that for the PSF-aligned component it has risen to the level we found for the second component in the sky frame. This indicates a dependence of the measured bias on

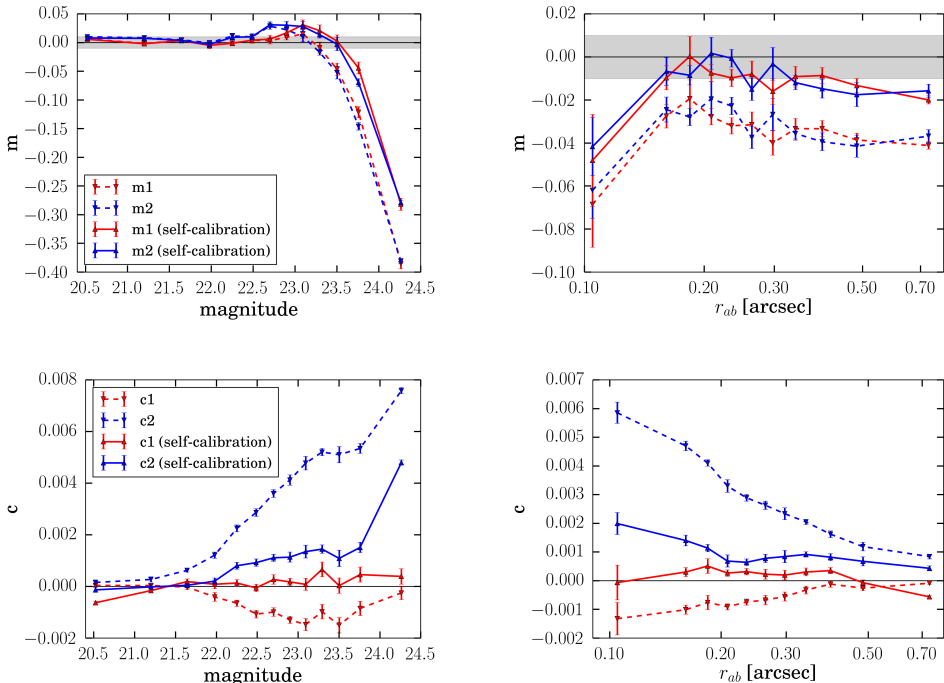


Figure 3.8: The multiplicative bias m (top) and additive bias c (bottom) as a function of simulation input galaxy properties. The left panels shows the bias with and without *lensfit* self-calibration as a function of input magnitude. The right panels shows the same measurements as a function of input size. The grey band in the top panels indicates the requirement on the knowledge of the multiplicative bias set by Hildebrandt et al. (2016b) in the context of a cosmic shear analysis.

PSF properties and motivates a more detailed investigation in §3.4.6.

To explore the dependencies on input parameters, Fig. 3.8 shows the bias in m and c as a function of input magnitude and size. Selection effects are clearly important for the multiplicative bias for faint galaxies, although it should be noted that the most dramatic effects arise at magnitudes $m > 23$, where the galaxy detection is incomplete (Fig. 3.3) and where the weighted contribution to shear measurement is small. In the case of the additive bias, in particular, the utility of self-calibration is evident, as the dependences on input parameters are significantly reduced.

3.4.5 Multiplicative shear bias calibration

The self-calibrated *lensfit* already delivers excellent results in terms of total residual shear bias, as shown in Table 3.2. However, emphasised by Fig. 3.7 and Fig. 3.8, multiplicative biases significantly larger than 5% are still possible, most prominently for faint and small galaxies, although we must be cautious in interpreting any size dependence, owing to the selection bias demonstrated in §3.4.3. We aim here to derive a calibration for the residual multiplicative bias after self-calibration as a function of *lensfit*-measured SNR and \mathcal{R} . While \mathcal{R} is a good choice for characterising the size of a galaxy with respect to the PSF (Massey et al. 2013), one could consider flux-related

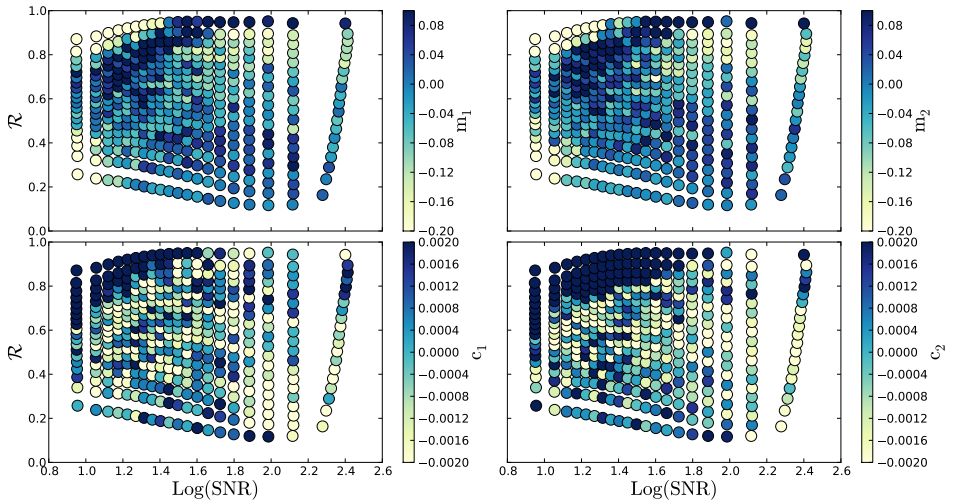


Figure 3.9: The 2D bias surface as a function of model SNR and \mathcal{R} . The *top panels* show the multiplicative bias surface, m_1 on the *left* and m_2 on the *right*. The *bottom panels* show the additive bias components, c_1 on the *left* and c_2 on the *right*. Each point in the plot has equal *lensfit* weight.

calibration quantities other than SNR, for example the observed magnitude, to use as a calibration parameter. However, as discussed in §3.3.5, the real KiDS imaging data has quite some variation of the pixel noise rms, mostly owing to varying observing conditions, while in the simulations we used a fixed value. As the shear bias depends on the noise level and not on the actual flux of the object, it is not possible to derive a robust calibration based on output magnitude.

We bin our simulated data according to the measured galaxy model SNR and \mathcal{R} , again requiring equal *lensfit* weight in each bin and we use the self-calibrated *lensfit* measurements as the default. The two dimensional multiplicative bias surface as a function of SNR and \mathcal{R} is shown in Fig. 3.9. A crucial parameter in such analyses is the total number of bins used to characterise the bias surface. On the one hand, we would like to have a fine enough grid to capture every real feature in the bias surface, but, on the other hand, we have to ensure that there is enough statistical power in each bin so that measurements are not dominated by noise. We tried a variety of grids ranging from only two up to 40 bins on each axis. A coarse 10×10 binning scheme results in an average m-bias error of 2% in both components per bin and increases to an average 10% per bin for the 40×40 scheme. This results in a vanishing signal-to-noise ratio for bins with a small measured bias while using a very fine binning scheme. We found that a 20×20 bin grid provides the best compromise with an average signal-to-noise of 2.5 per bin over the full SNR- \mathcal{R} surface and enough resolution to capture the complicated structure of the bias surface in the low SNR, large \mathcal{R} regime.

Fig. 3.9 reveals that the multiplicative bias surface is complex. Our initial characterisation attempt is based on a fit of an analytic 2D function to the bias surface, as was done for example in Miller et al. (2013); Hoekstra et al. (2015); Jarvis et al.

(2015). Unfortunately, even a complex 16-parameter functional form

$$m_{1/2} = f_0 + f_1 \mathcal{R}^{-1} + f_2 \mathcal{R} + f_3 \mathcal{R}^2, \quad (3.10)$$

where the pre-factors f_i depend on the 16 parameters and the *lensfit* SNR

$$f_i = p_{4i+1} + p_{4i+2} \text{SNR}^{-1} + p_{4i+3} \text{SNR}^{-2} + p_{4i+4} \text{SNR}^{-1/2}, \quad (3.11)$$

for $i \in (0, 1, 2, 3)$ gave only a poor fit to the surface (χ^2 -values of 3.9 and 3.6 for m_1 and m_2 respectively). From now on we will refer to this form of characterisation of the bias surface as method *A*.

Our second attempt to characterise the surface, method *B*, is based on an interpolation of the bias surface. Simple spline interpolation fails to robustly interpolate the bias due to its complicated structure in SNR and \mathcal{R} space. We applied an interpolation scheme based on a Gaussian radial basis function with a spatially varying shape parameter (see Merten 2014, and references therein). The interpolation was trained beforehand using the best-fit analytic functional form of method *A*, to optimally adapt its shape-parameters to the spatial structure of the SNR- \mathcal{R} grid and the general features of the bias surface. The resulting interpolation allowed us to query the multiplicative bias in both components for any parameter pair, at least in the area covered by the given SNR and \mathcal{R} range shown in Fig. 3.9.

Finally, we tried a simpler calibration strategy, method *C*, which was to not fit or interpolate the bias surface, but rather to assign the bias determined in each of the 20×20 bins to the galaxies that fall in each bin.

We test the differing calibration strategies, by investigating the derived multiplicative bias as a function of SNR and \mathcal{R} according to methods *A*, *B* or *C*, for all galaxies with shape measurement in the simulation. In each bin of the analysis we calculate the *lensfit*-weighted average multiplicative bias correction and apply it to the average measured ellipticity in the bin according to equation 3.9. Afterwards, we recalculate the bias. The results for each method are presented in Table 3.2 in terms of the total bias and in Figs. 3.10 and 3.11 as a function of the key output and input quantities. The total multiplicative bias after we apply the calibration is around or below the percent level in both shear components for all three methods. It vanishes completely, by construction, within its error bars for the bin-based calibration method *C*. In terms of our 1% target window, method *A* fails to deliver a robust calibration over the full \mathcal{R} range. Methods *B* and *C* do clearly better and robustly calibrate the residual bias over the full \mathcal{R} range. An exception are extremely small, high \mathcal{R} objects, which represent only a small population in the image simulations. The very last bin in \mathcal{R} , where methods *B* and *C* show a residual bias of 2%, accounts for 7% of the total *lensfit* weight in the sample.

The picture is similar in terms of the calibration performance as a function of SNR. Method *C* performs best and only marginally falls out of our target accuracy for objects with $\text{SNR} < 7$. The reason why this method shows a residual bias at all, is the fact that the binning scheme we used for this analysis differs in both the number of bins and its 1-dimensional nature from the 20×20 SNR- \mathcal{R} binning scheme that we used to derive the calibration. The first SNR bin in Fig. 3.10, where methods *B* and *C* show residual multiplicative biases of -3.5% and 1.5%, respectively, contributes 7% to the *lensfit* weight in the full sample. In the extremely low SNR regime (~ 10), the interpolation based method *B* performs much worse than *C*, likely due to less robust interpolation result near the edges of the initial bias surface. In the final analysis and

considering all mentioned effects, we find that method C provides the most robust calibration of the multiplicative bias and it will be our default method.

In order to test the dependence of this calibration on the number of bins used to characterise the multiplicative bias surface, we investigated the measured bias as a function of the number of 2D bins used. We find that if the number of bins is too small, the calibration is not able to pick up all relevant features in the bias surface and hence existing residual bias remains uncalibrated. Using more than ten bins starts to remedy the problem and a 20 bin scheme is the first calibration that delivers a robust calibration within 1% for the full range of SNR and \mathcal{R} , with the exception of very small objects with $\mathcal{R} > 0.9$, which contribute only a small fraction of the sample’s total *lensfit* weight.

We might hope that when the residual bias, after applying the calibration, is measured as a function of input magnitude and size, it should be consistent with zero. However, this is not the case, as shown in Fig. 3.11. All the calibration schemes show a small positive bias for objects with bright input magnitudes ($m \lesssim 23$) and small galaxies ($r_{ab} \lesssim 0.2''$), and a negative bias at faint magnitudes which becomes large for galaxies below the selection completeness limit. The average weighted bias, however, for the entire simulation, is consistent with zero. The cause of this effect is that the calibration on noisy output quantities relies on there being a stationary correlation between the true quantities and their measured, noisy counterparts. At magnitudes below the completeness limit, the relationship between true size and measured size in the selected galaxies changes, which in turn impacts the calibration relation. In effect, there is a third axis of “magnitude” in our calibration space which has not been included in the calibration relation. In fact, it is not possible to reliably include this third axis, as the three quantities are highly correlated, and also correlated with galaxy ellipticity, and correct calibration in this space would require the joint distributions in the simulations and in the data to match precisely, which is difficult to achieve and is not the case in our simulations.

As by construction, the net residual bias after calibration in the simulations is zero, if the data that we seek to calibrate has the same distribution of true magnitude and size as the simulations, application of the calibration relation should also result in zero residual bias in the calibrated data. However, in reality the data and simulation distributions differ, as shown in Fig. 3.3, and in the cosmic shear analysis (Hildebrandt et al. 2016b) the data are divided into tomographic subsamples, with their own size and magnitude distributions. We investigate the amount of residual bias that might leak into the tomographic analysis presented in Hildebrandt et al. (2016b) via this effect in § 3.5.

3.4.6 Additive shear bias calibration and PSF properties

We have identified the 20×20 grid, bin-based method C as the most robust to calibrate for the remaining residual multiplicative bias. Using exactly the same methodology and by again following equation 3.9 we also characterise the small remaining additive bias not accounted for by *lensfit*’s self-calibration. When calibrating for both, multiplicative and additive bias, simultaneously, we find the residuals shown in the last line of Table 3.3, which is our best and final result.

Fig. 3.12 shows the residual additive bias as a function of SNR and \mathcal{R} before and after calibration and Fig. 3.13 shows the remaining multiplicative and additive bias as a function of PSF properties. This includes the two PSF ellipticity components, the PSF size and “pseudo-Strehl ratio” (defined as the fraction of light contained in the central

Table 3.3: The total multiplicative and additive bias after residual bias calibration.

method	m_1 [10^{-3}]	$\Delta m_1(\text{regr})/(\text{BS})$ [10^{-3}]	m_2 [10^{-3}]	Δm_2 [10^{-3}]
<i>A</i>	3.80	3.35/4.62	4.90	1.88/1.90
<i>B</i>	-1.99	3.35/3.72	-1.89	1.90/2.44
<i>C</i>	-0.008	3.37/3.89	-0.01	1.91/2.49
<i>C</i> (m+c)	-0.008	3.36/4.22	-0.005	1.90/2.72
method	c_1 [10^{-5}]	Δc_1 [10^{-5}]	c_2 [10^{-5}]	Δc_2 [10^{-5}]
<i>A</i>	–	–	–	–
<i>B</i>	–	–	–	–
<i>C</i>	–	–	–	–
<i>C</i> (m+c)	-0.007	9.51/9.38	0.014	5.37/6.66

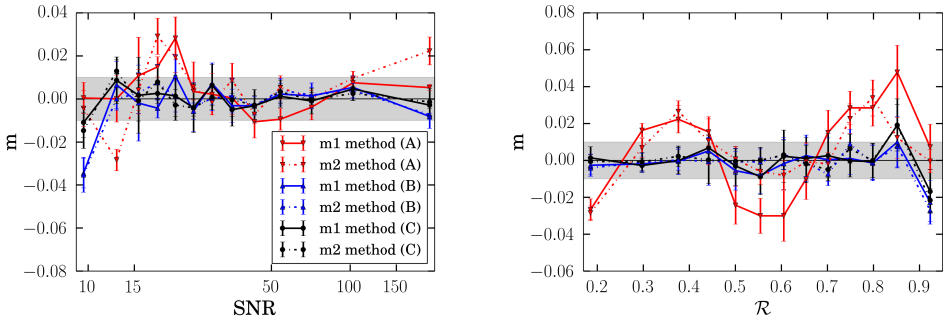


Figure 3.10: The multiplicative bias after empirical calibration using different methods. Method *A* is based on a function form fit to the bias surface, method *B* performs an interpolation of the bias surface and *C* assigns a constant bias correction in 2D bins. The *left panel* shows the residual multiplicative bias after calibration as a function of model SNR and the *right panel* as a function of \mathcal{R} . The grey band indicates the requirement on the knowledge of the multiplicative bias set by Hildebrandt et al. (2016b) in the context of a cosmic shear analysis.

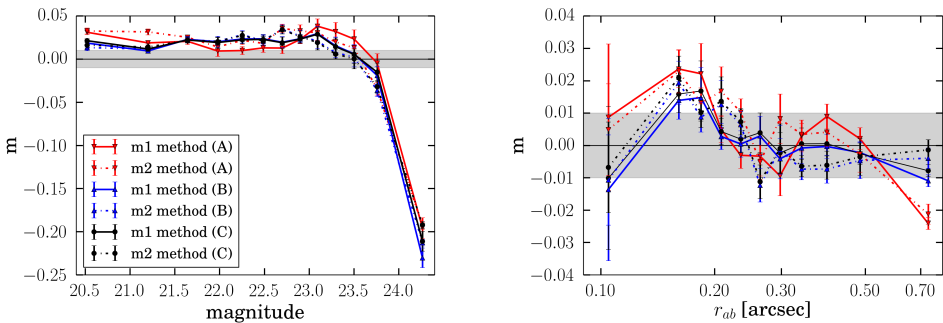


Figure 3.11: This plot is equivalent to Fig. 3.10, but shows the residual multiplicative bias as a function of input magnitude in the *left panel* and as a function of input size in the *right panel*.

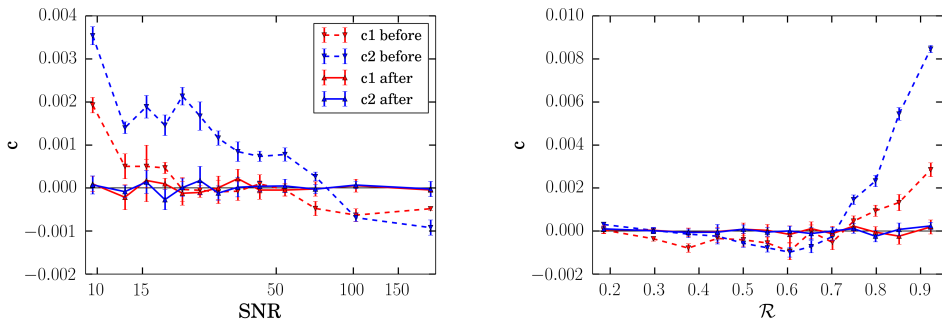


Figure 3.12: The residual additive shear bias before and after calibration using method C . The *left panel* shows residual bias as a function of model SNR and the *right panel* in bins of \mathcal{R} .

pixel of the PSF). All the analyses show no systematic dependence of m - and c -bias on PSF properties and all reported residual biases fulfil, within their errors, our target of 1% residual bias. However, as summarised earlier in Table 3.3, we do detect bias when performing the analysis in the PSF and not in the sky frame. This is expected from the additive selection bias of §3.4.3 and should also have a contribution arising from residual uncorrected noise bias (Miller et al. 2013). In order to characterise this effect we extend our bias description by including a PSF ellipticity dependent term α , following Jarvis et al. (2015):

$$g_j^{\text{meas}} = (1 + m_j)g_j^{\text{true}} + \alpha_j \epsilon_j^{\text{PSF}} + c_j. \quad (3.12)$$

We measure the two α components by subdividing the galaxy sample into bins of the respective PSF ellipticity component. For the full sample, without any further subdivision into bins of galaxy properties we determine $\alpha_1 = -0.006 \pm 0.002$ and $\alpha_2 = 0.005 \pm 0.003$ for the self-calibrated *lensfit* output. It is important to note that no additional residual bias calibration, as described in §3.4.5 and §3.4.6 is applied here. Fig. 3.14 shows the dependence of α , which is sometimes also called PSF leakage, on measured galaxy properties and Fig. 3.15 shows it as a function of simulation input quantities. Clearly, the measurement is significant over the full property range, but is most significant for the low SNR and the small size regime. Fig. 3.14 also shows the bias obtained when true, sheared ellipticity values are propagated through the analysis, as in §3.4.2. We observe that the α -dependence on SNR is well explained by the selection bias, but that there remains α -dependence on the relative galaxy size that appears to have an additional contribution to the selection bias.

In summary, referring to our preferred calibration scheme (method C), all m , c - and α biases vanish for the galaxy sample in its entirety. When looking closer into the biases as a function of measured galaxy properties we find small, of the order 2% residual multiplicative biases for extremely low SNR and extremely high \mathcal{R} objects. All c -biases vanish after our calibration and while residual α terms are presented in the self-calibrated *lensfit* output, they vanish after the additional residual bias calibration. We do expect the PSF-dependent additive biases to be sensitive to the PSF properties, and thus we recommend that the additive bias measured from the simulations is not simply applied blindly to any science analysis. In Hildebrandt et al. (2016b), the additive bias is investigated empirically in the data, and the results compared with

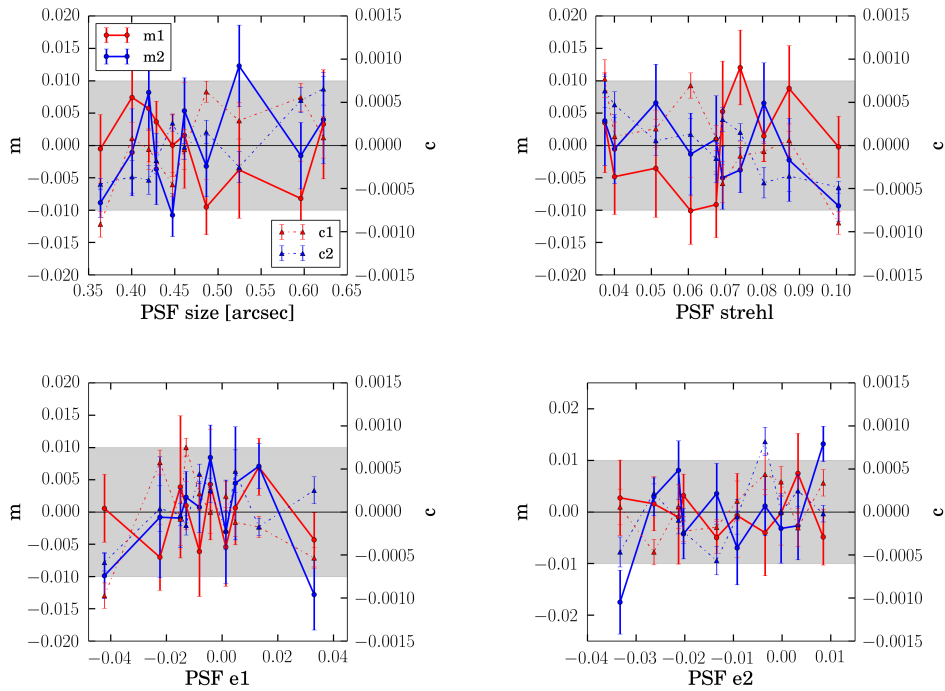


Figure 3.13: The residual bias as a function of PSF properties. The solid lines refer to the residual multiplicative bias with the scale given by the left y-axis. The dot-dashed lines refer to residual additive bias with the scale on the right y-axis in each plot, respectively. The four panels show the biases in clock-wise order starting on the top-right as a function of: measured PSF size, PSF pseudo-Strehl ratio, second PSF ellipticity component and first ellipticity component.

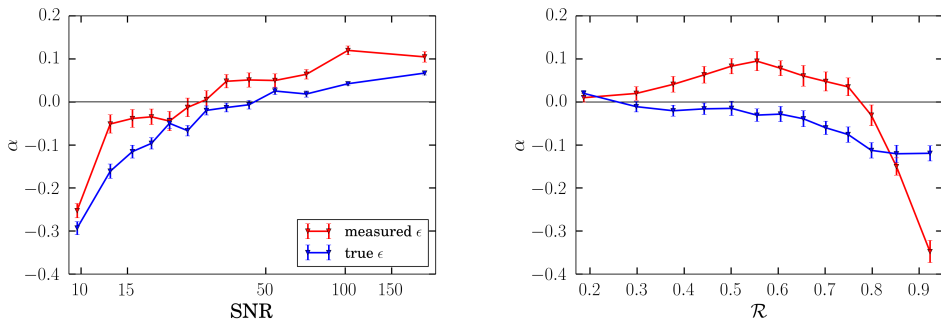


Figure 3.14: The average of the two PSF leakage components, α , as a function of measured galaxy properties, showing the leakage deduced from measured *lensfit* ellipticities (red curves and points) and from true, sheared input ellipticities (blue curves and points), as a test of selection bias. The *left panel* shows α as a function of model SNR, the *right panel* as a function of \mathcal{R} .

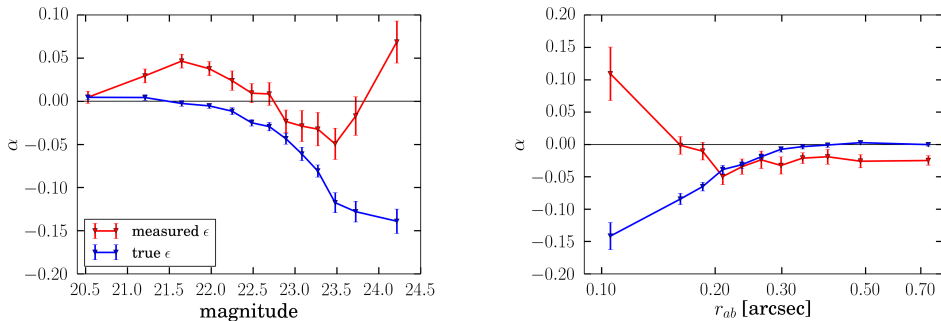


Figure 3.15: The PSF leakage for measured and true ellipticities as a function of simulation input quantities. Input magnitude in the *left panel* and input size in the *right panel*.

those from the simulations, rather than relying on the simulations to be an exact representation of the data regarding its PSF and noise properties.

3.5 Calibration by resampling the simulated catalogue

3.5.1 A resampling approach to calibration

Once the bias has been characterised in terms of relevant observed properties, it can be applied to virtually any selection of the real galaxies used to measure shear. For example, a tomographic cosmic shear analysis requires splitting the galaxy sample into redshift bins; a galaxy-galaxy lensing analysis requires selecting a source sample behind lenses at a given redshift. However, as we saw in §3.4, the bias surface may be complex and thus difficult to characterize, and may itself be biased (see §3.4.3). This may be a concern, given the tight requirements from current and especially future lensing surveys.

The *lensfit* measurements are, however, made for individual objects, and as an

alternative to the approach presented in §3.4, we may instead resample the output from the image simulations, such that the measured galaxy parameter distributions match those of any (sub-)selection of galaxies. The multiplicative and additive biases may then be calculated from the resampled catalogues and applied to the galaxy sample of interest. Note, however, that this approach will only give reliable results if the multi-dimensional parameter space of simulated galaxy properties covers the full parameter space of the real galaxies. Whilst this approach is less flexible than the one described in §3.4, as the simulations need to be resampled for each galaxy sample used to measure shear, it avoids having to characterise the bias as a function of galaxy properties.

Comparison of the biases determined using the different schemes provides an important check on the robustness of the calibration. As described in more detail below, we therefore implemented the resampling approach and applied it to the four tomographic bins used in the cosmic shear analysis presented in Hildebrandt et al. (2016b).

3.5.2 Application to the multiplicative bias in KiDS data

For a given selection of real galaxies, the population of simulated galaxies may be resampled using a k -nearest neighbour search of an N -dimensional volume, defined by a combination of N observed properties of the simulated galaxies. As the search is done by minimising the Euclidian distance between the simulated and real galaxies in that space, it is important to map the distributions of the chosen properties onto a unit length vector. Moreover, there are two important points to consider in order to successfully apply this technique:

- The galaxy properties that define the N -dimensional volume must be correlated with the shear bias;
- The N -dimensional volume of the simulations has to be at least as large as the corresponding volume defined using the properties of the real galaxy sample.

Motivated by the results presented in §3.4, we define the resampling volume based on the galaxy SNR and the ratio of the PSF size and observed galaxy size (\mathcal{R}), for which the simulations cover the same space as the data, as we have shown in §3.3.5. We apply the resampling technique to the selection of galaxies defined by the four tomographic bins used for the cosmic shear analysis presented in Hildebrandt et al. (2016b). Our simulations do not contain any simulated redshift information: we implicitly assume that matching the size and SNR distributions of each tomographic bin is adequate, and that there is no redshift dependence of the bias beyond that conveyed by the bias as a function of SNR and size.

The tomographic bins are defined using the peak of the posterior photometric redshift distribution z_B as measured by BPZ (Benítez 2000) using observations in four optical bands *ugri* (Kuijken et al. 2015). The KiDS-450 data are further divided in five contiguous regions on the sky (designated G9, G12, G15, G23 and GS). We resample the simulations using each region individually, in order to test the robustness of the method, although we note that the SNR and \mathcal{R} distributions are very similar between the regions.

The top panels in Fig. 3.16 show the SNR and \mathcal{R} distributions measured from the KiDS-450 data (all regions combined) and those obtained from the resampled simulations for the third tomographic bin, $0.5 < z_B \leq 0.7$, used in Hildebrandt et al. (2016b).

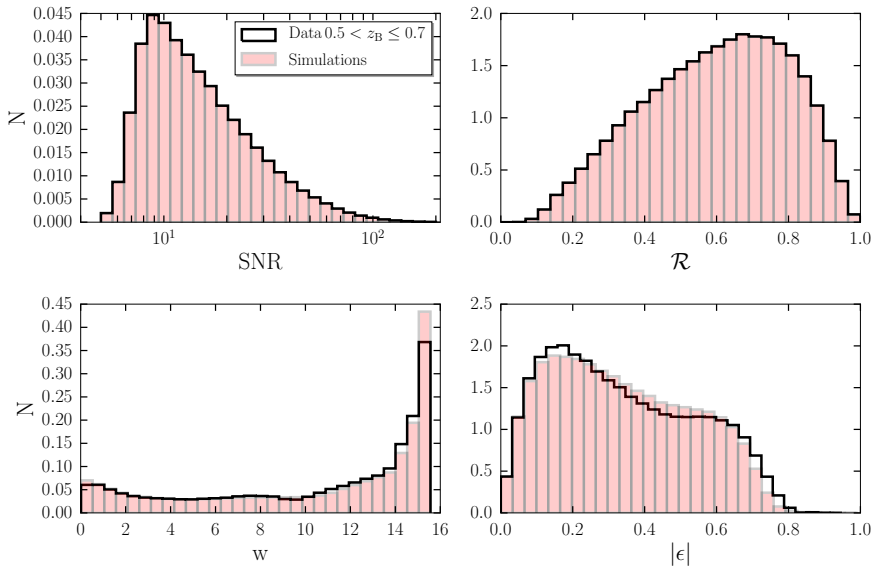


Figure 3.16: *Top panels:* SNR and \mathcal{R} distributions measured from the KiDS-450 data (black line) and using the resampled simulations (red histogram). *Bottom panels:* The distribution of *lensfit* weight (left) and weighted ellipticity (right) measured from the KiDS-450 (black line) and using the resampled simulations (red histogram). All distributions are computed using galaxies in the redshift range $0.5 < z_B \leq 0.7$, which corresponds to the third tomographic bin used in the cosmic shear analysis presented in Hildebrandt et al. (2016b)

The excellent agreement between them validates the resampling technique and confirms that the simulations are representative of the data. In the bottom panels of Fig. 3.16 we show the distributions of the *lensfit* weight and the weighted distribution of the modulus of the ellipticity. As those two quantities were not used in the resampling, it is not surprising that the distributions differ slightly. However, the amplitude of the noise bias depends on the galaxy ellipticity distribution (Viola et al. 2014): we will assess the possible impact of this mismatch on the derived average biases in §3.6.3.

3.5.3 Robustness of the tomographic calibration

From the k -nearest neighbour search we can define a ‘resampling’ weight w^{res} , which is the number of times that a simulated object was matched to an object in the data. We use this new weight in combination with the *lensfit* weight to measure the shear from the resampled simulations:

$$g_j^{\text{obs,res}} \equiv \frac{\sum_i w_i w_i^{\text{res}} \epsilon_{ij}}{\sum_i w_i w_i^{\text{res}}}, \quad (3.13)$$

and compute the multiplicative and additive bias using equation 3.9. We verified that the estimate for the bias is robust against the choice of the number of nearest neighbours. The errors on the biases are also unchanged for $k > 4$. Unless explicitly stated, all the results quoted in this paper have been derived using $k = 5$.

The measured multiplicative bias does not depend on the PSF properties, in agreement with what we found in §3.4. As an additional test we compared the average biases derived from resampling each individual PSF set individually with the results derived from resampling the whole simulation volume. Also in this case we found statistically equivalent results. Fig. 3.17 shows the multiplicative bias derived using the resampling technique and the calibration method presented in §3.4. The hatched regions, centered on the bias measured using the resampling technique indicate the requirements in the knowledge of the multiplicative bias as derived by Hildebrandt et al. (2016a). We compare the results from the two calibration schemes for the four tomographic bins used in Hildebrandt et al. (2016b). The average difference, combining all tomographic bins, is $\Delta m = -0.001 \pm 0.003$.

3.6 Calibration sensitivity analyses

3.6.1 Sensitivity to the magnitude distribution

In §3.4.5 we noted that there might be a residual shear bias that arises from differences between the magnitude distributions of the simulations and of the selection of galaxies in the tomographic bins. We estimate this effect by first applying the method C calibration scheme to the simulations. Then, a new resampling weight is derived for each galaxy, by comparing the *lensfit*-weighted distributions of measured magnitudes in the simulations and in the KiDS-450 data in each tomographic bin, and reweighting the simulated galaxies so that those distributions match.

We measure the residual bias in these reweighted simulations, for each tomographic bin. First, we confirm that the residual bias is consistent with zero in the absence of any magnitude reweighting, as expected. Then, for each tomographic bin reweighting, we find residual bias levels of approximately $-0.001, 0.001, 0.0004, -0.012$ in each of the four bins. The residual bias is consistent with zero in the first three bins, but shows a

percent-level residual in the highest-redshift bin. We cannot know whether this effect is as large in the data as in the simulations, for two reasons: first, we have reweighted using noisy, measured magnitudes rather than true magnitudes, and second we know that the simulations become incomplete at a slightly brighter magnitude limit than the data, so the residual bias effect is expected to be larger in the simulations than in the data. However, this test does indicate the possible size of the residual bias, which is either much smaller than (tomographic bins 1 – 3) or comparable to (tomographic bin 4) our nominal requirement on calibration accuracy.

To explore further the effect of the simulation magnitude limit on the measured shear bias we run another suite of simulations, which are identical to the reference simulations described in Section 3.3, except that we change the noise level, such that the magnitude limit increases by 0.3 magnitude. These simulations are 0.2 magnitude deeper than the KiDS-450 data. We apply the method *C* to these new simulations and we compute the multiplicative shear bias in the four tomographic bins. Compared to the fiducial results we find a change in the bias of -0.008 , -0.003 , -0.006 , -0.014 in each of the four bins. We can use this result to estimate the sensitivity of the bias to the magnitude limit from which we can calculate that the 0.1 magnitude limit difference between the reference simulations and the KiDS-450 data should result in sub-percent residual biases of -0.003 , -0.001 , -0.002 , -0.005 in the four bins.

3.6.2 Sensitivity to the galaxy size distribution

The output galaxy size distribution also differs between the data and the simulations, as shown in Fig. 3.3, which might arise from a difference between the input size distribution we used to create the simulations and the true size distribution of the KiDS-450 galaxies. To examine in more detail the impact of such a difference, we again reweight the galaxies such that the output size distributions of data and simulations match. However, in this case we cannot simply weight by the distribution of output size, as that would not capture correctly the joint dependence of the correlated output size and ellipticity measurements. Instead, we choose to reweight simulated galaxies as a function of their true, input size. We first define an alternative target input size distribution and calculate a ‘size weight’ that may be applied to each galaxy, such that the fiducial input size distribution is transformed from the nominal distribution to the target distribution. The size weight is just the ratio of the values of the target and nominal distributions for each galaxy. The target distribution was varied until a good match of output size distributions was found. The simplest target distribution that was tried had the same functional form as the input size distribution, but with a shift of the median relation by a constant factor to larger sizes, while preserving the magnitude dependence. The factor was varied to obtain the best match between the simulation and data size distributions (as measured by the Kolmogorov-Smirnov statistic), however differences in the distributions remained.

Hence, we also tested a lognormal target distribution, where the median size was again scaled by some factor and where the standard deviation of the distribution of the logarithm was also varied to obtain the best match between data and simulations. This produced a better match, but with some magnitude dependence: a final sophistication then was to allow the slope of the $r_{\text{med}} - m$ relation to vary. The new relation was found to be $r_{\text{med}} = \exp(-1.07 - 0.19(m - 23))$ with standard deviation of the logarithm $\sigma = 0.48$. A good match was then found between the size distributions of the data and the reweighted simulations. The size reweighting also causes some variation in the measured distributions of other quantities, but does not on its own remove the

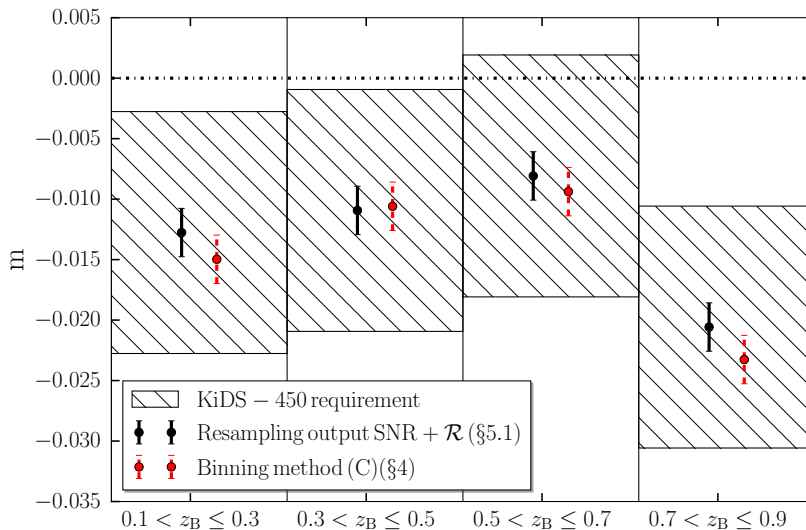


Figure 3.17: Multiplicative bias calculated using the resampling technique and the bias calculated employing the calibration scheme described in §3.4 as a function of the tomographic bins used in the cosmic shear analysis described in Hildebrandt et al. (2016b). The hatched area indicates the requirement on the knowledge of the multiplicative bias for KiDS-450.

discrepancies between the data and simulations in the distributions of magnitude and SNR.

To test the possible effect on the deduced bias, we apply the size reweighting globally to the entire simulation, repeat the bias estimation using method C, and then deduce again the bias for each tomographic bin, as described above. The reweighted bias values differ from the nominal values by -0.0011 , -0.0014 , -0.0013 , 0.0085 in each tomographic bin. The differences in the first three bins are again negligible, with only a sub-percent level effect in the final tomographic bin. That effect has the opposite sign to that found in the magnitude reweighting, which suggests that the joint effect of magnitude- and size-reweighting may be close to zero in all tomographic bins. We conclude that the effect of the uncertainty in either the size or magnitude distributions does not impact our tomographic bin calibration at the level of accuracy required here.

3.6.3 Sensitivity to accuracy of the galaxy ellipticity distribution

A remaining concern is that the recovered ellipticity distribution in the simulations does not match precisely those from the KiDS-450 observations. This may indicate either that the intrinsic ellipticity distribution in the simulations is not the same as in the real Universe, or that some other observed property that is correlated with ellipticity is biasing the distribution. Such a discrepancy in the ellipticity distribution may result in a bias measured from the simulations which may not be applicable to the observations (Melchior & Viola 2012; Viola et al. 2014). To quantify how our results change for different input ellipticity distributions, we perform a further resampling sensitivity analysis, similar to those done by Bruderer et al. (2015) and Hoekstra et al. (2015), that investigates the effect of possible variations in the ellipticity distribution

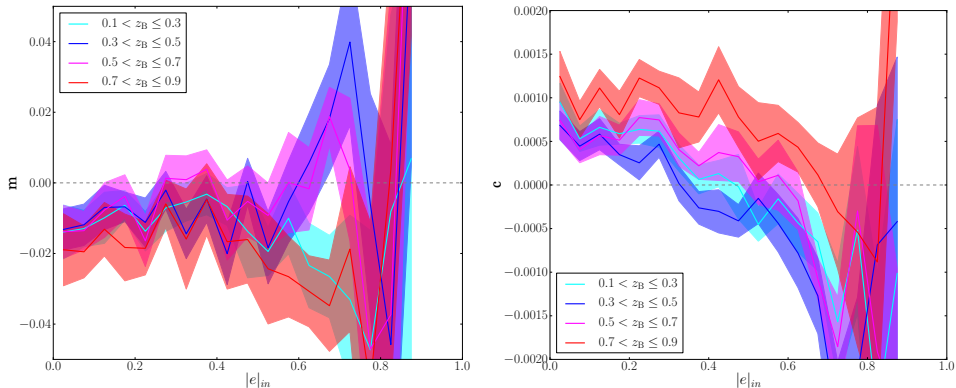


Figure 3.18: Multiplicative bias (left panel) and additive bias (right panel) for bins in input ellipticity for the four tomographic resampled catalogues with 1σ uncertainties. A redder colour indicates a higher redshift tomographic bin.

on the resampling calibration, in tomographic bins (§3.5).

We first quantify the sensitivity of the shear measurement to the input ellipticity distribution, by binning the simulated galaxies according to their input ellipticity, ϵ^s , and computing the multiplicative and additive bias in each ellipticity bin. The results are presented in Fig. 3.18 for the resampled catalogues for the four tomographic bins (see §3.5). Thanks to the resampling, these catalogues have the same observed SNR and resolution distributions as the KiDS-450 data in each tomographic bin. The multiplicative bias depends only weakly on the intrinsic ellipticity for objects with low ellipticities, although the biases differ between tomographic bins. For the additive bias we observe a clear trend with ϵ^s , but we note that the amplitude is low and we do not, in any case, apply our simulated additive bias measurements directly to the data. These findings are in line with the expectations from Viola et al. (2014) and show that modest changes to the input ellipticity distribution should result in at most a percent level effect on the overall multiplicative bias.

The results for the four tomographic bins shown in Fig. 3.18 indicate that the sensitivity of the multiplicative bias to the adopted intrinsic ellipticity distribution is small. Nonetheless, we aim to quantify this further by considering possible variations of the input ellipticity distributions in the simulations. To do so, we follow a similar method to that in §3.6.2, by applying additional weights to the catalogue entries as a function of their input intrinsic ellipticity, and then computing the new, reweighted bias. The difficulty in this approach is that there may be many possible variations of the true ellipticity distribution that result in the same, or similar, measured ellipticity distributions. So, although the principle of resampling is analogous to that done in §3.6.2, here we follow a Monte-Carlo approach to the reweighting, in which we test many possible variations of the true ellipticity distribution, only selecting those that produce a match with the KiDS-450 data. As the input ellipticity is uncorrelated to any other input galaxy property in the simulations, the new weight does not introduce any further bias due to selection effects in our measurements. Here we focus on the ellipticity distribution, but note that this method could be used for other, or multiple, distributions, provided that the simulated volume is large enough. The steps for our sensitivity analysis procedure are as follows:

- We bin the *lensfit* weighted input ellipticity distribution in equally spaced bins $P_i^s(|\epsilon|)$.
- For each input ellipticity bin we determine the corresponding observed ellipticity distribution $\overline{P}_i^{\text{out}}(|\epsilon|)$.
- We assign a weight \tilde{w}_i to each input ellipticity bin, resulting in a modification of both the input and output ellipticity distributions.

In this way we can mimic image simulations with differing input ellipticity distributions, without the need to create and analyse such simulations. For our analysis we have chosen to use 50 bins in input ellipticity. The weights \tilde{w}_i are chosen such that the simulated output ellipticity distribution matches the observed ellipticity distribution in the KiDS-450 data. The intrinsic ellipticity distribution in the Universe varies due to cosmic variance, which limits the precision with which the bias can be determined from our sensitivity analysis. An estimate for cosmic variance can be obtained from the variation in the observed ellipticity distributions between the KiDS-450 patches. We found that these variations are very similar to the Poisson errors on the observed ellipticity distribution. When comparing the ellipticity distributions from simulations and data we therefore assign Poisson errors to the latter.

Matching the observed and simulated ellipticity distributions can only be done reliably if the full range of ellipticities found in the data is encompassed by the simulations. In the course of performing the analysis, we found that the KiDS-450 data contain a small fraction of galaxies with $\epsilon > 0.8$, which are absent in the simulations (see the inset in the lower left panel of Fig. 3.3). In the simulations, such high ellipticities are caused either by measurement noise or by blending of galaxies with close neighbours. To check whether the objects in the data are also caused by noise or blending, we inspected *HST* images of the COSMOS field (Scoville et al. 2007) for which we also have VST *r*-band data. To ensure a fair comparison, we restricted the comparison to images in the *F606W* filter, which is similar to the *r*-band.

Unfortunately, the *F606W* imaging in the COSMOS field only covers 240 arcmin², resulting in a comparison sample of only about 100 galaxies. We found that 70% of these objects were genuinely high-ellipticity, edge-on galaxies, while the rest were either spurious detections or blended objects. The likely cause is that there exists a distribution of the ratio of galaxy disk scale-heights to their scale-lengths (e.g. Unterborn & Ryden 2008), with a tail of galaxies having very thin disks, which are not represented by the nominal ellipticity prior that we assume. Even though the comparison sample is small, this test suggests that the high-ellipticity tail of the *lensfit* prior is not representative of the Universe in this regime. However, the sample is too small to allow us to derive an updated ellipticity prior. Instead, to compensate this incompleteness, we augment our catalogues with very elliptical objects. We created and analysed additional simulations with 2000 galaxies per exposure, adopting a flat input ellipticity distribution with $0.5 \leq |\epsilon| \leq 0.95$. All other properties of the simulations remained unchanged from what has been described in §3.3. Note that the number density of these very elliptical galaxies does not reflect reality, but rather was chosen to provide adequate information for the sensitivity analysis.

We use Monte Carlo Markov Chains (MCMCs) to sample the \tilde{w}_i parameter space. We found that convergence was slow, and the resulting input ellipticity distribution very irregular and spiky if no priors on \tilde{w}_i were imposed. This result is not physical, and does not agree with our limited knowledge of the ellipticity distribution based on high quality data, which indicates a much smoother distribution. To speed up the

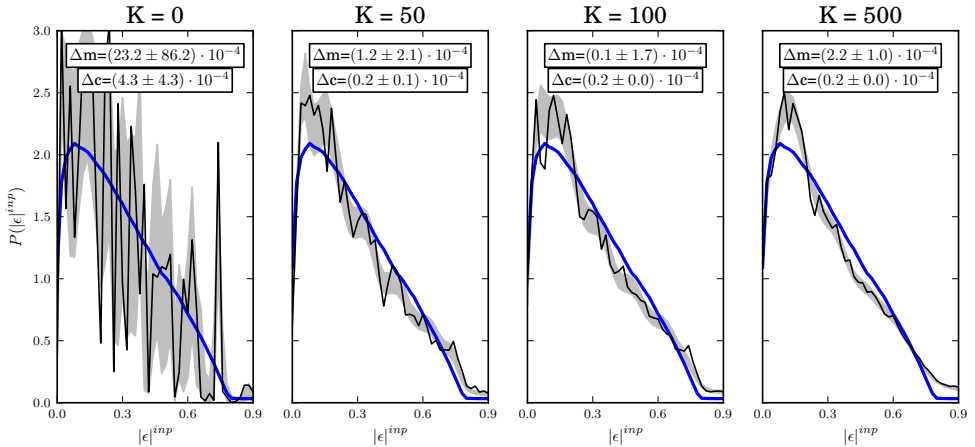


Figure 3.19: Results from the sensitivity analysis based on $0.5 \leq Z_B < 0.7$ galaxies in the G15 patch of the KiDS DR3 data. The intrinsic ellipticity distribution in the resampled catalogue in blue and the distribution which best fits the measured KiDS data and the grey band shows the possible variations from the MCMC tests. To suppress the spiky nature of the best fit we demanded smoother intrinsic ellipticity distribution, finding a strength of the smoothness prior $K = 500$ to be adequate, as indicated at the top of the plot. The bottom row shows how similar the observed ellipticity distribution is to the KiDS-450 data for the resampled catalogue in blue and the best fit in black. The textboxes show the difference in multiplicative (top box) and additive (bottom box) bias between the blue and black distribution. The biases change with K , but all biases are much smaller than the 1% required for cosmic shear.

MCMC runs in finding a more physical solution, we applied a prior to regularise the result. The form of the prior is

$$\pi(K, |\epsilon^s|) := K \times \left| 1 - \frac{P_{i+1}(|\epsilon^s|)}{P_i(|\epsilon^s|)} \right| \frac{|\epsilon^s|_i}{|\epsilon^s|_{i+1}}, \quad (3.14)$$

which penalises a spiky distribution where subsequent bins have very different values. The extra factor of $|\epsilon^s|_i/|\epsilon^s|_{i+1}$ lessens the effect of the prior near $|\epsilon| = 0$, where the distribution turns over. The strength of the prior K should be chosen so that the prior does not dominate. We explored several values of K and found a good compromise for $K = 500$; this choice produced physical distributions in a reasonable amount of computing time.

The third tomographic bin ($0.5 < z_B \leq 0.7$) shows the largest discrepancy between the observed ellipticity distribution in the simulations and KiDS DR3 data and thus serves as a worst case scenario for the sensitivity analysis. We use the ellipticity distribution from patch G15 in the sensitivity analysis and use the 1σ variation between the patches as the error on the distribution. The results of our sensitivity analysis and the effect of the smoothing prior are shown in Fig. 3.19, which shows the input ellipticity distribution of the SCHOOL simulations $P(|\epsilon^s|)$ in blue and the best fit model $\sum_i \tilde{w}_i P(|\epsilon^s|)_i$ from the MCMC results in black. The MCMC chains converged for every run, so that the observed ellipticity distribution was identical to the KiDS ellipticity distribution within the errorbars.

The MCMC framework was able to match the simulations to the data. For the family of modified ellipticity distributions from the MCMC, we compute the standard

deviation in input ellipticity for each bin and show this as the grey band. From left to right the strength of the smoothness prior increases, resulting in smoother distributions. Importantly, the unphysical spike around $|\epsilon^s| = 0.75$ is no longer present in this case. For 1% of the $\sim 2 \times 10^7$ MCMC solutions we computed the shear bias from the corresponding (observed) ellipticity distributions. The difference between the average bias and that measured from the resampled catalogue is shown in the boxes and the error is the 1σ spread of all the computed biases. The difference in ellipticity distribution thus results in only a small change in bias. The biases also change very little as a function of the applied smoothing; the change in multiplicative and additive bias never exceeds 0.3% and 0.01%. These tests show that the shear measurement is quite insensitive to changes in the intrinsic ellipticity distribution and any reasonable variations are within the 1% errors. The discrepancy between the observed ellipticity distribution in the simulations and the data is therefore not a concern for the cosmic shear analysis.

3.7 Conclusions

The large areas covered by ongoing and future imaging surveys dramatically reduce the statistical uncertainties in the measurement of the alignments of galaxies caused by lensing by intervening large-scale structure. This increase in precision needs to be matched by a corresponding improvement in the accuracy with which weak lensing shear can be measured. This can only be achieved by evaluating the performance of shear measurement algorithms on realistic mock data (e.g. Miller et al. 2013; Hoekstra et al. 2015). In this paper we use extensive image simulations created using GALSIM (Rowe et al. 2015), to test and calibrate the *lensfit* algorithm used by Hildebrandt et al. (2016b) to analyse 450 deg^2 (360.3 deg^2 after accounting for masking) of KiDS-450 data. This large survey area implies that the multiplicative bias needs to be determined to better than about 1 percent.

We have shown that the average multiplicative bias over the simulation volume using the self-calibrating *lensfit* algorithm is $\sim 2\%$, and the average additive bias is $\sim 5 \times 10^{-4}$. Although this is close to the required level of accuracy, a final correction is nonetheless required. We have investigated the behaviour of the bias as a function of observed properties of galaxies, such as SNR and size. The measured bias as a function of galaxy properties is a combination of measurement bias, caused by noise, and selection bias, caused by the inability to measure small galaxies and by the weighting of galaxies in the shear measurement process. While it is possible to disentangle those effects in the simulations, it is not possible to do the same in the data. In our analysis, we find that selection bias is at least as important as measurement bias, which implies that even shear measurement methods that are free from, or that perfectly correct for, noise bias may still show shear biases that are present at the percent level or larger.

We have successfully derived a calibration relation that corrects for the dependence of bias on galaxy properties, but we have also shown that this calibration itself may be biased by its use of noisy, measured galaxy properties rather than their unobservable true properties, and these ‘calibration bias’ effects need to be assessed when deriving any new shear calibration. We have tested the accuracy of the application of the calibration relation, including the effect of calibration bias, by a number of resampling tests that were designed to test the accuracy in the four tomographic bins used in the cosmic shear analysis presented by Hildebrandt et al. (2016b). Although there are sub-percent uncertainties in the calibrations arising from the differences between

the data and the simulations, and from the effects of calibration bias, the accuracy of the calibration appears to satisfy the specification required for cosmic shear analysis of the KiDS-450 data set, at 1 percent accuracy of multiplicative bias. In deriving cosmological constraints it is therefore necessary to marginalise over the uncertainty in the shear bias employing a gaussian prior with $\sigma_m = 0.01$. As the SNR and \mathcal{R} distributions in the four tomographic bins are very broad, the shear biases derived from the simulations described in this paper are strongly correlated among tomographic bins. For this reason we conservatively recommend to assume a correlation coefficient of $r=0.99$ between all bins.

Acknowledgements

We would like to thank Joe Zuntz for his detailed referee report which helped improving the paper substantially. We would like to thank Ami Choi, Thomas Erben, Catherine Heymans, Hendrik Hildebrandt and Reiko Nakajima for pipeline testing during the development of self-calibrating *lensfit* using images from the KiDS survey, in addition to the KiDS Collaboration for providing the meta-data upon which the image simulations were based. We would also like to thank Dr Alessio Magro, from the Institute of Space Sciences and Astronomy for many hours of help and support whilst running the simulations pipeline on the cluster. This research has been carried out using computational facilities procured through the European Regional Development Fund, Project ERDF-080 “A supercomputing laboratory for the University of Malta”. JM has received funding from the People Programme (Marie Curie Actions) of the European Unions Seventh Framework Programme (FP7/2007-2013) under REA grant agreement number 627288. HH, RH, MV acknowledge support from the European Research Council FP7 grant number 279396. This work is supported by the Netherlands Organisation for Scientific Research (NWO) through grants 614.001.103 (MV). LM is supported by STFC grant ST/N000919/1.

Author contributions: All authors contributed to the development and writing of this paper. The authorship list is given in alphabetical order (IFC, RH, HHo, JM, LM, MV). MV leads the image simulation working group in the KiDS collaboration.

Bibliography

- Bacon D. J., Refregier A. R., Ellis R. S., 2000, MNRAS, 318, 625
- Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291
- Becker M. R., et al., 2015, preprint, ([arXiv:1507.05598](https://arxiv.org/abs/1507.05598))
- Benítez N., 2000, ApJ, 536, 571
- Bernstein G. M., Jarvis M., 2002, AJ, 123, 583
- Bertin E., 2010, SWarp: Resampling and Co-adding FITS Images Together, Astrophysics Source Code Library ([ascl:1010.068](https://arxiv.org/abs/1010.068))
- Bertin E., 2013, PSFEx: Point Spread Function Extractor, Astrophysics Source Code Library ([ascl:1301.001](https://arxiv.org/abs/1301.001))
- Bertin E., Arnouts S., 1996, A&AS, 117, 393

- Blandford R. D., Saust A. B., Brainerd T. G., Villumsen J. V., 1991, *MNRAS*, 251, 600
- Bridle S., et al., 2010, *MNRAS*, 405, 2044
- Bruderer C., Chang C., Refregier A., Amara A., Berge J., Gamper L., 2015, preprint, ([arXiv:1504.02778](https://arxiv.org/abs/1504.02778))
- Clampitt J., et al., 2016, preprint, ([arXiv:1603.05790](https://arxiv.org/abs/1603.05790))
- Czekaj M. A., Robin A. C., Figueras F., Luri X., Haywood M., 2014, *A&A*, 564, A102
- Dawson W. A., Schneider M. D., Tyson J. A., Jee M. J., 2014, preprint, ([arXiv:1406.1506](https://arxiv.org/abs/1406.1506))
- Fu L., et al., 2008, *A&A*, 479, 9
- Heymans C., et al., 2006, *MNRAS*, 368, 1323
- Heymans C., et al., 2012, *MNRAS*, 427, 146
- Heymans C., et al., 2013, *MNRAS*, 432, 2433
- Hildebrandt H., et al., 2016b, preprint, ([arXiv:1606.05338](https://arxiv.org/abs/1606.05338))
- Hildebrandt H., et al., 2016a, preprint, ([arXiv:1603.07722](https://arxiv.org/abs/1603.07722))
- Hirata C., Seljak U., 2003, *MNRAS*, 343, 459
- Hoekstra H., et al., 2006, *ApJ*, 647, 116
- Hoekstra H., Herbonnet R., Muzzin A., Babul A., Mahdavi A., Viola M., Cacciato M., 2015, *MNRAS*, 449, 685
- Jarvis M., et al., 2015, preprint, ([arXiv:1507.05603](https://arxiv.org/abs/1507.05603))
- Jee M. J., Tyson J. A., Hilbert S., Schneider M. D., Schmidt S., Wittman D., 2015, preprint, ([arXiv:1510.03962](https://arxiv.org/abs/1510.03962))
- Joachimi B., et al., 2015, *Space Sci. Rev.*, 193, 1
- Kacprzak T., Bridle S., Rowe B., Voigt L., Zuntz J., Hirsch M., MacCrann N., 2014, *MNRAS*, 441, 2528
- Kaiser N., 1992, *ApJ*, 388, 272
- Kaiser N., 2000, *ApJ*, 537, 555
- Kaiser N., Wilson G., Luppino G. A., 2000, preprint, ([arXiv:astro-ph/0003338](https://arxiv.org/abs/astro-ph/0003338))
- Kilbinger M., 2015, *Reports on Progress in Physics*, 78, 086901
- Kitching T. D., Miller L., Heymans C. E., van Waerbeke L., Heavens A. F., 2008, *MNRAS*, 390, 149
- Kitching T. D., et al., 2012, *MNRAS*, 423, 3163
- Kuijken K., et al., 2015, *MNRAS*, 454, 3500

- Mandelbaum R., et al., 2015, MNRAS, 450, 2963
- Massey R., et al., 2007, MNRAS, 376, 13
- Massey R., et al., 2013, MNRAS, 429, 661
- Melchior P., Viola M., 2012, MNRAS, 424, 2757
- Merten J., 2014, preprint, ([arXiv:1412.5186](https://arxiv.org/abs/1412.5186))
- Miller L., Kitching T. D., Heymans C., Heavens A. F., van Waerbeke L., 2007, MNRAS, 382, 315
- Miller L., et al., 2013, MNRAS, 429, 2858
- Miralda-Escude J., 1991, ApJ, 380, 1
- Rafelski M., et al., 2015, AJ, 150, 31
- Refregier A., Kacprzak T., Amara A., Bridle S., Rowe B., 2012, MNRAS, 425, 1951
- Rix H.-W., et al., 2004, ApJS, 152, 163
- Robin A. C., Reyl e C., Derri ere S., Picaud S., 2003, A&A, 409, 523
- Rowe B. T. P., et al., 2015, Astronomy and Computing, 10, 121
- Schrabback T., Hartlap J., Joachimi B., et al. 2010, A&A, 516, A63
- Scoville N., et al., 2007, ApJS, 172, 1
- Unterborn C. T., Ryden B. S., 2008, ApJ, 687, 976
- Van Waerbeke L., et al., 2000, A&A, 358, 30
- Viola M., Kitching T. D., Joachimi B., 2014, MNRAS, 439, 1909
- Wittman D. M., Tyson J. A., Kirkman D., Dell’Antonio I., Bernstein G., 2000, Nature, 405, 143
- Zuntz J., Kacprzak T., Voigt L., Hirsch M., Rowe B., Bridle S., 2013, MNRAS, 434, 1604
- de Jong J. T. A., et al., 2015, A&A, 582, A62

3.A Model bias

The measurements used for KiDS-450 may suffer from “model bias”, if the assumed model surface brightness distributions are mismatched to the true distributions of galaxies (e.g. Kacprzak et al. 2014; Zuntz et al. 2013). Results from the GREAT3 challenge suggest that the amplitude of such bias is sub-percent and hence is subdominant compared to the ~ 1 percent systematic uncertainties on the shear calibration arising from other effects that we estimate in this work. To verify this, here we describe a differential measurement between the shear recovered from a population of synthetic galaxies generated by GALSIM (Rowe et al. 2015) using *Hubble Space Telescope* (HST)

images of faint galaxies and the shear recovered from a population of galaxies made with synthetic bulge-plus-disk models whose distributions of sizes and shapes match the HST galaxies.

First, a simulation was created using postage stamps of high resolution HST galaxies, with i-band magnitude between 20 and 24.5, which are available in GALSIM. Each galaxy was sheared and convolved with the median KiDS PSF (FWHM=0.64", Moffat $\beta=3.14$, $\epsilon_1=0.08$, $\epsilon_2=-0.05$) and rendered to a pixel scale of 0.214". The flux is the same for each object and set high enough with respect to the noise level, so that noise bias in the measurements is small. The simulated images consist of a grid of approximately 50 000 isolated galaxies, so that blended galaxy isophotes do not influence the shape measurement. As was done for the fiducial simulations (see §3.3), four rotations of each galaxy were used to reduce shape noise and the same 8 shear values were tested. Given the high SNR of the galaxies and the use of four rotations, the simulated volume is large enough to achieve per mille precision in the shear bias determination.

SEXTRACTOR was run on the simulated images with the same configuration used in the analysis of the KiDS-450 data. About 1% of the HST galaxies were incorrectly segmented and flagged by *lensfit* in the subsequent analysis as blended. We visually inspected several postage stamps and indeed confirmed that these HST images showed unphysical features, such as a large number of negative pixels, creating problems for SEXTRACTOR. Furthermore another $\sim 1\%$ of objects were flagged by *lensfit* and assigned a weight of zero. In order to retain the rotational symmetry we used in the subsequent analysis only galaxies for which all the 32 renditions (4 rotations time 8 shears) have a weight larger than zero and are unflagged, as would be the case in a survey of real galaxies.

We then reran the same simulation without applying the shear to the galaxies. This was necessary to determine the distributions of intrinsic galaxy properties for the input for the synthetic galaxy simulation. The modulus of the intrinsic ellipticity of each HST galaxy was obtained by averaging the modulus of the measured *lensfit* ellipticity of the four rotations. As before, only if all four rotations were properly detected and had non-zero weight, were they included in the average. Similarly we obtained the intrinsic scale lengths and bulge fractions.

The comparison set of simulations were created using synthetic galaxies, adopting a bulge plus disk model. The modulus of the intrinsic ellipticity, the size and the bulge fraction were drawn from the measured distribution in the real galaxy simulation. The intrinsic position angle of galaxies was randomly assigned from a uniform distribution. This procedure ensures that the distributions between the first and the second set of simulations are the same and it also removes any bias in the *lensfit* measurements correlated with the shear. These galaxies were sheared, in the same way as it was done for the HST galaxy simulations, and convolved with the same PSF.

Finally, the same analysis was run as described in Section §3.4 on the two catalogues and we compared the average biases. The HST galaxies showed an average multiplicative bias $m = -0.002 \pm 0.002$, while the bulge-plus-disk galaxy simulations the average bias was $m = -0.001 \pm 0.002$. We conclude that there is no evidence of a *lensfit* multiplicative bias larger than couple of permille. This is in line with the previous results achieved on the GREAT3 benchmark simulations.

4

Multi Epoch Nearby Cluster Survey: Weak lensing masses for 48 local galaxy clusters

Large surveys have detected significant samples of galaxy clusters with well-understood selection functions. These can be used to constrain cosmological parameters, provided that their masses can be measured robustly. To extend the calibration of cluster masses using weak gravitational lensing we present results for 48 clusters with $0.05 < z < 0.15$, observed as part of the Multi Epoch Nearby Cluster Survey (MENeCS). Our measurements benefit from the low cluster redshifts, which decrease contamination of the source sample by cluster members and reduce the sensitivity to uncertainties in the source redshift distribution. Combined with advances in shape measurements we estimate that the systematic uncertainties in the lensing signal are less than 3%, sufficient for the size of the MENeCS sample. We compute physical cluster properties by fitting parametric models to the contamination corrected weak lensing signal. The weak lensing masses and velocity dispersions are in fair agreement with estimates based on galaxy dynamics and we find consistent relations for MENeCS and the Canadian Cluster Comparison Project. We derive a scaling relation with hydrostatic masses using *Planck* measurements and find a bias in the hydrostatic masses $1 - b = 0.90 \pm 0.05(\text{stat}) \pm 0.03(\text{syst})$ when combining both cluster samples. The data support a decreasing trend of $1 - b$ with mass, which is in agreement with other observations.

R. Herbonnet, C. Sifón, H. Hoekstra, R. F. J. Van der Burg
to be submitted

4.1 Introduction

The growth rate of massive structures is sensitive to cosmology as gravitational build-up of overdensities in the initial density distribution is counteracted by the expansion of the Universe. Numerical simulations can predict the abundance of massive structures for varying cosmologies and linking these to such objects in the real Universe allows for cosmological tests. Although the bulk of the mass in these structures is in the form of dark matter, they are observable across the electro-magnetic spectrum because they contain large amounts of baryons that manifest their presence in various ways, such as clusters of galaxies and hot gas. Studies of the number of clusters as a function of mass and redshift (cluster mass function) have put tight constraints on the energy density of matter Ω_m and normalisation of the matter power spectrum σ_8 and the redshift evolution of the mass function can constrain the abundance and the equation of state of dark energy and the number of neutrino species (Borgani & Guzzo 2001; Vikhlinin et al. 2009b; Planck Collaboration et al. 2014; Mantz et al. 2015; Planck Collaboration et al. 2016b; de Haan et al. 2016). See also Allen et al. (2011) for a general review on galaxy clusters as a cosmological tool.

Determination of the cluster mass function requires a large sample of clusters representative of the whole population *and* accurate mass estimates of those clusters. The number of observed clusters is steadily increasing thanks to optical searches for overdensities of (red) galaxies (e.g. Gladders & Yee 2005; Rykoff et al. 2016), and X-ray surveys looking for diffuse hot intracluster gas (e.g. Böhringer et al. 2004; Vikhlinin et al. 2009a). In recent years millimeter wavelength observations have added greatly to the number of detected clusters (Hasselfield et al. 2013; Bleem et al. 2015; Planck Collaboration et al. 2016c). In galaxy clusters photons from the cosmic microwave background (CMB) undergo inverse Compton scattering off the hot intracluster gas thereby obtaining a slight net boost in energy (Sunyaev & Zeldovich 1972, SZ). This SZ effect introduces a characteristic distortion in the millimetre part of the spectral energy distribution, which is a tell-tale sign of a massive galaxy cluster. CMB photons are present at all observable redshifts and the SZ signal scales linearly with gas density making it observable even for high redshift clusters with relatively low gas density.

All these surveys detect clusters based on a selection function (such as signal to noise ratio cuts), which can make the sample unrepresentative of the underlying distribution of clusters. The selection function needs to be taken into account, lest the cluster mass function be biased (see e.g. Mantz et al. 2010; Battaglia et al. 2016 for the effects of selection functions).

The second requirement for robust estimates of cosmological parameters is a well calibrated relation between survey observable and mass¹. In fact, the lack of a reliable scaling relation is the main limitation for the full exploitation of the all-sky *Planck* cluster catalogue. The total mass of clusters can be computed using kinematics of cluster members under the assumption of dynamical equilibrium (e.g. Ruel et al. 2014; Bocquet et al. 2015; Sifón et al. 2016; Amodeo et al. 2017) or using caustics (Rines et al. 2016). The X-ray temperature and surface brightness profile can be connected to mass, but this is usually done under the assumption of hydrostatic equilibrium (see e.g. Battaglia et al. (2012) for an assessment of the impact of non thermal pressure on cluster mass estimation). As clusters formed through mergers, most clusters will not be in a state of equilibrium. Numerical hydrodynamical simulations consistently suggest

¹Actually, because of degeneracy between cosmological and astrophysical parameters in the estimation, the scaling relation should be inferred simultaneously with cosmological parameters (Mantz et al. 2010, e.g.)

that X-ray masses can be biased low by $\sim 10\%$ - 35% depending on the dynamical state of the cluster (Nagai et al. 2007; Rasia et al. 2012; Henson et al. 2017).

Weak gravitational lensing can provide the total mass of a cluster against which other mass proxies can be calibrated. A galaxy cluster acts as a lens because its gravitational potential distorts the surrounding space-time which deflects photons from their straight line trajectories. This phenomenon introduces a coherent distortion (shear) in the observed shape of background galaxies. The lensing signal is not sensitive to the dynamical state of the cluster and can be a source for unbiased mass estimates. However, lensing is a probe for the gravitational potential projected along the line of sight and the triaxial distribution of mass introduces an uncertainty of $\sim 10\%$ - 30% in the weak lensing estimates (Corless & King 2007; Meneghetti et al. 2010; Becker & Kravtsov 2011; Rasia et al. 2012; Henson et al. 2017). Moreover, uncorrelated large scale structure also affects the lensing signal introducing extra scatter in the mass estimates (Hoekstra 2001; Hoekstra et al. 2011a). For a large sample of clusters these uncertainties should average out, so reliable scaling relations can only be produced for large samples of clusters.

A number of studies have estimated weak lensing masses for galaxy clusters with the aim to constrain scaling relations (Jee et al. 2011; Okabe et al. 2013; von der Linden et al. 2014b; Kettula et al. 2015; Hoekstra et al. 2015; Okabe & Smith 2016; Penna-Lima et al. 2016; Schrabback et al. 2016; Sereno et al. 2017). The large numbers of studied clusters help to reduce the statistical uncertainties, but also to quantify the intrinsic scatter of the scaling relations, which can help to reveal the underlying astrophysical origins. The largest weak lensing surveys have targeted several tens of galaxy clusters so that statistical errors can become comparable to systematic uncertainties. Therefore, the main focus for the Weighing the Giants survey (WtG, von der Linden et al. 2014a; Applegate et al. 2014), the Canadian Cluster Comparison Project (CCCP, Hoekstra et al. 2012, 2015), the Local Cluster Substructure Survey (LoCuSS, Okabe & Smith 2016) and Cluster Lensing and Supernovae survey with Hubble (CLASH, Umetsu et al. 2014) has been to robustly assess sources of systematic errors. Weak lensing experiments measure the shear by averaging the shapes of galaxies behind the clusters, and combine these with distance estimates for the background galaxies in order to reconstruct the mass profile. The background galaxies are predominantly faint objects, so the distances are computed using photometric redshifts. Systematics are thus introduced by biased measurements of the galaxy shapes or of the galaxy redshifts, a false classification of objects as background galaxies, and an incorrect estimation of the mass profile. These effects have become more important as different teams have found inconsistent mass estimates for the same clusters. To investigate the discrepancies between Applegate et al. (2014) and Hoekstra et al. (2012), the CCCP cluster mass estimates were revisited by Hoekstra et al. (2015, hereafter H15). They used a large set of simulations of telescope images to calibrate their shape measurement algorithm to an accuracy of 2%. Unlike the WtG data, the CCCP data did not have enough colour data for their galaxies to compute photometric redshifts. This lack of redshifts was a major concern, but was overcome with a separate high fidelity redshift catalogue and a correction for the contaminated sample of background galaxies. With these improvements they found that shear- and masses estimates were consistent between the two teams within the measurement uncertainties.

With the robust pipeline of H15, systematic errors are low enough that they are comparable to the statistical uncertainties. In this work, we build on the work of H15 by studying another sample of clusters, which was observed with the Canada-France-

Hawaii Telescope (CFHT), as was CCCP, and analyse it with the same pipeline. The Multi Epoch Nearby Cluster Survey (MENeCS) provides excellent quality optical imaging data in the g and r -band for a sample of 58 X-ray selected clusters at $0.05 < z < 0.15$. MENeCS presents a significant collection of clusters allowing for a precise determination of the average cluster mass. However, as was the case for CCCP, the trade-off for the large sample size is the lack of colour information required to estimate photometric redshifts for all observed galaxies. Fortunately, the systematic errors due to the lack of individual redshifts are much less severe thanks to the low redshifts of the clusters compared to the CCCP analysis. The local MENeCS cluster galaxies are spread over a larger part of the sky and thus the number density of cluster members contaminating the background population should be small. The lensing signal for low redshift clusters is also fairly insensitive to the actual distribution of redshifts of background galaxies (Hoekstra et al. 2011a). These qualities make MENeCS an excellent sample of galaxy clusters with which to estimate robust weak lensing masses.

The MENeCS observations are briefly described in Section 4.2, where we also present details on the pipeline used to determine galaxy shapes. The MENeCS observations lack the colour information required for accurate photometric redshifts and instead we determine a distribution of redshifts for the background galaxy population using ancillary data. This process is described in Section 4.3. Without reliable photometric redshifts, galaxies cannot be separated into a population associated to the cluster and a population of gravitationally lensed background galaxies. We derive a correction for this contamination in Section 4.4. In Section 4.5 we determine the physical properties of the MENeCS clusters by fitting parametric density models to the corrected weak lensing signal. We then assess the robustness of our results by comparing them against estimates derived using galaxy dynamics. Finally we compute a scaling relation with the masses estimated using the *Planck* CMB measurements (Planck Collaboration et al. 2016c) and we conclude in Section 4.6. Throughout the paper we assume a flat Λ cold dark matter cosmology where $H_0=70h_{70}$ km/s/Mpc and the current energy densities of matter and dark energy are $\Omega_m(z=0) = 0.3$ and $\Omega_\Lambda(z=0) = 0.7$, respectively. All masses and radii scale as h_{70}^{-1} .

4.2 Data and shape analysis

The Multi Epoch Nearby Cluster Survey (MENeCS) is a deep, wide-field imaging survey of a sample of X-ray selected clusters with $0.05 < z < 0.15$. The data were obtained with two main science objectives in mind. The first, the study of the dark matter halos of cluster galaxies using weak gravitational lensing, defined the required total integration time and image quality, as well as the redshift range; with ground-based observations such studies are best done with low redshift clusters. The results of this analysis are presented in **Chapter 5**. Taking advantage of the queue scheduling of CFHT observations, however, the observations were spread over a two-year period, which enabled a unique survey to study the rate of supernovae in clusters (Sand et al. 2012; Graham et al. 2012), including intra-cluster supernovae (Sand et al. 2011). To do so, typically two 120s exposures in the g and r -band were obtained for each epoch (which are a lunation apart). The full sample comprises the 58 most X-ray luminous clusters that were catalogued at the start of the survey, whilst observable with the CFHT. A detailed description of the survey is presented in Sand et al. (2012)². All

²We note that Sand et al. (2012) did not present results for Abell 401, which we have included in the total sample.

clusters used in our analysis are listed in Table 4.1 together with their redshift and the coordinates of the brightest cluster galaxy (BCG) which we take as the centre of the cluster. A possible complication is the misclassification of a galaxy as the BCG (see e.g. Bildfell et al. 2008). However, it was straightforward for the MENeCS clusters from visual inspection which galaxy was the BCG.

In this paper we use the r -band data to determine the cluster masses using weak gravitational lensing. The individual exposures are pre-processed using the **Elixir** pipeline (Magnier & Cuillandre 2004), and we refine the astrometry using **Scamp** (Bertin 2006). Although the CFHT observations were typically obtained when the seeing was below $1''$, some exposures suffer from a larger PSF. As this is detrimental for accurate shape measurements, these exposure were excluded when co-adding the data. For each cluster the 20 frames with the best image quality were selected and combined into a single deep mosaic using **Swarp** (Bertin 2010). However, if additional frames had a seeing full width at half maximum less than 0.80 arcseconds they were added to the stack. The minimal depth of each mosaic is 40 minutes of exposure time. The magnitudes we use are corrected for Galactic extinction using the Schlafly & Finkbeiner (2011) recalibration of the Schlegel et al. (1998) infrared-based dust map. For the analysis presented here, we excluded 9 clusters based on their r -band Galactic dust extinction A_r . The threshold value $A_r < 0.2$ was chosen to reflect the range in which we can reliably correct for contamination (see Section 4.4 and Appendix 4.A). Finally, the cluster Abell 763 contained no significant overdensity of galaxies and was removed from the sample. Table 4.1 lists for all selected clusters the characteristics of the image quality of the mosaic: the half-light radius of the PSF, the root mean square (r.m.s.) noise level and the Galactic extinction.

Objects were detected in the mosaics using the pipeline described in Hoekstra et al. (2012). To measure the weak lensing signal around the clusters we select objects with an r -band magnitude $20 \leq m_r \leq 24.5$. Following H15 an upper limit of 5 pixels on the galaxy half-light radius is imposed. A lower limit for the size is set by the size of the PSF, which removes stars and small galaxies that have highly biased shapes. Galaxies are assigned a lensing weight

$$w = \left[\langle \epsilon_{\text{int}}^2 \rangle + \left(\frac{\sigma_\chi}{P^\gamma} \right)^2 \right]^{-1}, \quad (4.1)$$

where $\langle \epsilon_{\text{int}}^2 \rangle = 0.25^2$ is the dispersion in the distribution of intrinsic ellipticities and σ_χ is an estimate of the uncertainty in the measured polarisation value χ due to noise in the image (Hoekstra et al. 2000). The polarisation is measured using a weight function to reduce the effect of noise, which introduces a bias in the final shear estimate. The shear polarisability P^γ corrects the polarisation for the use of the weight function and for PSF smearing. The shear is then computed as the weighted average of the corrected polarisations

$$g_i = \frac{\sum_n w_n \chi_{i,n} / P_n^\gamma}{\sum_n w_n}, \quad (4.2)$$

where the index i indicates the two Cartesian components of the shear and the sum runs over all galaxies in the sample. In practice, we measure the reduced shear $g_i = \gamma_i / (1 - \kappa_i)$ (Bartelmann & Schneider 2001), which deviates from the true shear γ near the cluster centre, but for most radii of interest the convergence κ is negligible, although we take it into account in our analysis. We decompose the shear into a cross and tangential component relative to the lens, where the tangential shear g_t can be related to the

projected mass of the lens and the cross shear can be used to find systematic errors (Schneider 2003).

The galaxy polarisations and polarisabilities are measured from the mosaics using the shape measurement algorithm detailed in H15, which is based on the moment-based method of Kaiser et al. (1995). H15 used extensive image simulations to quantify the multiplicative bias that arises from noise in the data and the imperfect correction for the blurring by the PSF. The MENeACS data are similar in terms of depth and image quality compared to the observations of the CCCP that were analysed in H15; therefore we use the same correction scheme. A potentially important difference with the CCCP analysis is that the individual exposures are offset from one another. This could lead to a complicated PSF pattern in the combined images. However, tests on the CCCP data indicate that this results in a negligible change in the mass estimates. Moreover, the large number of exposures, combined with the smooth PSF pattern results in a smooth PSF when measured from the mosaics.

Galaxy magnitudes are corrected for background light by subdividing pixels without galaxy light in an annulus between 16 and 32 pixels into four quadrants and fitting the quadrants with a plane to allow for spatial variation of the background. We found that bright neighbouring objects affect this local background subtraction, which in turn affects the shape measurement. When we examined the performance of the algorithm near bright cluster members in image simulations for the purpose of studying the lensing signal around such galaxies (see **Chapter 5**), there were cases where m_{find} , the apparent magnitude as measured by the detection algorithm differed from m_{shape} , the magnitude measured by the shape measurement algorithm. No background light was present in the simulations and instead the local background subtraction was affected by the proximity of bright cluster galaxies. We introduced a flag that identified galaxies for which the shape measurement is biased due to the background subtraction. A cut based on $\Delta m = m_{\text{find}} - m_{\text{shape}}$ of

$$\Delta m > -49.0 - 7.0m_{\text{shape}} + 0.3m_{\text{shape}}^2 - 0.005m_{\text{shape}}^3 \quad (4.3)$$

efficiently removed problematic objects. We therefore apply this cut to the full lensing catalog, which removes approximately 10% of the sources. We applied the same selection to the image simulations studied in H15 and found that the biases are unchanged. Consequently, we use the same parameters to correct for the biases in the method. H15 estimate that the systematic uncertainties in the cluster masses caused by the shape measurements is less than 2%, which is also adequate for the results presented here.

4.3 Photometric source redshift distribution

Gravitational lensing is a geometric phenomenon and the amplitude of the effect depends on the distances involved. This dependency is parametrised by the critical surface density

$$\Sigma_{\text{crit}} = \frac{c^2}{4\pi G} \frac{D_{\text{os}}}{D_{\text{ol}} D_{\text{ls}}}, \quad (4.4)$$

where the lensing efficiency $\beta = \max(0, D_{\text{ls}}/D_{\text{os}})$ contains the redshift information about the background galaxy (termed the ‘source’). The angular diameter distances $D_{\text{os}}, D_{\text{ls}}, D_{\text{ol}}$ are measured between observer ‘o’, lens ‘l’ and/or source ‘s’. The definition of β is such that objects in front of the cluster, which are not gravitationally sheared,

do not contribute to the measured signal. For an increasing source redshift the lensing efficiency β rises sharply when the source redshift is comparable to the lens redshift, but it flattens off when source and lens are far apart.

We lack photometric redshifts for individual objects in our catalogue and we cannot determine the critical surface density for each source lens pair. However, as the galaxies are averaged to obtain a shear estimate, we can use the average lensing efficiency $\langle\beta\rangle$ to compute the critical surface density for the full source population. This assumption introduces a bias in our shear estimates which can be approximately corrected for by multiplying our reduced shear estimates by

$$1 + \left(\frac{\langle\beta^2\rangle}{\langle\beta\rangle^2} - 1 \right) \kappa, \quad (4.5)$$

(Equation 7 in Hoekstra et al. 2000). The width of the distribution of the lensing efficiency $\langle\beta^2\rangle$ corrects the reduced shear for the use of a single value of $\langle\beta\rangle$. For our local clusters most sources are so distant that there is little variation in the value of β . Indeed, we find that the ratio $\langle\beta^2\rangle/\langle\beta\rangle^2 \approx 1$ in the range $0.05 \leq z \leq 0.15$ and so the correction is negligible for our analysis. However, when we compute cluster properties we do correct our reduced shear estimates using Equation 4.5 for completeness.

A reference sample of field galaxies can serve as a proxy for the source population in the observations from which the average lensing efficiency can be computed. For this we use the Cosmological Evolution Survey (COSMOS) field which has received dedicated spectroscopic coverage so that reliable redshift estimates are available. In our analysis we use the COSMOS2015 catalogue of Laigle et al. (2016), which contains photometric redshifts of galaxies in the COSMOS field based on over 30 different wavelengths. This catalogue has two important benefits for our analysis. First, near infrared data from the UltraVISTA DR2 are included, so that the Lyman and Balmer/4000 Å breaks can be distinguished. The additional knowledge on these features helps to robustly assess high redshift galaxies, which would otherwise tend to be assigned a low redshift. Second, the catalogue also includes the CFHT r filter, so that we can easily match it to our data. Although the objects in the COSMOS2015 catalogue were not selected based on their r -band magnitude, we find that the catalogue is complete down to $m_r \approx 25$, which is sufficient to cover our magnitude range $20 \leq m_r \leq 24.5$. From comparisons to spectroscopic data Laigle et al. (2016) found that their redshift estimates are accurate to better than a percent, which is sufficient for this study.

However, the COSMOS2015 catalogue is not representative of our MENeACS lensing catalogues, as the latter are subject to various cuts. Gruen & Brimiouille (2016) have shown that these selection effects can introduce a bias in the mass estimates. To quantify the impact, we ran our lensing pipeline on r -band observations of the CFHT Legacy Survey (CFHTLS) D2 field which covers ~ 1 square degree of the COSMOS field and matched the lensing catalogue to the COSMOS2015 catalogue. This enabled us to perform exactly the same cuts on the redshift distribution as were applied to the lensing data. We found that applying the cuts introduces a difference in the lensing efficiency of only $\Delta\beta \leq 1\%$ for all clusters. We use the matched catalogue for our analysis, but note that this does not significantly impact our results, nor the results in H15.

We select galaxies from the matched catalogue using the TYPE parameter, which classifies objects as either stars or galaxies. The observed galaxy density as function of magnitude varies for the MENeACS clusters due to the different observing conditions,

which is not reflected in the source population in COSMOS. As a galaxy's magnitude depends on its redshift, using the COSMOS galaxies to compute $\langle\beta\rangle$ directly may lead to an incorrect value for our data. Therefore we customise our COSMOS galaxy population by reweighting them to match the magnitude distributions for individual MENeACS clusters. The redshift catalogue is divided into magnitude bins and for each magnitude bin we compute the sum of the lensing weights of the COSMOS galaxies (as the appropriate proxy for number of galaxies) and the mean lensing efficiency $\langle\beta\rangle_{\text{bin}}$. Then the final estimate $\langle\beta\rangle$ for each cluster is the average of $\langle\beta\rangle_{\text{bin}}$ weighted by the sum of the lensing weight for all bins. For each cluster the value of $\langle\beta\rangle$ is listed in column 9 of Table 4.1. We use $\langle\beta\rangle$ to compute the average critical surface density with which we estimate cluster masses. In order to apply Equation 4.5 we also require $\langle\beta^2\rangle$, which is calculated the same way and listed in column 10 of Table 4.1.

The redshift distribution in our catalogue based on 1 square degree of the COSMOS field might not be representative for all source populations in our observations. This cosmic variance introduces an uncertainty in the mean lensing efficiency $\langle\beta\rangle$. We estimate the impact of cosmic variance using the photometric redshift catalogues of Coupon et al. (2009) for the four CFHTLS DEEP fields. We also analysed these fields with our own weak lensing pipeline and matched these catalogues to introduce the lensing selections. These photometric redshifts are based on five optical bands and hence are not as reliable as the COSMOS2015 catalogue. However, as the four fields were analysed consistently they may serve as an estimate of the variation in redshift distributions due to cosmic variance. For each cluster we compute the weighted average $\langle\beta\rangle$ for the 4 fields and use the standard deviation between them as the error due to cosmic variance.

In addition to cosmic variance, there are Poisson errors in the $\langle\beta\rangle$ due to finite statistics. The Poisson errors are estimated by comparing the lensing efficiency in the CFHTLS D2 field with the lensing efficiency in the remainder of the COSMOS field, where we assume that within two square degrees cosmic variance is subdominant. We compare the lensing efficiency for galaxies between $20 \leq m_r \leq 24.5$ for both regions and use the difference as a measure of the Poisson error. As we do not have KSB parameters for the full COSMOS2015 catalogue we do not impose any other constraints besides the magnitude limits. We assume that the lensing cuts would affect both samples in the same way.

The uncertainties from cosmic variance and finite statistics are roughly equal in amplitude and we estimate our final uncertainty by summing both quadratically, assuming they are independent. The uncertainty in the average lensing efficiency for each cluster is listed in column 6 in Table 4.1 and is at most $\sim 1.5\%$. Strictly speaking, this is a conservative estimate, as our cosmic variance errors are also affected by Poisson errors. An error $\delta\beta \sim 1.5\%$ is a marked improvement over H15 who found an uncertainty of 4.2%. In part, this can be attributed to our use of better photometric redshift data. However, H15 found that most of their uncertainty was driven by faint galaxies and the shallower data for MENeACS help to reduce the error. Finally, for the low redshift MENeACS clusters the distribution of β is very peaked which reduces the uncertainty in the lensing efficiency.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	cluster	z	RA ^{BCG} (J2000)	Dec ^{BCG} (J2000)	A_r [mag]	R_{PSF} [pixels]	r.m.s. noise [counts]
1	A7	0.106	00:11:45.2	+32:24:56.4	0.086	1.907	1.279
2	A21	0.095	00:20:37.3	+28:39:28.3	0.083	2.005	1.325
3	A85	0.055	00:41:50.4	-09:18:11.3	0.082	1.964	1.447
4	A119	0.044	00:56:16.1	-01:15:19.1	0.084	2.041	1.403
5	A133	0.057	01:02:41.7	-21:52:54.4	0.040	2.148	1.405
6	A646	0.129	08:22:09.5	+47:05:52.9	0.089	2.171	1.381
7	A655	0.127	08:25:29.0	+47:08:00.4	0.076	2.056	1.389
8	A754	0.054	09:08:32.3	-09:37:47.4	0.150	2.338	1.326
9	A780	0.054	09:18:05.7	-12:05:45.2	0.086	2.542	1.470
10	A795	0.136	09:24:05.3	+14:10:21.8	0.062	2.282	1.514
11	A961	0.124	10:16:22.8	+33:38:17.3	0.039	2.245	1.536
12	A990	0.144	10:23:39.8	+49:08:37.8	0.014	2.468	1.482
13	A1033	0.126	10:31:44.2	+35:02:28.0	0.037	2.064	1.627
14	A1068	0.138	10:40:44.5	+39:57:11.2	0.046	1.937	1.582
15	A1132	0.136	10:58:23.6	+56:47:41.8	0.024	2.144	1.600
16	A1285	0.106	11:30:23.8	-14:34:52.8	0.090	2.587	1.588
17	A1348	0.119	11:41:24.2	-12:16:38.5	0.066	2.590	1.535
18	A1361	0.117	11:43:39.5	+46:21:20.4	0.050	1.926	1.401
19	A1413	0.143	11:55:18.0	+23:24:18.6	0.052	2.074	1.423
20	A1650	0.084	12:58:41.5	-01:45:40.9	0.036	2.413	1.316
21	A1651	0.085	12:59:22.4	-04:11:45.8	0.060	2.871	1.625
22	A1781	0.062	13:44:52.5	+29:46:15.3	0.035	2.308	1.199
23	A1795	0.062	13:48:52.5	+26:35:33.2	0.028	2.162	1.402
24	A1927	0.095	14:31:06.7	+25:38:01.3	0.084	1.953	1.284
25	A1991	0.059	14:54:31.4	+18:38:32.3	0.071	2.116	1.465
26	A2029	0.077	15:10:56.1	+05:44:41.0	0.083	2.058	1.330
27	A2033	0.082	15:11:26.5	+06:20:56.7	0.081	1.937	1.237
28	A2050	0.118	15:16:17.9	+00:05:20.8	0.119	1.953	1.296
29	A2055	0.102	15:18:45.7	+06:13:56.2	0.082	1.935	1.300
30	A2064	0.108	15:20:52.2	+48:39:38.4	0.036	2.196	1.129
31	A2065	0.073	15:22:29.1	+27:42:27.6	0.086	2.079	1.230
32	A2069	0.116	15:24:07.4	+29:53:20.2	0.053	1.945	1.152
33	A2142	0.091	15:58:20.0	+27:14:00.3	0.098	1.971	1.259
34	A2420	0.085	22:10:18.7	-12:10:13.7	0.127	2.130	1.638
35	A2426	0.098	22:14:31.5	-10:22:26.2	0.129	2.293	1.757
36	A2440	0.091	22:23:56.9	-01:34:59.7	0.174	2.200	1.449
37	A2443	0.108	22:26:07.8	+17:21:23.4	0.136	1.944	1.352
38	A2495	0.078	22:50:19.7	+10:54:14.1	0.167	1.938	1.366
39	A2597	0.085	23:25:19.7	-12:07:27.0	0.066	2.106	1.649
40	A2627	0.126	23:36:42.0	+23:55:29.0	0.168	2.025	1.360
41	A2670	0.076	23:54:13.6	-10:25:08.5	0.097	2.422	1.520
42	A2703	0.114	00:05:23.9	+16:13:09.2	0.103	1.881	1.326
43	MKW3S	0.045	15:21:51.8	+07:42:31.8	0.077	2.046	1.253
44	RXJ0132	0.149	01:32:41.1	-08:04:04.8	0.066	1.901	1.444
45	RXJ0736	0.118	07:36:38.0	+39:24:52.6	0.104	2.198	1.418
46	RXJ2344	0.079	23:44:18.2	-04:22:49.1	0.077	2.198	1.515
47	ZWCL1023	0.143	10:25:57.9	+12:41:08.4	0.101	2.274	1.528
48	ZWCL1215	0.075	12:17:41.1	+03:39:21.2	0.036	2.730	1.484

(1)	(2) cluster	(3) z	(9) $\langle\beta\rangle$	(10) $\langle\beta^2\rangle$	(11) $\delta\beta$
1	A7	0.106	0.714	0.536	0.008
2	A21	0.095	0.738	0.568	0.007
3	A85	0.055	0.841	0.718	0.004
4	A119	0.044	0.871	0.767	0.004
5	A133	0.057	0.834	0.708	0.005
6	A646	0.129	0.665	0.469	0.009
7	A655	0.127	0.669	0.475	0.008
8	A754	0.054	0.842	0.720	0.005
9	A780	0.054	0.841	0.718	0.005
10	A795	0.136	0.648	0.449	0.010
11	A961	0.124	0.675	0.481	0.008
12	A990	0.144	0.632	0.428	0.010
13	A1033	0.126	0.670	0.475	0.008
14	A1068	0.138	0.646	0.446	0.010
15	A1132	0.136	0.648	0.448	0.010
16	A1285	0.106	0.709	0.528	0.009
17	A1348	0.119	0.683	0.492	0.007
18	A1361	0.117	0.693	0.505	0.007
19	A1413	0.143	0.636	0.434	0.010
20	A1650	0.084	0.763	0.603	0.007
21	A1651	0.085	0.756	0.592	0.007
22	A1781	0.062	0.821	0.688	0.005
23	A1795	0.062	0.821	0.688	0.005
24	A1927	0.095	0.740	0.571	0.007
25	A1991	0.059	0.828	0.698	0.005
26	A2029	0.077	0.782	0.631	0.007
27	A2033	0.082	0.772	0.615	0.006
28	A2050	0.118	0.690	0.502	0.008
29	A2055	0.102	0.724	0.548	0.008
30	A2064	0.108	0.710	0.530	0.008
31	A2065	0.073	0.793	0.646	0.007
32	A2069	0.116	0.696	0.510	0.007
33	A2142	0.091	0.747	0.581	0.007
34	A2420	0.085	0.759	0.597	0.007
35	A2426	0.098	0.726	0.551	0.007
36	A2440	0.091	0.745	0.577	0.007
37	A2443	0.108	0.708	0.527	0.008
38	A2495	0.078	0.778	0.625	0.007
39	A2597	0.085	0.760	0.598	0.007
40	A2627	0.126	0.672	0.478	0.008
41	A2670	0.076	0.780	0.628	0.007
42	A2703	0.114	0.699	0.513	0.007
43	MKW3S	0.045	0.868	0.763	0.004
44	RXJ0132	0.149	0.622	0.419	0.010
45	RXJ0736	0.118	0.686	0.497	0.008
46	RXJ2344	0.079	0.775	0.620	0.007
47	ZWCL1023	0.143	0.633	0.430	0.010
48	ZWCL1215	0.075	0.783	0.631	0.007

Table 4.1: Basic information on the MENeaCS clusters, parameters governing the quality of our observations and the lensing efficiency β computed in Section 4.3. (2 & 3) cluster name and redshift; (4)&(5) coordinates of the BCG, which is taken to be the cluster centre; (6) Galactic extinction in r -band magnitude; (7) half-light radius of the PSF in pixels averaged over the entire image; (8) r.m.s. of the background noise in counts; (9) average β used to estimate the critical surface density; (10) average β^2 used to correct the shear for the lack of individual source redshifts; (11) error on β (see text for more details).

4.4 Contamination of the source population by cluster members

The galaxy catalogue from the lensing analysis contains both field galaxies and cluster members. Cluster members are not sheared by the gravitational potential of the cluster and keeping them in the sample will alter the shear signal. If cluster galaxies are not intrinsically aligned (as has been shown by Sifón et al. 2015), their presence dilutes the shear signal, biasing the shear estimate low, where the size of the bias depends on the relative overdensity of cluster members compared to background galaxies. At low redshift, cluster galaxies are spread out over the field of view, so their number density per unit solid angle is low.

With reliable redshifts for individual galaxies, cluster members can be identified and removed from the sample. However, we lack the multi-band observations to acquire reliable photometric estimates. Instead, we can apply a ‘boost correction’ to statistically correct for cluster member contamination (see appendix A1 of Leauthaud et al. (2016) for a discussion of the boost correction). This approach offsets the dilution of the shear by boosting the shear signal based on the fraction of cluster members to background galaxies. The application of the boost correction relies on the assumption that only cluster members affect the galaxy counts. We investigate the effects that violate this assumption in the next sections and take them into account to obtain a reliable estimate of the density of cluster members relative to the density of background galaxies, from which we compute the boost correction.

As noted in Section 4.2 close proximity to bright objects can affect the measured shapes of galaxies, changing the measured shear signal. We incorporate this effect by quantifying the boost correction in terms of the sum of the lensing weights per square arcminute, which we call the weight density ξ .

4.4.1 Magnification

Gravitational lensing near the cluster core magnifies the background sky. This phenomenon increases the observed flux of background galaxies, but it also reduces the actual area behind the cluster that is observed. These two features counteract each other in their effect on the observed number density of sources. The net effect depends on the number of galaxies scattered into the magnitude range that we designate for our lensing study. The observed number of galaxies increases with the magnification μ as $\mu^{2.5\alpha-1}$ (Mellier 1999). Hence, for a slope of the magnitude distribution $\alpha = d\log N_{\text{source}}/dm_{\text{shape}} = 0.40$ the net effect is negligible. For the MegaCam r -band data H15 computed that the slope is close to 0.40 and so we can safely ignore the effect of magnification on the source population.

4.4.2 Obscuration

Cluster members are large foreground objects and obscure part of the background sky, thereby reducing the number density of observed background galaxies. This phenomenon is especially important for MENeACS as the low redshift cluster members are very spread out. As the cluster member density increases towards the cluster core, obscuration reduces the apparent contamination signal as a function of projected cluster-centric distance. To address this issue we use the results of **Chapter 5**, where we used image simulations of the MENeACS clusters to compute the effect of obscuration. Their cluster image simulations were designed to mimic the observations as

closely as possible to accurately predict the effect of obscuration. For each simulated cluster image the seeing and noise level were set to the values measured in the data. Background galaxies were created with the image simulations pipeline of H15, which is based on the GALSIM software (Rowe et al. 2015), and cluster galaxies were added to the images. Where available, the GALFIT (Peng et al. 2002) measurements of Sifón et al. (2015) were used to create surface brightness profiles for cluster members, which were identified through spectroscopy or as part of the red sequence. For cluster members without (reliable) GALFIT measurements, galaxy properties were randomly sampled from parametric curves which follow the observed distributions of those properties. The analysis pipeline is run on both the background image and the cluster image producing two lensing catalogues. By matching these catalogues all background galaxies can be selected and the effect of cluster members on the weight density of the background population can be determined. The obscuration is defined as

$$f_{\text{obsc}} = 1 - \frac{\xi^{\text{cl}}}{\xi^{\text{bg}}}, \quad (4.6)$$

where ξ^{cl} and ξ^{bg} are the weight densities of all observed background galaxies in the cluster simulation and in the background simulation, respectively.

In Figure 4.1 we show the resulting obscuration in bins of projected cluster centric distance for individual clusters R in gray, and in black the average for all clusters. The effect of obscuration is greatest in the very lowest radial bins, which is expected because of the presence of the low redshift BCGs. At radii larger than 1 Mpc the obscuration flattens out but does not reach zero. We do not expect cluster members to obscure $\sim 5\%$ of all background galaxies in these outer regions. Instead, this plateau is caused by field galaxies entering the cluster member sample, as Sifón et al. (2015) showed that their sample of red sequence selected cluster members is contaminated at large radii. The simulated sample of cluster members lacks faint blue galaxies, but we expect that their obscuration is minimal and we thus ignore their contribution.

We determine an obscuration correction for the background weight density in the MENeaCS data by fitting a smooth function to the individual cluster obscuration profiles shown in gray in Figure 4.1. We find that the expression

$$f_{\text{obsc}}(R) = n_{\Delta} + n_0 \left(\frac{1}{R + R_c} - \frac{1}{R_{\text{max}} + R_c} \right), \quad (4.7)$$

worked well to describe the obscuration. We use the same expression in Section 4.4.4 to model the contamination and we discuss it in more detail there. On average $R_c \approx 0.04$ Mpc and $n_0 \approx 0.04$ for the best fits to the obscuration profiles. The parameter $R_{\text{max}} = 3$ Mpc and n_{Δ} was fit but then set to zero to renormalise the data such that f_{obsc} is consistent with zero beyond 1.5 Mpc. The best fits to the obscuration profiles to individual clusters were then used to correct the background galaxy counts in the MENeaCS data.

4.4.3 Excess galaxy weight density

Now that we have a correction for the decreased weight density due to obscuration, we can determine the excess weight density of all sources in the MENeaCS lensing catalogues relative to the weight density of background objects as a function of cluster-centric distance. This then provides the boost correction for the shear signal to correct for contamination of the source sample by cluster members.

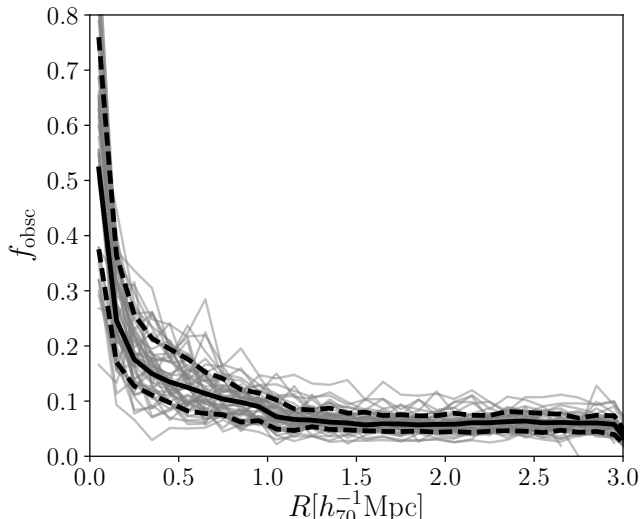


Figure 4.1: Fraction of source galaxies obscured by cluster members in realistic image simulations of MENeACS clusters as a function of radial distance to the BCG, which is assumed to be the cluster centre. Gray lines show the obscuration profile for individual clusters and the black solid (dashed) curve shows the average (1σ spread) for all clusters. The region of interest for our lensing analysis is between 0.5 Mpc and 2 Mpc, where obscuration is on average a $\lesssim 10\%$ effect.

The first step to compute the excess weight density is to determine the weight density of background objects. H15 computed that at 4 Mpc the structure associated to the cluster is a negligible contribution to the number density of field galaxies and used the area outside that 4 Mpc to estimate the galaxy density. The low redshift of the MENeACS sample means that the field of view does not encompass 4 Mpc for all clusters. Only the clusters at $z > 0.1$ have any significant area outside 3 Mpc in the mosaics with full exposure time and only the highest redshift clusters have sufficient area outside 3 Mpc for statistically meaningful estimates. To compensate for this lack of data, we use ancillary publicly available observations of blank fields to obtain an estimate of the weight density of field galaxies (as was also suggested by Schrabback et al. 2016). We analysed ~ 33 square degrees of deep CFHT data with our lensing pipeline and we derive a parametric model for the field galaxy weight density ξ^{field} in Appendix 4.A. The value of ξ^{field} is a function of the Galactic extinction, depth of the observations, and the seeing, and it predicts the mean density with an uncertainty of 1%. We use this model to predict the weight density of field galaxies for each cluster based on the seeing, noise level and the Galactic extinction in the observations.

In Figure 4.2 we show the excess weight density ξ/ξ^{field} (the obscuration corrected weight density normalised to the weight density of field galaxies), as a function of the distance to the BCG. Points with errorbars show the average excess weight density for all clusters and blue (red) shaded regions show the average excess weight density for clusters at $z < 0.1$ ($z \geq 0.1$). The contamination by cluster members is benign for the MENeACS clusters; the excess weight density is higher than 20% only within the inner 500 kpc. For the lensing analysis we only use sources beyond 500 kpc (see Section 4.5)

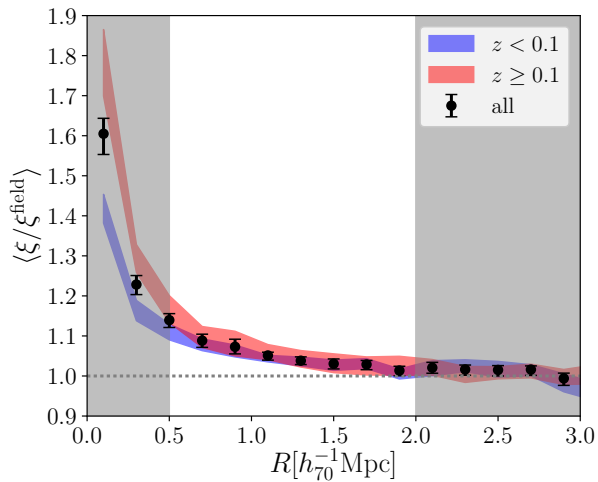


Figure 4.2: Excess weight densities of all sources in the magnitude range $20 \leq m_r \leq 24.5$ in the lensing catalogues as a function of radial distance to cluster centre, which we assume to be the position of the BCG. The excess weights are determined from the ratio of the weight density ξ corrected for obscuration and the average weight density for field galaxies. The blue (red) shaded area shows average excess weight density for all $z < 0.1$ ($z \geq 0.1$) clusters, black points with errorbars show the full MENeCS sample. The dotted line shows unity. The width of the coloured regions and the errorbars show the 1σ uncertainty in the excess weight density. The region shown in white between 0.5 Mpc and 2 Mpc is used for the lensing analysis, in which the contamination is on average $\sim 5\%$.

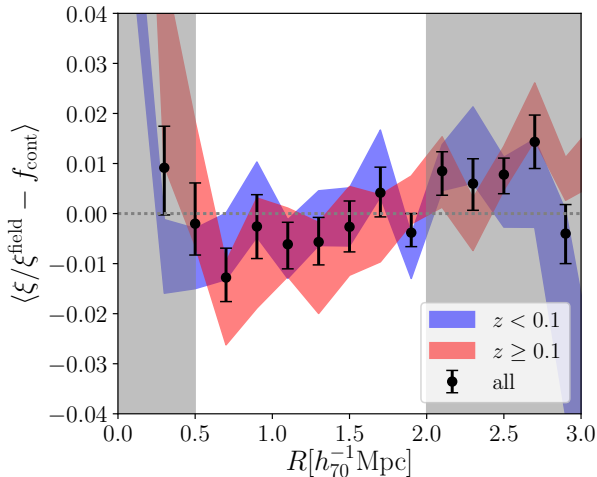


Figure 4.3: Average weight density of galaxies as a function of cluster-centric distance after the best fit model for contamination for each individual cluster has been subtracted. Blue (red) shaded area shows residual weight density for $z < 0.1$ ($z \geq 0.1$) clusters, points show the full MENeACS sample. The width of the coloured regions and the errorbars show the 1σ uncertainty. The region of interest for our lensing analysis is shown in white, where the residual contamination is consistent with zero.

and so the effect of contamination is small. The difference in the red and blue shaded regions clearly shows that this low level of contamination is thanks to the low redshift of the MENeACS clusters.

The weight density from the blank fields predicts the average weight density of field galaxies very well, but it may not be representative of specific environments of the MENeACS clusters. Specifically, we may miss large scale structure in the cluster background, which would have been properly accounted for if the density was normalised to the area outside the cluster. Therefore we checked the gain of combining the blank fields and any available area in the MENeACS data outside 3 Mpc from the BCG for the highest redshift clusters. We redefined the field galaxy weight density ξ^{field} as the inverse variance weighted average of the blank fields and the available uncontaminated MENeACS area. The spread of the blanks around the best fit is 6.4%, which is used as the square root of the variance for the blank fields. The variance for the uncontaminated data is estimated by subdividing the available area outside 3 Mpc into quadrants and taking the variance between the quadrants. The addition of the cluster data helps to reduce the uncertainty in the boost correction. However, in our analysis we are dominated by the uncertainty in the shear estimate due to the finite number of galaxies that are averaged. Hence we use the blank field prediction for ξ^{field} for all clusters to compute the excess weight density.

4.4.4 Boost correction

The excess weight density per cluster is a noisy measurement and using it directly to boost the shear signal can produce a spurious signal. Instead, we assume that the density of cluster members is a smooth function of the cluster-centric radius. This

assumption will not be valid if the cluster has local substructure, but any additional uncertainty this introduces, will average out for the full ensemble of clusters. A reasonable and simple approximation for the density of cluster members is the singular isothermal sphere (SIS; see also § 4.5.1). H15 found that the projected singular isothermal sphere was not able to describe the excess density profile to sufficient accuracy and introduced a function with additional parameters, which is shown in Equation 4.7. In this case, n_0 describes the amplitude of the contamination and the cluster core radius R_c is fitted for each cluster individually. The maximum radius $R_{\max} = 3$ Mpc is the limit beyond which the function is set to n_Δ . In Figure 4.2 the excess weight density already vanishes beyond 2 Mpc, so setting $R_{\max} = 3$ Mpc is reasonable for *MENeaCS*. All CCCP clusters were small enough in angular coordinates so that H15 could set $n_\Delta = 1$. Our prediction for field galaxy weight density has an intrinsic scatter and so we do not expect the excess weight density for individual clusters to converge to 1 at large radii. Therefore we add n_Δ as a free parameter in our analysis. We find that the relative spread in n_Δ is 7.2%, which is in agreement with the 6.4% scatter expected from the blank fields.

The ensemble averaged residual, after subtracting the best fit profile for each cluster, is shown in Figure 4.3. Again, we separate the sample in low redshift ($z < 0.1$, blue) and high redshift ($z \geq 0.1$, red) clusters and the full sample is denoted by the black points. For most radii the average residual is consistent with zero within the errors, regardless of the mean redshift of the sample. This shows that Equation 4.6 is a decent description of the density of cluster members. At $R \approx 3$ Mpc the available area for $z < 0.1$ clusters is decreasing which greatly increases the errorbars and the crowded cluster centre is not accurately described by the fitting function. However, for the lensing analysis we restrict ourselves to 0.5 - 2 Mpc for which the residual is consistent with zero with an uncertainty of $\sim 1\%$. The best fit profiles will serve as a boost correction for the shear signal of clusters to statistically correct for contamination of the source population by unlensed cluster members.

4.5 *MENeaCS* cluster masses

In the previous sections we have computed the corrections for the lack of individual redshift estimates for the source galaxies and the presence of cluster members in the source sample. We now apply these corrections to the measured tangential reduced shear and use the resulting shear as a function of cluster-centric distance to estimate the weak lensing masses.

To compute cluster masses we fit parametric models to our estimated shear profile. These symmetric models are not a perfect description of full dark matter haloes and may be poor fits to individual clusters. In fact numerical simulations have shown that fitting models can lead to masses underestimated by $\sim 5\%$ - 10% (Becker & Kravtsov 2011; Schrabback et al. 2016; Henson et al. 2017). To mitigate potential biases we restrict the scales at which we fit our models to 0.5 - 2 Mpc scales, where the models are reasonable descriptions of dark matter distributions. At large radii the density models go to zero whereas large scale structure surrounding the cluster contributes to the shear signal in the real Universe. Estimated masses will be biased due to this mismatch, unless the fit is restricted to intermediate scales (Corless & King 2007; Becker & Kravtsov 2011; Rasia et al. 2012). At small radii the shears can become very large for which our shape measurement algorithm was not calibrated, so we also exclude these scales. By excluding the cluster core we also limit the impact of miscentring in

our mass estimates (see Section 4.5.4).

Uncorrelated structures along the line of sight do not bias the lensing signal (Hoekstra 2001), but introduce an additional uncertainty that cannot be distinguished from the cluster signal. We use the predictions from Hoekstra et al. (2011a) to incorporate the effect of distant large scale structure into the errorbars on our weak lensing masses.

4.5.1 Singular isothermal sphere

The simplest profile to fit to the tangential shear profile is the projected singular isothermal sphere (SIS)

$$\gamma_t(R) = \frac{R_E}{2R}. \quad (4.8)$$

The SIS profile is useful to consider because it can be directly linked to dynamical estimates of the cluster. The Einstein radius R_E in radians is related to the velocity dispersion of the cluster projected along the line of sight σ_v , assuming isotropic orbits and spherical symmetry,

$$R_E = 4\pi\langle\beta\rangle\frac{\sigma_v^2}{c^2}, \quad (4.9)$$

where c is the speed of light. For the MENeACS clusters Sifón et al. (2015) compiled a catalogue of cluster members and computed the velocity dispersions from spectroscopic redshifts. We fit the SIS model to our weak lensing data to compare with the dynamical velocity dispersions.

The convergence κ for a SIS is given by the same expression as the shear and we fit the reduced shear $g = \gamma/(1 - \kappa)$ (corrected for using Equation 4.5) to our measured reduced shears. The best fit Einstein radii and resulting velocity dispersions are listed for all clusters in Table 4.2. For the radial range 0.5-2 Mpc the convergence is negligible and we note that instead fitting the reduced shear directly with Equation 4.8 would alter the Einstein radii by only 0.5%.

The comparison between the velocity dispersions from the SIS fit to the dynamical velocity dispersions from Sifón et al. (2015) is shown in Figure 4.4. Black points show results for the MENeACS clusters and gray points show CCCP clusters from H15. The MENeACS clusters have on average lower velocity dispersions (and thus lower masses according to the virial theorem) than the CCCP clusters. The volume from which the local MENeACS clusters are drawn is small compared to the volume for the CCCP clusters. As the mass function decreases rapidly for high masses, and both surveys target the most massive clusters in their respective volumes, we expect on average lower masses for MENeACS clusters.

To fit the data points with errorbars on both axes, we represent each data point as an elliptical Gaussian distribution where the width is set by the errorbars and we randomly sample from the distribution to recreate our measurements many times. For each mock measurement we fit a straight line with only a slope and no offset. The analysis shows that there is decent agreement between the velocity dispersion estimates for the MENeACS clusters. On average the weak lensing estimates are roughly 2σ higher than the dynamical estimates, where σ was determined from the 16th and 84th percentiles of the distribution. The full sample of MENeACS and CCCP are consistent with a one-to-one relation. We perform a formal Kolmogorov-Smirnov test on the difference between the two velocity dispersion estimates for the two clusters samples. The p -value is 0.18, so we can be confident that the two samples are drawn from the same underlying distribution.

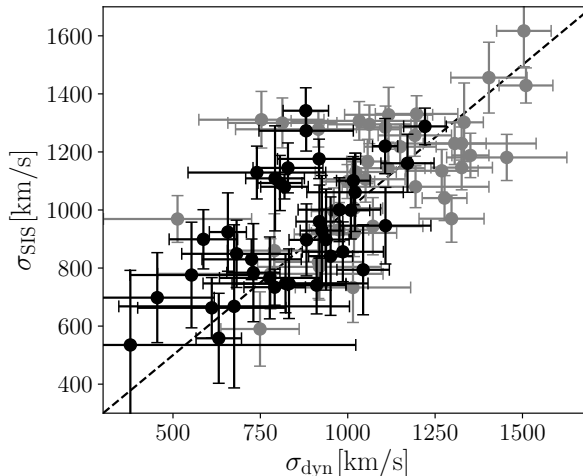


Figure 4.4: Velocity dispersion σ from SIS fit to the weak lensing against the dynamical velocity dispersion taken from Sifón et al. (2015). Black points show our results and grey points show the same for clusters in the CCCP sample of H15. The dashed line shows a one to one relation.

4.5.2 Navarro-Frenk-White profile

An often used profile to describe dark matter haloes is the Navarro-Frenk-White (NFW) profile, which, unlike the SIS profile, is known to be a good fit to observational data (e.g. Okabe et al. 2013; Umetsu et al. 2014; Viola et al. 2015). In numerical simulations Navarro et al. (1997) found a universal profile for the density of dark matter haloes

$$\frac{\rho(r)}{\rho_0} = \frac{\delta_c(\Delta)}{(r/r_s)(1+r/r_s)^2}, \quad (4.10)$$

where the radial shape of the profile is defined by the scale radius r_s ³. The amplitude of the profile is set by the characteristic overdensity

$$\delta_c(\Delta) = \frac{\Delta}{3} \frac{c_\Delta^3}{\ln(1+c_\Delta) + c_\Delta/(1+c_\Delta)}, \quad (4.11)$$

which depends on the concentration c_Δ . For a fixed number Δ the concentration is the ratio of the radius r_Δ enclosing a sphere of density $\Delta\rho_0$ and the scale radius: $c_\Delta = r_\Delta/r_s$. The mass within this region can be obtained from:

$$M_\Delta = \Delta\rho_0 \frac{4\pi}{3} r_\Delta^3. \quad (4.12)$$

The density ρ_0 is usually set to the critical density of the Universe $\rho_{\text{crit}} = 3H(z)^2/8\pi G$ or the mean density of the Universe $\rho_{\text{mean}} = \Omega_m\rho_{\text{crit}}$ in combination with some characteristic value for Δ . For our analysis we use the mean density $\rho_0 = \rho_{\text{mean}}$ in Equations 4.10 and 4.12 and the overdensity factor $\Delta_{\text{vir}} = (18\pi^2 + 82\Omega_\Lambda(z) - 39\Omega_\Lambda(z)^2)/\Omega_m(z)$, where

³Here the scale radii r and r_s are three dimensional quantities in contrast to the two dimensional Einstein radius R_E and cluster-centric radius R .

$\Omega_\Lambda(z)$ and $\Omega_m(z)$ are the energy densities of dark energy and matter at redshift z , respectively (e.g. Coe 2010). The resulting mass estimate from the NFW profile is the virial mass M_{vir} .

We follow the definitions in Wright & Brainerd (2000) to fit a projected NFW profile to our lensing signal. We combine their expressions for γ and κ to create an NFW profile for the tangential reduced shear, again with the additional terms given in Equation 4.5. The free parameters in the NFW model are correlated and the concentration depends on redshift. In practice, the concentration is constrained using numerical dark matter simulations (e.g. Duffy et al. 2008; Dutton & Macciò 2014; Diemer & Kravtsov 2015). We follow H15 and use the mass concentration relation found by Dutton & Macciò (2014), which is in good agreement with later work by Diemer & Kravtsov (2015) for our low redshift clusters. With the addition of the mass-concentration relation, our fitting function only has the virial mass as a free parameter. We fit our corrected reduced shear signals with an NFW model and list the best fit virial masses and corresponding virial radii r_{vir} in Table 4.2. We find that taking into account the uncertainty in the masses due to distant large scale structure inflates the errorbars by $\sim 20\%$.

In Figure 4.5 we show the inverse variance weighted average shear signal multiplied by the critical surface density to account for the different lensing efficiencies of the clusters. Each individual reduced shear profile is converted to shear γ using the individual best fit mass estimates to compute κ . The best fit NFW profiles are also shown and are visually a good representation of the data. The average NFW virial mass is $8.26 \pm 0.40 \times 10^{14} M_\odot$ and is shown as the red point in the inset. If we leave the concentration as a free fit parameter we obtain $7.93 \pm 0.60 \times 10^{14} M_\odot$ (blue contours). Our data shows an average mass and concentration consistent with the results we get when using the results of Dutton & Macciò (2014).

It is instructive to compare our best fit NFW masses other available mass estimates. Rines et al. (2016) have used spectroscopic redshifts to identify caustics, which can be related to the escape velocity in the cluster potential. They provide M_{200} dynamical masses for 24 MENeaCS clusters. We convert our NFW virial masses to M_{200} using the expressions given in Hu & Kravtsov (2003)⁴. For M_{200} this amounts to equating $\rho_{\text{mean}} \delta_c(\Delta_{\text{vir}}) = \rho_{\text{crit}} \delta_c(200)$ and solving for c_{200} . The resulting M_{200} masses are listed in Table 4.2 and compared to the dynamical estimates in Figure 4.6. The 24 MENeaCS clusters are shown in black and 18 CCCP clusters are shown in gray. Simple linear regression shows that the weak lensing masses are higher than the dynamical masses by a factor of $\sim 20\%$. This discrepancy is consistent for both the MENeaCS and the CCCP sample. H15 discussed that the discrepancy could be reduced (but not removed) by excluding outliers, which were also commented upon by Rines et al. (2013). The bulk of the MENeaCS clusters have consistently higher weak lensing mass compared to the dynamical mass, making it difficult to explain the difference based on individual clusters. Moreover, we could not find discussions on the dynamical masses or dynamical states for all clusters with a large difference between the masses estimates. One of these outlying clusters, Abell 2243, is a known merging system, so that the dynamical mass might be poorly constrained. There is no satisfactory explanation for the discrepancy of the mass estimates, but we do find that the overall discrepancy is consistent for both the MENeaCS and the CCCP clusters.

⁴The conversion is not strictly necessary here as Dutton & Macciò (2014) derive a relation for $c_{200}(M_{200})$, so that it is possible to fit the NFW profile directly for M_{200} . However, the same cannot be done for M_{500} , which we shall use in Section 4.5.3, so we employ the conversion already here.

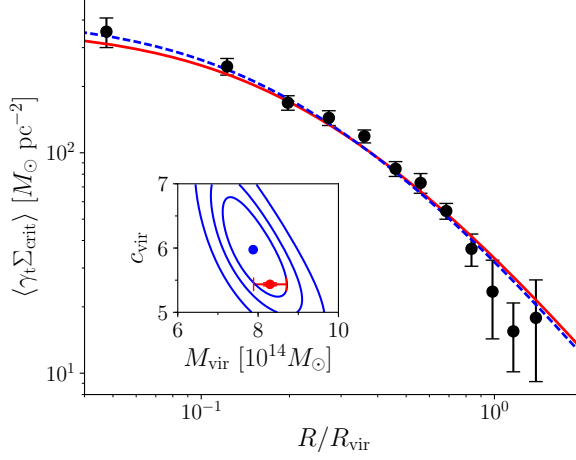


Figure 4.5: Stacked profile of the shear times critical surface density (also known as the excess surface density profile) as a function of the virial radius. Errorbars were computed using a sufficiently large number of bootstraps for each bin. The red solid line shows the best fit NFW profile using the Dutton & Macciò (2014) mass-concentration relation, and the blue dashed line shows the best fit, when leaving the concentration as a free fit parameter. The inset shows for the same colours the best fit mass and concentration and contours show the 68%, 90%, and 99% confidence intervals.

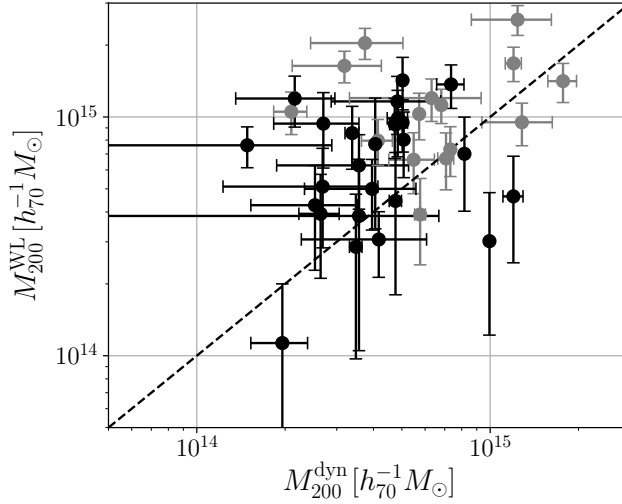


Figure 4.6: Comparison of the weak lensing masses M_{200}^{WL} and the dynamical masses M_{200}^{dyn} from Rines et al. (2016). Black points show our results and gray points show the results for CCCP clusters from H15. The dashed line shows unity slope.

4.5.3 Hydrostatic mass comparison

The *Planck* all-sky survey has produced a large catalogue of clusters detected through the SZ effect (Planck Collaboration et al. 2016c). Planck Collaboration et al. (2016b) used 439 clusters to constrain cosmological parameters by measuring the cluster mass function at different epochs. Cluster masses were computed using a scaling relation between the hydrostatic X-ray mass and the SZ observable Y_{SZ} (integrated Compton y-profile) based on a universal pressure profile (Arnaud et al. 2010). X-ray mass estimates can be biased because the underlying assumption of hydrostatic equilibrium is violated in galaxy clusters by bulk gas motions and non-thermal pressure support (e.g. Rasia et al. 2012), or due to uncertainties related to the calibration of X-ray observables (Mahdavi et al. 2013), and possibly by the assumption of a pressure profile. Planck Collaboration et al. (2016b) find that a bias $M_{\text{Planck}}/M_{\text{true}} \equiv 1 - b = 0.58 \pm 0.04$ is required to attain consistency between cosmological parameter constraints obtained with the cluster mass function and those obtained using primary CMB measurements (Planck Collaboration et al. 2016a). Such a low bias is not fully supported by independent mass measurements. von der Linden et al. (2014b) find a bias $1 - b = 0.69 \pm 0.07$, which is marginally consistent, but H15 find a higher value $1 - b = 0.76 \pm 0.05(\text{stat}) \pm 0.06(\text{syst})$, and Smith et al. (2016) find that b is consistent with zero. However, Battaglia et al. (2016) showed that adding a correction for Eddington bias would bring the results of WtG and CCCP more in line with the required *Planck* value.

The *Planck* cluster masses are based on the Compton y-profile integrated out to a radius of r_{500} . Therefore we have computed M_{500} estimates based on the NFW profile for a direct comparison (the values are listed in column 10 in Table 4.2). In Figure 4.7 our mass estimates M_{500}^{WL} are shown as a function of the *Planck* masses M_{500}^{SZ} in black for 30 clusters in MENeCS in common with the *Planck* catalogue and in gray we show the same for 37 CCCP clusters.

We fit a linear scaling relation between the two mass estimates accounting for errors in both mass estimates and allowing for intrinsic scatter in the hydrostatic masses. For MENeCS we find $M_{500}^{\text{WL}} = (1.03 \pm 0.08)M_{500}^{\text{SZ}}$ with an intrinsic scatter of $29 \pm 10\%$. This value is notably different from the scaling relation $0.76 \pm 0.05(\text{stat}) \pm 0.06(\text{syst})$ found by H15. If we repeat our analysis for the CCCP clusters using the computed NFW masses and the updated *Planck* masses (Planck Collaboration et al. 2016c), we find a bias of 0.83 ± 0.06 , which is still inconsistent with the results for MENeCS. A similar difference in scaling relations between samples of clusters at different redshifts was already noted by Smith et al. (2016). They suggested a redshift dependent hydrostatic mass bias, possibly arising due to systematic errors in weak lensing measurements or departures from self-similar evolution (Andreon 2014).

Alternatively, von der Linden et al. (2014b) advocated a mass dependence in the hydrostatic bias, which was also seen in other observations by H15 and Mantz et al. (2016) (who used the WtG weak lensing measurements) and hinted at in numerical simulations of clusters. Henson et al. (2017) find that the bias in X-ray masses increases from 20% to 40% with the true cluster mass for clusters in a mass range from $\sim 10^{14} - 10^{15} h_{70}^{-1} M_{\odot}$. Nagai et al. (2007) find lower hydrostatic biases of 5 – 20% for a relatively low mass sample with $\langle M_{500} \rangle = 4.14 h_{70}^{-1} M_{\odot}$, whereas Rasia et al. (2012) examine a sample of more massive clusters and find a higher range of biases around $\sim 25 - 35\%$. As the CCCP clusters are more massive than the MENeCS clusters, such a mass dependent hydrostatic mass bias could explain the difference in the scaling relations.

We take advantage of the large range in masses in the combined sample of CCCP

and MENeACS to investigate a potential mass dependence of the hydrostatic bias. For consistency we use the masses from NFW fits for the CCCP clusters, but H15 showed that the mass estimates are consistent. We fit the combined sample with a linear relation, finding $M_{500}^{\text{WL}} = (0.90 \pm 0.05)M_{500}^{\text{SZ}}$ and an intrinsic scatter of $34 \pm 5\%$. To allow for a mass dependence we fit a power law and find

$$\frac{M_{500}^{\text{SZ}}}{10^{15}M_{\odot}} = (0.84 \pm 0.07) \left(\frac{M_{500}^{\text{WL}}}{10^{15}M_{\odot}} \right)^{(0.82 \pm 0.08)} \quad (4.13)$$

and an intrinsic scatter of $29 \pm 6\%$. Both lines are shown in Figure 4.7. The power law fit clearly favors a power different than one, indicating that the data prefer mass evolution. This is also backed up by the slightly lower intrinsic scatter found for the power law. Both the amplitude 0.84 ± 0.07 and the slope of the power law 0.82 ± 0.08 are consistent with the results from H15: 0.76 ± 0.04 and 0.64 ± 0.17 , respectively. The slope is also consistent with the WtG results of $0.68^{+0.15}_{-0.11}$ (von der Linden et al. 2014b) and 0.73 ± 0.02 (Mantz et al. 2016), but, as for CCCP, this is mainly due to the large errorbars.

A potential caveat in our analysis is the use of NFW profiles. Although the NFW profile is a good description for a stack of clusters, simulations suggest they are biased low (e.g. Bahé et al. 2012). Gravitational lensing measures the density contrast and the NFW profile might be biased for merging systems (see e.g. Hoekstra et al. 2000 for how substructure affects the mass estimates). We find several known merging systems as outliers in the population and omitting them in the analysis reduces the slope and the power for the power law by $\sim 0.5\sigma$. An alternative approach would be to use the aperture mass (Fahlman et al. 1994; Clowe et al. 1998), which estimates the mass inside an aperture and should be more robust against the state of the matter distribution. H15 showed that aperture masses were consistent with NFW masses and in future work we shall verify this for the MENeACS clusters. In addition, a more careful analysis should incorporate the effect of Eddington bias, the importance of which has been shown by Battaglia et al. (2016). At low masses, not all MENeACS have measurements of M_{500}^{SZ} and such selection effects could be mistaken for a mass dependence of the hydrostatic mass bias.

4.5.4 Systematic error budget

A large part of this work has been devoted to corrections for systematic effects. Here we review their impact on our mass estimates.

In our analysis we have assumed that the centre of the cluster is given by the location of the BCG. If the BCG is not in the bottom of the gravitational potential the NFW profile is miscentred and the mass estimates will be biased. However, the red line in Figure 4 in Hoekstra et al. (2011b) shows that for our conservative choice of 0.5 Mpc as the lower limit of the fit range the bias is only $\sim 5\%$ if the BCG is 100 kpc from the true cluster centre. Mahdavi et al. (2013) and Bildfell (2013) looked at the distance between the BCG location and the peak in the X-ray surface brightness, and for small distances they are a good indicator for the centre of the gravitational potential of the cluster (George et al. 2012). They found that most of the CCCP clusters have a BCG offset smaller than 100 kpc. If we assume a similar distribution for the MENeACS clusters, we expect a bias of much less than 5% in our average cluster mass.

The uncertainty in the shear estimates for our pipeline was tested by H15 and they found an accuracy of $\sim 1\%$. They conservatively assign a 2% uncertainty in their

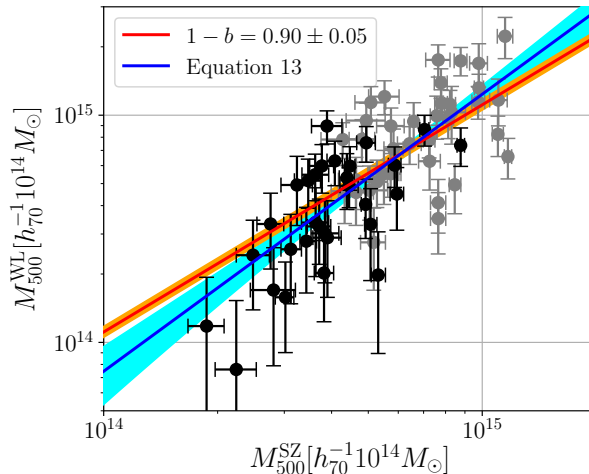


Figure 4.7: Comparison of the weak lensing masses M_{500} and the SZ masses from *Planck*. Black points show our results and gray points show the same for clusters in the CCCP sample of H15. The red line shows the best fit scaling relation using a constant hydrostatic mass bias and the scatter is shown as the orange band. The blue line and cyan shaded region show the same for the best power law fit.

analysis and we do the same. Thanks to a new high-fidelity photometric redshift catalogue, the uncertainty in our source redshift distribution is $\lesssim 1.5\%$, much better than the 4.2% uncertainty found for CCCP. The boost corrections applied to our tangential shear profiles are accurate to $\sim 1\%$. We find that not applying the boost corrections, contamination of cluster members in the source sample would reduce our mass estimates by $\sim 10\%$. This highlights the benefit of MENeACS as H15 found for the same fit range a $\sim 20\%$ effect for the higher redshift CCCP clusters. These sources of errors are uncorrelated and we add them quadratically to find a systematic error of 2.7%. The conservatively estimated 2% uncertainty in the shear measurement is the dominant source of error. A more cavalier view would state that the uncertainties in the photometric redshifts and the shear measurement are comparable.

The overall error in our analysis of $\sim 3\%$ is remarkably low, despite the lack of reliable photometric redshifts for individual galaxies. This deficiency is less severe for the low redshift MENeACS clusters and trading off multi-wavelength information against number of observed clusters has proven worth-while. We note that if the MENeACS observations had been deeper, the available photometric redshift catalogues would not have sufficient depth to cover our source population. For weak lensing analyses of clusters at much higher redshifts multi-wavelength observations will be necessary.

4.6 Conclusions

Cluster counts have the potential to put tight constraints on cosmological parameters, if large numbers of clusters with accurate mass estimates are observed. The large survey area of cosmic microwave background experiments provides a large sample

(1)	(2)	(3)	(4)	(5)	(6)
	cluster	R_E [arcsec]	σ_{SIS} [km/s]	$r_{\text{vir}} [h_{70}^{-1} \text{ Mpc}]$	$M_{\text{vir}} [10^{14} h_{70}^{-1} M_{\odot}]$
1	A7	19.0 ± 3.6	959 ± 92	2.86 ± 0.28	7.97 ± 2.34
2	A21	13.7 ± 2.5	803 ± 73	2.44 ± 0.22	4.85 ± 1.29
3	A85	24.6 ± 4.9	1006 ± 100	3.20 ± 0.33	10.17 ± 3.17
4	A119	20.4 ± 5.7	901 ± 126	2.82 ± 0.42	6.87 ± 3.04
5	A133	13.1 ± 2.2	737 ± 63	2.30 ± 0.19	3.78 ± 0.96
6	A646	12.0 ± 5.2	791 ± 172	2.33 ± 0.50	4.48 ± 2.89
7	A655	16.0 ± 6.2	909 ± 177	2.71 ± 0.52	7.05 ± 4.06
8	A754	30.0 ± 4.5	1111 ± 84	3.57 ± 0.27	14.07 ± 3.16
9	A780	13.6 ± 3.7	749 ± 103	2.33 ± 0.31	3.92 ± 1.56
10	A795	35.0 ± 4.3	1367 ± 85	4.02 ± 0.27	23.21 ± 4.61
11	A961	25.6 ± 4.2	1146 ± 95	3.40 ± 0.29	13.81 ± 3.52
12	A990	30.6 ± 3.6	1295 ± 76	3.76 ± 0.24	19.36 ± 3.68
13	A1033	24.3 ± 8.2	1121 ± 190	3.28 ± 0.59	12.39 ± 6.67
14	A1068	16.9 ± 4.9	952 ± 138	2.74 ± 0.41	7.37 ± 3.30
15	A1132	26.7 ± 3.2	1194 ± 72	3.49 ± 0.23	15.27 ± 3.01
16	A1285	21.0 ± 4.6	1013 ± 112	3.08 ± 0.34	9.89 ± 3.33
17	A1348	11.5 ± 3.8	763 ± 125	2.27 ± 0.35	4.04 ± 1.87
18	A1361	16.4 ± 3.8	906 ± 105	2.68 ± 0.31	6.69 ± 2.35
19	A1413	21.3 ± 5.6	1076 ± 140	3.14 ± 0.42	11.20 ± 4.47
20	A1650	26.2 ± 2.1	1091 ± 43	3.35 ± 0.14	12.33 ± 1.54
21	A1651	19.2 ± 7.9	939 ± 193	2.92 ± 0.59	8.16 ± 4.98
22	A1781	6.8 ± 6.6	536 ± 259	1.74 ± 0.73	1.65 ± 2.08
23	A1795	31.7 ± 4.9	1156 ± 89	3.65 ± 0.29	15.27 ± 3.66
24	A1927	14.9 ± 2.6	836 ± 71	2.54 ± 0.22	5.43 ± 1.39
25	A1991	14.5 ± 6.9	779 ± 184	2.43 ± 0.57	4.49 ± 3.17
26	A2029	38.6 ± 3.8	1308 ± 64	4.12 ± 0.20	22.56 ± 3.23
27	A2033	12.3 ± 3.3	744 ± 101	2.27 ± 0.30	3.82 ± 1.49
28	A2050	14.9 ± 3.6	865 ± 105	2.58 ± 0.31	5.95 ± 2.16
29	A2055	11.8 ± 3.8	752 ± 122	2.28 ± 0.37	3.99 ± 1.92
30	A2064	9.1 ± 7.8	668 ± 286	2.02 ± 0.79	2.83 ± 3.33
31	A2065	31.6 ± 5.5	1175 ± 102	3.69 ± 0.33	16.13 ± 4.39
32	A2069	12.9 ± 5.1	800 ± 157	2.40 ± 0.46	4.75 ± 2.74
33	A2142	33.2 ± 5.2	1240 ± 96	3.86 ± 0.29	19.00 ± 4.34
34	A2420	27.0 ± 6.7	1111 ± 136	3.45 ± 0.43	13.45 ± 5.04
35	A2426	20.1 ± 7.4	979 ± 181	2.92 ± 0.56	8.32 ± 4.82
36	A2440	26.3 ± 4.7	1107 ± 98	3.40 ± 0.30	13.01 ± 3.48
37	A2443	28.4 ± 5.7	1178 ± 118	3.55 ± 0.37	15.20 ± 4.80
38	A2495	7.0 ± 3.9	560 ± 156	1.75 ± 0.44	1.75 ± 1.32
39	A2597	16.1 ± 4.4	855 ± 118	2.62 ± 0.36	5.87 ± 2.43
40	A2627	9.7 ± 2.8	706 ± 101	2.10 ± 0.30	3.24 ± 1.40
41	A2670	21.0 ± 5.7	965 ± 130	2.94 ± 0.41	8.20 ± 3.45
42	A2703	17.6 ± 5.2	933 ± 138	2.78 ± 0.42	7.43 ± 3.33
43	MKW3S	11.1 ± 3.5	665 ± 105	2.09 ± 0.32	2.77 ± 1.27
44	RXJ0132	7.9 ± 3.4	663 ± 142	1.95 ± 0.40	2.71 ± 1.66
45	RXJ0736	9.8 ± 4.4	703 ± 157	2.10 ± 0.46	3.23 ± 2.11
46	RXJ2344	17.2 ± 5.0	878 ± 126	2.68 ± 0.39	6.27 ± 2.74
47	ZWCL1023	10.9 ± 4.0	772 ± 143	2.26 ± 0.41	4.17 ± 2.25
48	ZWCL1215	16.3 ± 7.9	848 ± 207	2.67 ± 0.63	6.13 ± 4.37

(1)	(2)	(7)	(8)	(9)	(10)
	cluster	$r_{200} [h_{70}^{-1} \text{ Mpc}]$	$M_{200} [h_{70}^{-1} \text{ Mpc}]$	$r_{500} [h_{70}^{-1} \text{ Mpc}]$	$M_{500} [h_{70}^{-1} \text{ Mpc}]$
1	A7	1.58 ± 0.16	5.00 ± 1.64	1.01 ± 0.10	3.24 ± 1.07
2	A21	1.35 ± 0.13	3.07 ± 0.94	0.86 ± 0.08	2.02 ± 0.62
3	A85	1.74 ± 0.19	6.27 ± 2.16	1.11 ± 0.12	4.05 ± 1.40
4	A119	1.53 ± 0.23	4.27 ± 1.99	0.98 ± 0.15	2.79 ± 1.30
5	A133	1.26 ± 0.11	2.39 ± 0.70	0.81 ± 0.07	1.58 ± 0.46
6	A646	1.30 ± 0.28	2.86 ± 1.89	0.84 ± 0.18	1.88 ± 1.24
7	A655	1.51 ± 0.30	4.45 ± 2.65	0.97 ± 0.19	2.89 ± 1.72
8	A754	1.93 ± 0.16	8.61 ± 2.31	1.23 ± 0.10	5.52 ± 1.48
9	A780	1.27 ± 0.17	2.48 ± 1.05	0.82 ± 0.11	1.64 ± 0.70
10	A795	2.22 ± 0.17	14.24 ± 3.53	1.40 ± 0.11	8.96 ± 2.22
11	A961	1.88 ± 0.17	8.56 ± 2.52	1.19 ± 0.11	5.47 ± 1.61
12	A990	2.09 ± 0.15	11.95 ± 2.88	1.32 ± 0.10	7.56 ± 1.82
13	A1033	1.82 ± 0.33	7.70 ± 4.30	1.15 ± 0.21	4.93 ± 2.76
14	A1068	1.53 ± 0.23	4.65 ± 2.20	0.98 ± 0.15	3.02 ± 1.42
15	A1132	1.94 ± 0.14	9.47 ± 2.34	1.23 ± 0.09	6.03 ± 1.49
16	A1285	1.70 ± 0.20	6.16 ± 2.27	1.08 ± 0.13	3.98 ± 1.46
17	A1348	1.26 ± 0.20	2.58 ± 1.25	0.81 ± 0.13	1.70 ± 0.83
18	A1361	1.49 ± 0.18	4.22 ± 1.61	0.95 ± 0.12	2.75 ± 1.05
19	A1413	1.75 ± 0.24	7.01 ± 2.98	1.11 ± 0.15	4.50 ± 1.91
20	A1650	1.83 ± 0.10	7.61 ± 1.48	1.17 ± 0.06	4.89 ± 0.95
21	A1651	1.60 ± 0.33	5.09 ± 3.20	1.02 ± 0.21	3.31 ± 2.07
22	A1781	0.96 ± 0.41	1.06 ± 1.35	0.62 ± 0.26	0.72 ± 0.91
23	A1795	1.98 ± 0.17	9.33 ± 2.63	1.26 ± 0.11	5.97 ± 1.68
24	A1927	1.40 ± 0.13	3.43 ± 1.01	0.90 ± 0.08	2.25 ± 0.66
25	A1991	1.33 ± 0.32	2.83 ± 2.04	0.85 ± 0.20	1.87 ± 1.35
26	A2029	2.24 ± 0.13	13.69 ± 2.83	1.41 ± 0.08	8.65 ± 1.79
27	A2033	1.25 ± 0.17	2.43 ± 1.01	0.81 ± 0.11	1.60 ± 0.67
28	A2050	1.43 ± 0.18	3.76 ± 1.48	0.92 ± 0.12	2.46 ± 0.96
29	A2055	1.27 ± 0.21	2.54 ± 1.28	0.81 ± 0.13	1.68 ± 0.84
30	A2064	1.13 ± 0.44	1.82 ± 2.15	0.73 ± 0.29	1.21 ± 1.43
31	A2065	2.01 ± 0.20	9.87 ± 3.06	1.27 ± 0.12	6.29 ± 1.95
32	A2069	1.33 ± 0.26	3.02 ± 1.80	0.85 ± 0.17	1.98 ± 1.18
33	A2142	2.11 ± 0.18	11.62 ± 3.17	1.33 ± 0.11	7.37 ± 2.01
34	A2420	1.89 ± 0.24	8.29 ± 3.34	1.20 ± 0.16	5.31 ± 2.14
35	A2426	1.61 ± 0.32	5.20 ± 3.11	1.03 ± 0.20	3.37 ± 2.02
36	A2440	1.86 ± 0.18	8.03 ± 2.46	1.18 ± 0.11	5.15 ± 1.58
37	A2443	1.95 ± 0.22	9.38 ± 3.27	1.24 ± 0.14	5.98 ± 2.08
38	A2495	0.97 ± 0.25	1.13 ± 0.87	0.63 ± 0.16	0.76 ± 0.59
39	A2597	1.44 ± 0.21	3.69 ± 1.62	0.92 ± 0.13	2.42 ± 1.06
40	A2627	1.17 ± 0.17	2.08 ± 0.95	0.75 ± 0.11	1.38 ± 0.63
41	A2670	1.61 ± 0.23	5.11 ± 2.28	1.03 ± 0.15	3.32 ± 1.48
42	A2703	1.54 ± 0.24	4.67 ± 2.21	0.98 ± 0.15	3.03 ± 1.43
43	MKW3S	1.14 ± 0.18	1.76 ± 0.85	0.73 ± 0.11	1.18 ± 0.57
44	RXJ0132	1.10 ± 0.23	1.75 ± 1.10	0.71 ± 0.15	1.16 ± 0.73
45	RXJ0736	1.18 ± 0.26	2.07 ± 1.39	0.76 ± 0.17	1.37 ± 0.92
46	RXJ2344	1.47 ± 0.22	3.93 ± 1.82	0.94 ± 0.14	2.57 ± 1.19
47	ZWCL1023	1.27 ± 0.23	2.67 ± 1.49	0.81 ± 0.15	1.76 ± 0.98
48	ZWCL1215	1.47 ± 0.35	3.85 ± 2.80	0.94 ± 0.22	2.52 ± 1.83

Table 4.2: Physical properties measured from the weak lensing signal of the MENeCS clusters. (2) cluster name; (3) Einstein radius from the best fit SIS profile; (4) velocity dispersion from the best fit SIS profile; (5) & (7) & (9) Radius of a sphere where the cluster is overdense by a factor Δ compared to the mean (for column 5) or critical (for columns 7 and 9) density of the Universe; (6) & (8) & (10) Mass enclosed within the radius R_{Δ} .

of clusters, whose mass can be efficiently determined from scaling relations of cluster observables. Weak gravitational lensing can produce unbiased mass estimates of galaxy clusters, such that scaling relations can be calibrated. However, projection effects of overdensities along the line of sight to the cluster introduce a large uncertainty in the measurement, so that robust calibration requires large samples of clusters.

The Multi Epoch Nearby Cluster Survey (MENeACS) provides high quality optical imaging data in the g and r filters observed using the Canada-France-Hawaii Telescope (CFHT) for a sample of 58 galaxy clusters. We performed a thorough weak lensing analysis on 48 clusters in this sample, excluding some of the clusters because of their very high Galactic extinction, which prevented us from establishing a robust correction for contamination by cluster members. Our shape measurement pipeline was extensively tested in Hoekstra et al. (2015) and this gives us confidence that our systematic uncertainty in the shear estimates is $\sim 2\%$.

The MENeACS observations lack the colour information for reliable photometric redshifts and we rigorously check and correct for the uncertainties this introduces into the lensing analysis. First, the redshift distribution for the background sources in our data is approximated by the galaxy population in the COSMOS field, for which redshift information accurate to 1% is available. We consistently apply the lensing quality cuts to our photometric redshift distribution to avoid introducing biases from selection effects. To assess how representative the redshift distribution in the COSMOS field is, we use additional observations of four CFHT Legacy Survey Deep fields and find that the redshift uncertainty for our source population is at most 1.5%.

Second, without redshifts for individual sources we cannot discriminate between cluster members and field galaxies. We statistically correct for the contamination of our source sample by applying a boost correction to the shear signal. The lowest redshift clusters fill the entire field of view, so that the weight density in the cluster fields cannot be compared to the weight density of source galaxies. To overcome this issue, we used additional deep observations of blank parts of the sky from the CFHT archive to compute the average weight density for field galaxies. We determine the expected field galaxy weight density as a function of the seeing, depth of the observations and the Galactic extinction and find only a marginal uncertainty of 1% on our best fit. The intrinsic scatter in the weight density of around $\sim 6\%$ is smaller than the statistical uncertainty in the shear estimates. The final boost corrections have a residual contamination of cluster members into source sample of $\lesssim 1\%$.

The radial profiles of the boost corrected tangential shear were fit with parametric models to estimate physical properties of the clusters. Velocity dispersions from an SIS fit and masses from an NFW fit agree reasonably well with estimates based on dynamics of cluster members. The CCCP and MENeACS samples were observed under the same conditions with the same telescope and analysed with the same pipeline. We check that the two samples are similar and find no evidence for a systematic difference.

Finally, we perform an analysis to compute a scaling relation between weak lensing masses and hydrostatic masses estimates from Planck Collaboration et al. (2016c), for 30 clusters in common between MENeACS and *Planck*, which results in a bias of $1 - b = 1.03 \pm 0.08(\text{stat}) \pm 0.03(\text{syst})$. This value is high compared what was found for the CCCP and Weighing the Giants cluster samples, but in agreement with results of LoCuSS. The variation in hydrostatic estimates hints at an evolution of the hydrostatic bias with mass, as has been found in numerical simulations of clusters. We combine the CCCP and MENeACS samples to extend the mass range of the clusters and find that a power law is a better fit to the data than a constant hydrostatic bias. In future

work we shall use another mass estimator to check the effect of structure in the cluster and incorporate a careful treatment of Eddington bias. This will help to verify the mass dependence of the hydrostatic mass bias. It is unclear what the impact of such a mass dependence is for the tension between the cosmological parameters estimated from primary CMB and cluster counts.

Acknowledgements

We thank Anja Von der Linden for helpful comments.

This work is based on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/IRFU, at the Canada-France-Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Science de l'Univers of the Centre National de la Recherche Scientifique (CNRS) of France, and the University of Hawaii. This work is based in part on data products produced at Terapix available at the Canadian Astronomy Data Centre as part of the Canada-France-Hawaii Telescope Legacy Survey, a collaborative project of NRC and CNRS.

RH, CS, HH acknowledge support from the European Research Council FP7 grant number 279396.

Bibliography

- Allen S. W., Evrard A. E., Mantz A. B., 2011, *ARA&A*, 49, 409
- Amodeo S., et al., 2017, preprint, ([arXiv:1704.07891](https://arxiv.org/abs/1704.07891))
- Andreon S., 2014, *A&A*, 570, L10
- Applegate D. E., et al., 2014, *MNRAS*, 439, 48
- Arnaud M., Pratt G. W., Piffaretti R., Böhringer H., Croston J. H., Pointecouteau E., 2010, *A&A*, 517, A92
- Bahé Y. M., McCarthy I. G., King L. J., 2012, *MNRAS*, 421, 1073
- Bartelmann M., Schneider P., 2001, *Phys. Rep.*, 340, 291
- Battaglia N., Bond J. R., Pfrommer C., Sievers J. L., 2012, *ApJ*, 758, 74
- Battaglia N., et al., 2016, *JCAP*, 8, 013
- Becker M. R., Kravtsov A. V., 2011, *ApJ*, 740, 25
- Bertin E., 2006, in Gabriel C., Arviset C., Ponz D., Enrique S., eds, *Astronomical Society of the Pacific Conference Series Vol. 351, Astronomical Data Analysis Software and Systems XV*. p. 112
- Bertin E., 2010, *SWarp: Resampling and Co-adding FITS Images Together*, *Astrophysics Source Code Library* ([ascl:1010.068](https://arxiv.org/abs/1010.068))
- Bildfell C. J., 2013, PhD thesis, University of Victoria (Canada)
- Bildfell C., Hoekstra H., Babul A., Mahdavi A., 2008, *MNRAS*, 389, 1637

- Bleem L. E., et al., 2015, *ApJS*, 216, 27
- Bocquet S., et al., 2015, *ApJ*, 799, 214
- Böhringer H., et al., 2004, *A&A*, 425, 367
- Borgani S., Guzzo L., 2001, *Nature*, 409, 39
- Clowe D., Luppino G. A., Kaiser N., Henry J. P., Gioia I. M., 1998, *ApJL*, 497, L61
- Coe D., 2010, preprint, ([arXiv:1005.0411](https://arxiv.org/abs/1005.0411))
- Corless V. L., King L. J., 2007, *MNRAS*, 380, 149
- Coupon J., et al., 2009, *A&A*, 500, 981
- Diemer B., Kravtsov A. V., 2015, *ApJ*, 799, 108
- Duffy A. R., Schaye J., Kay S. T., Dalla Vecchia C., 2008, *MNRAS*, 390, L64
- Dutton A. A., Macciò A. V., 2014, *MNRAS*, 441, 3359
- Fahlman G., Kaiser N., Squires G., Woods D., 1994, *ApJ*, 437, 56
- George M. R., et al., 2012, *ApJ*, 757, 2
- Gladders M. D., Yee H. K. C., 2005, *ApJS*, 157, 1
- Graham M. L., et al., 2012, *ApJ*, 753, 68
- Gruen D., Brimiouille F., 2016, preprint, ([arXiv:1610.01160](https://arxiv.org/abs/1610.01160))
- Gwyn S. D. J., 2008, *PASP*, 120, 212
- Hasselfield M., et al., 2013, *JCAP*, 7, 008
- Henson M. A., Barnes D. J., Kay S. T., McCarthy I. G., Schaye J., 2017, *MNRAS*, 465, 3361
- Hoekstra H., 2001, *A&A*, 370, 743
- Hoekstra H., Franx M., Kuijken K., 2000, *ApJ*, 532, 88
- Hoekstra H., Hartlap J., Hilbert S., van Uitert E., 2011a, *MNRAS*, 412, 2095
- Hoekstra H., Donahue M., Conselice C. J., McNamara B. R., Voit G. M., 2011b, *ApJ*, 726, 48
- Hoekstra H., Mahdavi A., Babul A., Bildfell C., 2012, *MNRAS*, 427, 1298
- Hoekstra H., Herbonnet R., Muzzin A., Babul A., Mahdavi A., Viola M., Cacciato M., 2015, *MNRAS*, 449, 685
- Hu W., Kravtsov A. V., 2003, *ApJ*, 584, 702
- Jee M. J., et al., 2011, *ApJ*, 737, 59
- Kaiser N., Squires G., Broadhurst T., 1995, *ApJ*, 449, 460
- Kettula K., et al., 2015, *MNRAS*, 451, 1460

- Laigle C., et al., 2016, *ApJS*, 224, 24
- Leauthaud A., et al., 2016, preprint, ([arXiv:1611.08606](https://arxiv.org/abs/1611.08606))
- Magnier E. A., Cuillandre J.-C., 2004, *PASP*, 116, 449
- Mahdavi A., Hoekstra H., Babul A., Bildfell C., Jeltema T., Henry J. P., 2013, *ApJ*, 767, 116
- Mantz A., Allen S. W., Ebeling H., Rapetti D., Drlica-Wagner A., 2010, *MNRAS*, 406, 1773
- Mantz A. B., et al., 2015, *MNRAS*, 446, 2205
- Mantz A. B., et al., 2016, *MNRAS*, 463, 3582
- Mellier Y., 1999, *ARA&A*, 37, 127
- Meneghetti M., Rasia E., Merten J., Bellagamba F., Ettori S., Mazzotta P., Dolag K., Marri S., 2010, *A&A*, 514, A93
- Nagai D., Vikhlinin A., Kravtsov A. V., 2007, *ApJ*, 655, 98
- Navarro J. F., Frenk C. S., White S. D. M., 1997, *ApJ*, 490, 493
- Okabe N., Smith G. P., 2016, *MNRAS*, 461, 3794
- Okabe N., Smith G. P., Umetsu K., Takada M., Futamase T., 2013, *ApJL*, 769, L35
- Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2002, *AJ*, 124, 266
- Penna-Lima M., Bartlett J. G., Rozo E., Melin J.-B., Merten J., Evrard A. E., Postman M., Rykoff E., 2016, preprint, ([arXiv:1608.05356](https://arxiv.org/abs/1608.05356))
- Planck Collaboration et al., 2014, *A&A*, 571, A20
- Planck Collaboration et al., 2016a, *A&A*, 594, A13
- Planck Collaboration et al., 2016b, *A&A*, 594, A24
- Planck Collaboration et al., 2016c, *A&A*, 594, A27
- Rasia E., et al., 2012, *New Journal of Physics*, 14, 055018
- Rines K., Geller M. J., Diaferio A., Kurtz M. J., 2013, *ApJ*, 767, 15
- Rines K. J., Geller M. J., Diaferio A., Hwang H. S., 2016, *ApJ*, 819, 63
- Rowe B. T. P., et al., 2015, *Astronomy and Computing*, 10, 121
- Ruel J., et al., 2014, *ApJ*, 792, 45
- Rykoff E. S., et al., 2016, *ApJS*, 224, 1
- Sand D. J., et al., 2011, *ApJ*, 729, 142
- Sand D. J., et al., 2012, *ApJ*, 746, 163
- Schlafly E. F., Finkbeiner D. P., 2011, *ApJ*, 737, 103

- Schlegel D. J., Finkbeiner D. P., Davis M., 1998, *ApJ*, 500, 525
- Schneider P., 2003, *A&A*, 408, 829
- Schrabback T., et al., 2016, preprint, ([arXiv:1611.03866](https://arxiv.org/abs/1611.03866))
- Sereno M., Covone G., Izzo L., Ettori S., Coupon J., Lieu M., 2017, preprint, ([arXiv:1703.06886](https://arxiv.org/abs/1703.06886))
- Sifón C., Hoekstra H., Cacciato M., Viola M., Köhlinger F., van der Burg R. F. J., Sand D. J., Graham M. L., 2015, *A&A*, 575, A48
- Sifón C., et al., 2016, *MNRAS*, 461, 248
- Smith G. P., et al., 2016, *MNRAS*, 456, L74
- Sunyaev R. A., Zeldovich Y. B., 1972, *Comments on Astrophysics and Space Physics*, 4, 173
- Umetsu K., et al., 2014, *ApJ*, 795, 163
- Vikhlinin A., et al., 2009a, *ApJ*, 692, 1033
- Vikhlinin A., et al., 2009b, *ApJ*, 692, 1060
- Viola M., et al., 2015, *MNRAS*, 452, 3529
- Wright C. O., Brainerd T. G., 2000, *ApJ*, 534, 34
- de Haan T., et al., 2016, *ApJ*, 832, 95
- von der Linden A., et al., 2014a, *MNRAS*, 439, 2
- von der Linden A., et al., 2014b, *MNRAS*, 443, 1973

4.A Blank field counts

The lack of deep multi-band data for galaxies in the MENeaCS observations prevents us from identifying cluster members. As they are unlensed and have random orientations these galaxies will dilute the shear signal and thus need to be corrected for. We correct for this contamination using a boost correction, for which we need to model the excess of galaxies in the cluster compared to the field as a function of cluster-centric radius. This approach was also used by H15, who used the observations beyond 4 Mpc from the cluster center to estimate the background level for the targets observed with MegaCam. However, the MENeaCS clusters are at much lower redshift, and for many targets the data do not extend that far out. Therefore, we follow a different approach to determine the expected background level.

We searched the CFHT MegaPipe (Gwyn 2008) archive for co-added r -band data with a total integration $T_{\text{exp}} > 3600\text{s}$ and image quality better than $0''.9$ of pointings that appeared to be blank fields (i.e. not targeting clusters). This query resulted in 46 suitable unique targets, but upon closer inspection five fields had to be rejected because of a high noise level or because the Galactic extinction was too high. The remaining 41 fields all have an r -band Galactic extinction A_r less than 0.2 magnitude

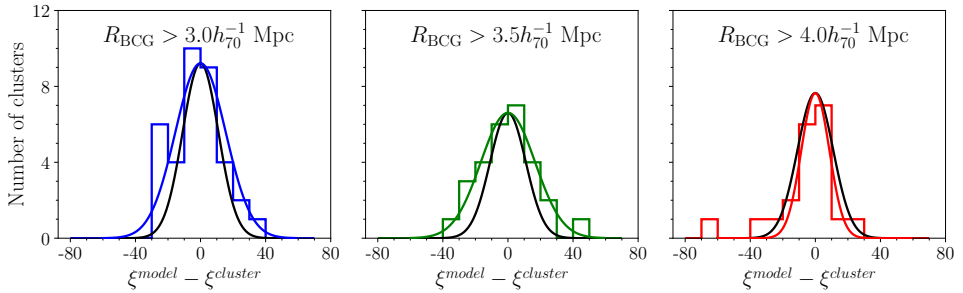


Figure 4.8: Histogram of the difference between the predicted weight density from the blank fields ξ^{model} and the weight density measured outside the cluster ξ^{cluster} for the highest redshift MENeCS clusters. As the full extent of the cluster is unknown we show ξ^{cluster} for different areas corresponding to a radius of 3 Mpc, 3.5 Mpc and 4 Mpc from the BCG, from left to right respectively. The decreasing number of clusters shows that only the highest redshift clusters have any area available and highlights the need for the model prediction ξ^{model} . The black curve is the same for all three panels and it shows the distribution expected from the 6% scatter around the best fit for the blank fields.

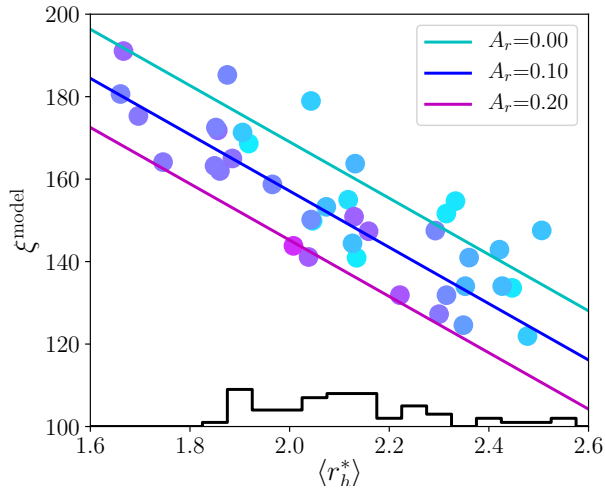


Figure 4.9: Weight density ξ as a function of the average PSF size in the image, color coded by the Galactic extinction for 41 CFHT observations homogenised to a noise rms of 1.4. Colored lines show the best fit to the data at the four different noise levels, using the same color code as the circles. The black histogram shows the distribution of PSF sizes in the MENeCS data, covering the same range as the blank fields.

and were visually inspected to mask out obvious artifacts, leaving an effective area of approximately 33 degrees.

The noise levels vary between the images and do not match the MENeACS data (they are typically deeper). To homogenise the data we added Gaussian noise so that all the blank fields had the same noise level. We considered four r.m.s. values $\sigma_n = (1.2, 1.4, 1.6, 1.8)$ to cover a range that matches most of the MENeACS data. For the two lowest noise levels we omitted ten and two fields, respectively. The resulting images were analysed in exactly the same way as the MENeACS data, resulting in catalogs with shape measurements and corresponding uncertainties. To quantify image quality we use the half-light radius of the PSF r_h^* .

Figure 4.9 shows the resulting weight density ξ for galaxies with $20 < m_r < 24.5$ as a function of the average PSF size in the image for a noise level of 1.4. Using the blanks at all 4 different noise levels we fit a simple model to the measurements and find

$$\xi^{\text{model}}(r_h^*, \sigma_n, A_r) = -40.6 \langle r_h^* \rangle - 68.4 \sigma_n - 122.8 A_r + 364.2, \quad (4.14)$$

where A_r is the Galactic extinction in the r -band. The color of the circles shows the extinction for each blank field and the lines show the prediction from the fit for different extinction levels in the same color scheme. For reference we also plot the distribution of PSF half-light radii found in the MENeACS data as a black histogram in the bottom of the plot.

For the full sample we find that the r.m.s. variation in the mean weight density is 6.4%, which is smaller than the typical statistical uncertainty in the lensing signal for an individual cluster. Hence observing clusters with a single band and modelling the excess weight as a function of cluster-centric distance is an efficient way to determine masses for a large sample of clusters. Although the contamination can be largely eliminated using multi-band data, the improvement in precision is modest. On the other hand, the limited number of blank fields limits our current precision of 1% with which the average background weight density can be determined.

The blank fields have been observed with different dither patterns than the MENeACS data. Consequently, the variations in depth will not exactly match the cluster data. To examine whether this leads to a significant systematic uncertainty, we compare the predicted weight density to the observations of high redshift clusters beyond the extent of the cluster. H15 found that the contribution from cluster members and associated structures is less than 0.5% for radii beyond 4 Mpc and we use this to define the areas in the MENeACS observations with mostly field galaxies. Due to the low redshift of MENeACS clusters there is very little area beyond 4 Mpc, so we vary the outer radius for a more robust comparison.

Figure 4.8 shows the distribution of the difference between the weight density in the blank field prediction and in the cluster data. For all three outer radii – 3.0 Mpc, 3.5 Mpc and 4.0 Mpc from left to right respectively – the scatter is centered around zero. The colored curves show the best fit Gaussian, which can be compared to the black curve, which shows the variation expected from the blank fields. We have fixed the centre of the Gaussian on zero, but leaving the centre as a free parameter changes little in the fit. The overall agreement between the scatter in the blank fields and the cluster outskirts is remarkably good. The scatter in all but the furthest outskirts is slightly larger than the scatter exhibited in the blank fields. The available area where $R_{\text{BCG}} > 4$ Mpc is so small that the lower scatter in the cluster outskirts might be due only to low number statistics. In fact, Figure 4.2 shows little evidence for cluster members outside 2 Mpc, so that the area $R_{\text{BCG}} > 3$ Mpc should also provide a clean

sample of source galaxies. Trusting the left panel in Figure 4.8, we investigate the larger scatter of the cluster data. We suspect it is caused by the limited area available in the cluster data. To check this, we re-evaluated the scatter around the blank field prediction using only part of the full blank fields. For half the area, the scatter around the best fit increased by $\sim 20\%$, which increased to $\sim 50\%$ when using only a quarter of the area. As expected, the scatter increases as we decrease the size of the used area, corroborating our hypothesis for the scatter of the cluster data. This again shows that our blank field prediction for the weight density of field galaxies is a reliable tool for the normalisation of the weight density in the cluster data.

The galaxy-subhalo connection in low-redshift galaxy clusters from weak gravitational lensing

We measure the gravitational lensing signal around satellite galaxies in a sample of galaxy clusters at $z < 0.15$ by combining high-quality imaging data from the Canada-France-Hawaii Telescope with a large sample of spectroscopically-confirmed cluster members. We use extensive image simulations to assess the accuracy of shape measurements of faint, background sources in the vicinity of bright satellite galaxies. We find a small but significant bias, as light from the lenses makes the shapes of background galaxies appear radially aligned with the lens. We account for this bias by applying a correction that depends on both lens size and magnitude. We also correct for contamination of the source sample by cluster members. We use a physically-motivated definition of subhalo mass, namely the mass bound to the subhalo, m_{bg} , similar to definitions used by common subhalo finders in numerical simulations. Binning the satellites by stellar mass we provide a direct measurement of the subhalo-to-stellar-mass relation, $\log m_{\text{bg}}/M_{\odot} = (11.66 \pm 0.07) + (0.94 \pm 0.15) \log[m_{\star}/(2 \times 10^{10} M_{\odot})]$. This best-fitting relation implies that, at a stellar mass $m_{\star} \sim 3 \times 10^{10} M_{\odot}$, subhalo masses are roughly 50 per cent lower than their central counterparts, and this fraction decreases at higher stellar masses. On the other hand, we find no statistically significant evidence for mass segregation when we bin lenses by their projected cluster-centric distance, contrary to recent claims, with an average total-to-stellar mass ratio $\langle m_{\text{bg}}/m_{\star} \rangle = 21.5_{-5.5}^{+6.3}$. We find that, once we account for projection effects and for differences between dark matter subhaloes and satellite galaxies, our results are consistent with theoretical predictions.

5.1 Introduction

According to the hierarchical structure formation paradigm, galaxy clusters grow by the continuous accretion of smaller galaxy groups and individual galaxies. Initially, each of these systems is hosted by their own dark matter halo, but as a galaxy falls into a larger structure, tidal interactions transfer mass from the infalling galaxy to the new host. The galaxy then becomes a satellite and its dark matter halo, a subhalo.

Detailed studies on the statistics of subhaloes from numerical N-body simulations have revealed that subhaloes are severely affected by their host haloes. Dynamical friction makes more massive subhaloes sink towards the centre faster, while tidal stripping removes mass preferentially from the outskirts of massive subhaloes closer to the centre. These two effects combined destroy the most massive subhaloes soon after infall (e.g., Tormen et al. 1998; Taffoni et al. 2003), a result exaggerated in simulations with limited resolution (e.g., Klypin et al. 1999; Taylor & Babul 2005; Han et al. 2016). Tidal stripping makes subhaloes more concentrated than field haloes of the same mass (e.g., Ghigna et al. 1998; Springel et al. 2008; Moliné et al. 2017), and counterbalances the spatial segregation induced by dynamical friction (van den Bosch et al. 2016).

One of the most fundamental questions is how these subhaloes are linked to the satellite galaxies they host, which are what we observe in the real Universe. Taking N-body simulations at face value results in serious inconsistencies with observations, the most famous of which are known as the “missing satellites” (Klypin et al. 1999; Moore et al. 1999) and “too big to fail” (Boylan-Kolchin et al. 2011) problems. It has since become clear that these problems may arise because baryonic physics has a strong influence on the small-scale distribution of matter. Energetic feedback from supernovae at the low-mass end, and active galactic nuclei at the high-mass end, of the galaxy population affect the ability of dark matter (sub)haloes to form stars and retain them. In addition, the excess mass in the centre of galaxies (compared to dark matter-only simulations) can modify each subhalo’s susceptibility to tidal stripping (e.g., Zolotov et al. 2012).

Despite these difficulties, given the current technical challenges of generating cosmological high-resolution hydrodynamical simulations (in which galaxies form self-consistently), N-body simulations remain a valuable tool to try to understand the evolution of galaxies and (sub)haloes. In order for them to be applied to real observations, however, one must post-process these simulations in some way that relates subhaloes to galaxies, taking into account the aforementioned complexities (and others). For instance, semi-analytic models contain either physical or phenomenological recipes whether or not to form galaxies in certain dark matter haloes based on the mass and assembly history of haloes (e.g., Bower et al. 2006; Lacey et al. 2015). A different method involves halo occupation distributions (HODs), which assume that the average number of galaxies in a halo depends only on host halo mass. Because they provide an analytical framework to connect galaxies and dark matter haloes, HODs are commonly used to interpret galaxy-galaxy lensing and galaxy clustering measurements through a conditional stellar mass (or luminosity) function (e.g., Seljak 2000; Peacock & Smith 2000; Mandelbaum et al. 2006; Cacciato et al. 2009; van den Bosch et al. 2013).

One of the key aspects of these prescriptions is the stellar-to-halo mass relation. While many studies have constrained the stellar-to-halo mass relation of central galaxies (e.g., Hoekstra et al. 2005; Heymans et al. 2006b; Mandelbaum et al. 2006, 2016; More et al. 2011; van Uitert et al. 2011, 2016; Leauthaud et al. 2012; Velander et al. 2014; Coupon et al. 2015; Zu & Mandelbaum 2015), this is not the case for satel-

lite galaxies, whose *subhalo-to-stellar* mass relation (SHSMR) remains essentially unexplored, and the constraints so far are largely limited to indirect measurements. Rodríguez-Puebla et al. (2012) used abundance matching (the assumption that galaxies rank-ordered by stellar mass can be uniquely mapped to [sub]haloes rank-ordered by total mass) to infer the SHSMR using the satellite galaxy stellar mass function, and Rodríguez-Puebla et al. (2013) extended these results using galaxy clustering measurements. They showed that the SHSMR is significantly different from the central stellar-to-total mass relation, and that assuming an average relation when studying a mixed population can lead to biased results (see also Yang et al. 2009).

Instead, only stellar dynamics and weak gravitational lensing provide direct ways to probe the total gravitational potential of a galaxy. However, the quantitative connection between stellar velocity dispersion and halo mass is not straightforward (e.g., Li et al. 2013b; Old et al. 2015), and only weak lensing provides a direct measurement of the total surface mass density (Fahlman et al. 1994; Clowe et al. 1998). Using deep Hubble Space Telescope (HST) observations, Natarajan et al. (1998, 2002, 2007, 2009) measured the weak (and also sometimes strong) lensing signal of galaxies in six clusters at $z = 0.2 - 0.6$. After fitting a truncated density profile to the ensemble signal using a maximum likelihood approach, they concluded that galaxies in clusters are strongly truncated with respect to field galaxies. Using data for clusters at $z \sim 0.2$ observed with the CFH12k instrument on the Canada-Hawaii-France Telescope (CFHT), Limousin et al. (2007) arrived at a similar conclusion. Halkola et al. (2007) and Suyu & Halkola (2010) used strong lensing measurements of a single cluster and a small galaxy group, respectively, and also found evidence for strong truncation of the density profiles of satellite galaxies. However, Pastor Mira et al. (2011) have argued that the conclusion that cluster galaxies are truncated from these (strong and weak) galaxy-galaxy lensing measurements are driven by the parametrization of the galaxy density profiles rather than constraints from the data themselves.

Recent combinations of large weak lensing surveys with high-purity galaxy group catalogues have allowed direct measurements of the average subhalo masses associated with satellite galaxies using weak galaxy-galaxy lensing (Li et al. 2014, 2016; Sifón et al. 2015a; Niemiec et al. 2017). However, these studies did not focus on the SHSMR but on the segregation of subhaloes by mass within galaxy groups, by measuring subhalo masses at different group-centric distances. The observational results are consistent, within their large errorbars, with the mild segregation of dark matter subhaloes seen in numerical simulations (Han et al. 2016; van den Bosch et al. 2016). However, it is not clear whether results based on subhaloes in N-body simulations can be directly compared to observations. In fact, van den Bosch (2017) has shown that the statistics of subhaloes inferred from N-body simulations are problematic even to this day, because of severe numerical destruction of subhaloes.

In this work, we present weak gravitational lensing measurements of the total mass of satellite galaxies in 48 massive galaxy clusters at $z < 0.15$. Our images were taken with the MegaCam instrument on the Canada-France-Hawaii Telescope (CFHT), which has a field of view of 1 sq. deg., allowing us to focus on very low redshift clusters and take advantage of the $< 1''$ seeing (corresponding to 1.84 kpc at $z = 0.1$) of our observations. We can therefore probe the lensing signal close to the galaxies themselves, at a physical scale equivalent to what can be probed in a cluster at $z \sim 0.5$ with HST, but out to the clusters' virial radii. In addition, the low-redshift clusters we use have extensive spectroscopic observations available from various data sets, compiled by Sifón et al. (2015b), so we do not need to rely on uncertain photometric

identification of cluster members.

This paper is organized as follows. We summarize the galaxy-galaxy lensing formalism in Section 5.2. We describe our data set in Section 5.3, taking a close look at the source catalogue and the shapes of background sources in Section 5.4. We present our modelling of the satellite lensing signal in Section 5.5, and discuss the connection between mass and light in satellite galaxies, in the form of the subhalo-to-stellar mass relation and subhalo mass segregation, in Sections 5.6 and 5.7, respectively. Finally, we summarize in Section 5.8.

We adopt a flat Λ cold dark matter (Λ CDM) cosmology with $\Omega_m = 0.315$, based on the latest results from cosmic microwave background observations by Planck Collaboration (2015), and $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$. In this cosmology, $10'' = \{9.8, 18.4, 26.1\} \text{ kpc}$ at $z = \{0.05, 0.1, 0.15\}$. As usual, stellar and (sub)halo masses depend on the Hubble constant as $m_\star \sim 1/H_0^2$ and $m \sim 1/H_0$, respectively.

5.2 Weak galaxy-galaxy lensing

Gravitational lensing distorts the images of background (“source”) galaxies as their light passes near a matter overdensity along the line-of-sight. This produces a distortion in the shape of the background source, called *shear*, and a *magnification* effect on the source’s size (and consequently its brightness). The shear field around a massive object aligns the images of background sources around it in the tangential direction. Therefore, starting from a measurement of the shear of an object in a cartesian frame with components (γ_1, γ_2) (see Section 5.3.3), it is customary to parametrize the shear as

$$\begin{pmatrix} \gamma_t \\ \gamma_\times \end{pmatrix} = \begin{pmatrix} -\cos 2\phi & -\sin 2\phi \\ \sin 2\phi & -\cos 2\phi \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix}, \quad (5.1)$$

where ϕ is the azimuthal angle of the lens-source vector, γ_t measures the ellipticity in the tangential ($\gamma_t > 0$) and radial ($\gamma_t < 0$) directions and γ_\times measures the ellipticity in directions 45° from the tangent. Because of parity symmetry, we expect $\langle \gamma_\times \rangle = 0$ for an ensemble of lenses (Schneider 2003) and therefore γ_\times serves as a test for systematic effects.

The shear is related to the excess surface mass density (ESD), $\Delta\Sigma$, via

$$\Delta\Sigma(R) \equiv \bar{\Sigma}(< R) - \bar{\Sigma}(R) = \gamma_t \Sigma_c, \quad (5.2)$$

where $\bar{\Sigma}(< R)$ and $\bar{\Sigma}(R)$ are the average surface mass density within a radius¹ R and within a thin annulus at distance R from the lens. The critical surface density, Σ_c , is a geometrical factor that accounts for the lensing efficiency,

$$\Sigma_c = \frac{c^2}{4\pi G} \frac{D_s}{D_l D_{ls}}, \quad (5.3)$$

where, D_l , D_s , and D_{ls} are the angular diameter distances to the lens, to the source and between the lens and the source, respectively. The ESD for each bin in lens-source separation is then

$$\Delta\Sigma = \frac{\sum_i w_i \Sigma_{c,i} \gamma_{t,i}}{\sum_i w_i}, \quad (5.4)$$

¹As a convention, we denote three-dimensional distances with lower case r and two-dimensional distances (that is, projected on the sky) with upper case R .

where the sums run over all lens-source pairs in a given bin and the weight of each source galaxy is given by

$$w_i = \frac{1}{\langle \epsilon_{\text{int}}^2 \rangle + (\sigma_{\gamma,i})^2}. \quad (5.5)$$

Here, σ_γ is the measurement uncertainty in γ_i , which results from the quadrature sum of statistical uncertainties due to shot noise in the images (see Section 5.3.3) and from uncertainties in the modelling of a measurement bias detailed in Sections 5.4.2 and 5.A.² We set the intrinsic root-mean-square galaxy ellipticity, $\langle \epsilon_{\text{int}}^2 \rangle^{1/2}$, to 0.25. In Equation 5.4, we use a single value for Σ_c for all satellites in each cluster (see Section 5.4.5).

In fact, the weak lensing observable is the *reduced* shear, $g \equiv \gamma/(1 - \kappa)$ (where $\kappa = \Sigma/\Sigma_c$ is the lensing convergence), but in the weak limit $\kappa \ll 1$ so that $g \approx \gamma$. However, close to the centres of galaxy clusters the convergence is of order unity, so this approximation is not accurate anymore. To account for this, the lensing model presented in Section 5.5 is corrected using

$$g(R) = \frac{\gamma(R)}{1 - \bar{\Sigma}(R)/\Sigma_c} = \frac{\Delta\Sigma(R)/\Sigma_c}{1 - \bar{\Sigma}(R)/\Sigma_c}. \quad (5.6)$$

5.2.1 Statistical errors: data covariance

Because the gravitational potential of satellites in a cluster is traced by the same background source galaxies, data points in the ESD are correlated. Following Viola et al. (2015), we can re-arrange Equation 5.4 to reflect the contribution from each *source* galaxy. The data covariance of measurements in a single cluster can then be written as

$$\mathbf{C}_{mij} = \Sigma_c^2 \langle \epsilon^2 \rangle \frac{\sum_s (C_{si,m} C_{sj,n} + S_{si,m} S_{sj,n})}{(\sum_s Z_{si,m})(\sum_s Z_{sj,n})}, \quad (5.7)$$

where index pairs m, n and i, j run over the observable bins (e.g., stellar mass) and lens-source separation, R , respectively, and C , S and Z are sums over the lenses:

$$\begin{aligned} C_{si} &= - \sum_l w_{ls} \cos 2\phi_{ls}, \\ S_{si} &= - \sum_l w_{ls} \sin 2\phi_{ls}, \\ Z_{si} &= \sum_l w_{ls}, \end{aligned} \quad (5.8)$$

where we explicitly allow for the possibility that the source weight, w , may be different for each lens-source pair (as opposed to a unique weight per source). This is indeed the case when we consider the corrections to the shape measurements from lens contamination discussed in Sections 5.4.2 and 5.A, although in practice differences are negligible. As implied by Section 5.7, we assign the same Σ_c to all galaxies that are part of the same cluster. The total ESD is then the inverse-covariance-weighted sum of the ESDs of individual clusters.

In addition to the data covariance there is, in principle, a contribution to the measurement uncertainty from sample variance. By comparing Equation 5.7 to uncertainties estimated by bootstrap resampling, Sifón et al. (2015a) have shown that the

²In practice, the latter is negligible in most cases.

contribution from sample variance is less than 10 per cent for satellite galaxy-galaxy lensing measurements when limited to small lens-source separations ($R \lesssim 2$ Mpc). Since the signal from satellites themselves is limited to $R \lesssim 300$ kpc (Figure 5.8; see also Sifón et al. 2015a), in this work we ignore the sample variance contribution to the lensing covariance.

5.3 Data set

In this section we describe the lens and source galaxy samples we use in our analysis. In the next section, we make a detailed assessment of the shape measurement and quality cuts on the source sample using extensive image simulations.

5.3.1 Cluster and lens galaxy samples

The Multi-Epoch Nearby Cluster Survey (MENeACS, Sand et al. 2012) is a targeted survey of 57 galaxy clusters in the redshift range $0.05 \lesssim z \lesssim 0.15$ observed in the g and r bands with MegaCam on CFHT. We only use the 48 clusters affected by r -band Galactic extinctions $A_r \leq 0.2$ mag, since we find that larger extinctions bias the source number counts and the correction for cluster member contamination (Section 5.4). The image processing and photometry are described in detail in van der Burg et al. (2013); most images have seeing $\lesssim 0.8''$. Sifón et al. (2015b) compiled a large sample of spectroscopic redshift measurements in the direction of 46 of these clusters, identifying a total of 7945 spectroscopic members. Since, Rines et al. (2016) have published additional spectroscopic redshifts for galaxies in 12 MENeACS clusters, six of which are included in Sifón et al. (2015b) but for which the observations of Rines et al. (2016) represent a significant increase in the number of member galaxies. We select cluster members in these 12 clusters in an identical way as Sifón et al. (2015b). The median dynamical mass of MENeACS clusters is $M_{200} \sim 6 \times 10^{14} M_\odot$ (Sifón et al. 2015b).

From the member catalogue of Sifón et al. (2015b), we exclude all brightest cluster galaxies (BCGs), and refer to all other galaxies as satellites. Because the shapes of background galaxies near these members are very likely to be contaminated by light from the BCG, we also exclude all satellite galaxies within $10''$ of the BCGs to avoid severe contamination from extended light. Finally, we impose a luminosity limit $L_{\text{sat}} < \min(2L^*, 0.5L_{\text{BCG}})$ (where $L^*(z)$ is the r -band luminosity corresponding to the characteristic magnitude, $m_{\text{phot}}^*(z)$ of the Schechter (1976) function, fit to red satellite galaxies in redMaPPer galaxy clusters over the redshift range $0.05 < z < 0.7$ (Rykoff et al. 2014)).³ We choose the maximum possible luminosity, $2L^*$, because the BCGs in our sample have $L_{\text{BCG}} \gtrsim 3L^*$, so this ensures we do not include central galaxies of massive (sub)structures that could, for instance, have recently merged with the cluster. In addition, we only include satellites within 2 Mpc of the BCG. At larger distances, contamination by fore- and background galaxies becomes an increasingly larger problem. Our final spectroscopic sample consists of 5414 satellites in 45 clusters.

In addition, we include red sequence galaxies in all MENeACS clusters in low Galactic extinction regions in order to improve our statistics. We measure the red sequence

³Equation 9 of Rykoff et al. (2014) provides a fitting function for the i -band $m_{\text{phot}}^*(z)$, which we convert to r -band magnitudes assuming a quiescent spectrum, appropriate for the majority of our satellites, using EZGAL (<http://www.baryons.org/ezgal/>, Mancone & Gonzalez 2012).

Table 5.1: Number of galaxies and average properties of stellar mass and cluster-centric distance bins used in Sections 5.6 and 5.7. Sub-columns correspond to the values of the fiducial spectroscopic-plus-red-sequence and the spectroscopic-only samples.

Binning observable	Bin label	Range	N_{sat}		$\langle R_{\text{sat}}/\text{Mpc} \rangle$		$\log\langle m_{\star}/M_{\odot} \rangle$	
			spec	spec	spec	spec	spec	spec
			+RS		+RS		+RS	
$\log(m_{\star}/M_{\odot})$	M1	[9.0 – 9.8)	2144	1010	0.66	0.88	9.51	9.51
	M2	[9.8 – 10.2)	2017	1315	0.67	0.87	10.01	10.03
	M3	[10.2 – 10.5)	1387	1146	0.80	0.91	10.36	10.35
	M4	[10.5 – 10.9)	1178	1052	0.83	0.89	10.67	10.67
	M5	[10.9 – 11.2]	278	265	0.93	0.98	11.01	11.01
$R_{\text{sat}} (\text{Mpc})$	D1	[0.1 – 0.35)	1346	664	0.23	0.23	9.97	10.20
	D2	[0.35 – 0.7)	1934	1139	0.52	0.52	10.03	10.20
	D3	[0.7 – 1.2)	1994	1397	0.90	0.94	10.07	10.22
	D4	[1.2 – 2.0)	1550	1529	1.55	1.55	10.24	10.25

by fitting a straight line to the colour-magnitude relation of red galaxies in each cluster using a maximum likelihood approach, based on the methodology of Hao et al. (2009). Following Sifón et al. (2015b), we include only red sequence galaxies brighter than $M_r = -19$ and within 1 Mpc of the BCG.⁴ When we include red sequence galaxies, we also use the six clusters without spectroscopic cluster members. Therefore our combined spectroscopic plus red sequence sample includes 7909 cluster members in 48 clusters (including three clusters without spectroscopic data). Throughout, we refer to the spectroscopic and spectroscopic plus red sequence samples as ‘spec’ and ‘spec+RS’, respectively.

For the purpose of estimating stellar masses and photometric redshifts, the original MENeCS observations in g and r were complemented by u - and i -band observations with the Wide-Field Camera on the Isaac Newton Telescope in La Palma (except for a few clusters with archival MegaCam data in either of these bands, see van der Burg et al. 2015, for details). Stellar masses were estimated by van der Burg et al. (2015) by fitting each galaxy’s spectral energy distribution using FAST (Kriek et al. 2009) assuming a Chabrier (2003) initial mass function.

In order to characterize the connection between satellite galaxies and their host subhaloes, we split the sample by stellar mass (Section 5.6) and cluster-centric distance (Section 5.7), each time splitting the sample in five bins. We show the stellar mass and cluster-centric distributions of the resulting subsamples in Figure 5.1, and list the average values in Table 5.1.

5.3.2 Source galaxy sample

We construct the source catalogues in an identical manner to Hoekstra et al. (2015), except for one additional flag to remove galaxies whose shape is significantly biased by the presence of a nearby bright object. This step is discussed in detail in Section 5.4.1. The biases in the shape measurements of the sources, depending on how the source sample is defined, have been characterized in great detail by Hoekstra et al. (2015).

⁴Here, M_r is the $k + e$ -corrected absolute magnitude in the r -band, calculated with EZGAL using a passively evolving Charlot & Bruzual (2007, unpublished, see Bruzual & Charlot 2003) model with formation redshift $z_f = 5$.

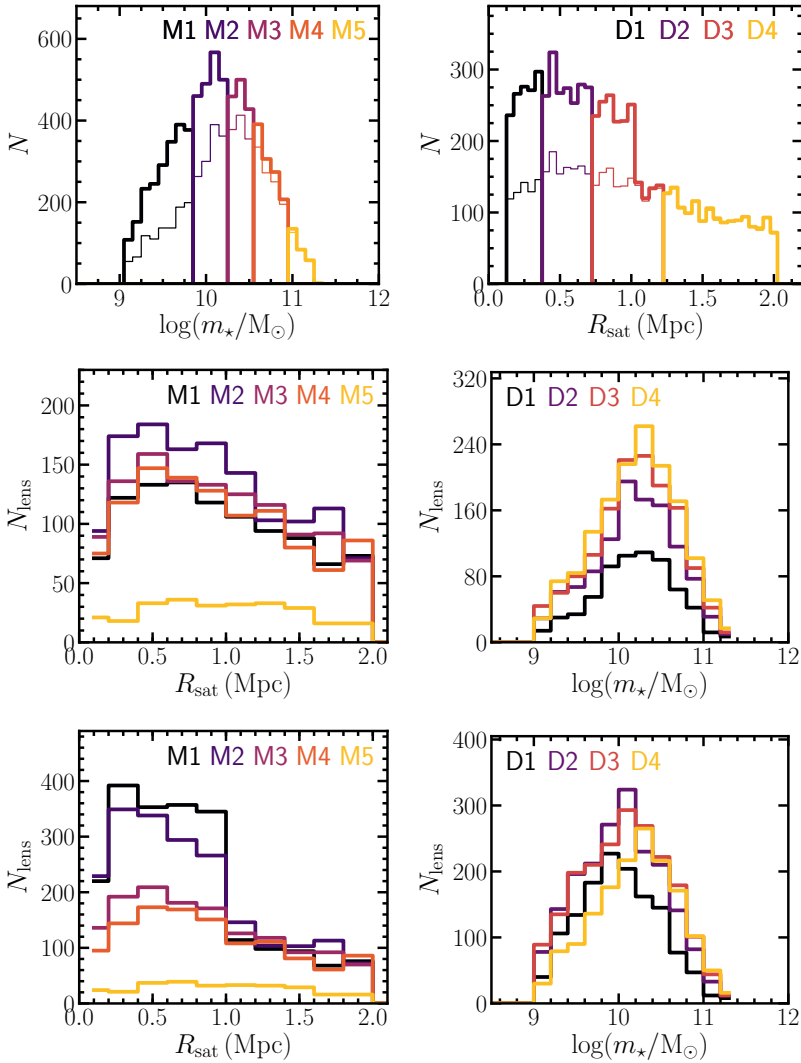


Figure 5.1: The five stellar mass bins used in Section 5.6 (lefthand plots) and the four cluster-centric distance bins used in Section 5.7 (righthand plots). Top panels show histograms for spectroscopic (‘spec’ sample, thin lines) and spectroscopic plus red sequence (‘spec+RS’ sample, thick lines) members. Middle and bottom panels show distribution of different stellar mass bins in cluster-centric distance (left) and of different cluster-centric radius bins in stellar mass (right), for the spec and spec+RS samples, respectively. Note the different vertical scales in each panel.

Although the study of Hoekstra et al. (2015) refers to a different cluster sample, both samples have been observed with the same instrument under very similar conditions of high image quality, so we can safely take the analysis of Hoekstra et al. (2015) as a reference for our study.

Specifically, we select only sources with r -band magnitudes⁵ $20 < m_{\text{phot}} < 24.5$, with sizes $r_h < 5$ pix and an additional constraint on δm_{phot} , the difference in estimated magnitude before and after the local background subtraction used for shape measurements (see Section 5.4.1). Compared to Hoekstra et al. (2015), who used $22 < m_{\text{phot}} < 25$, we choose different magnitude limits (i) at the bright end because our cluster sample is at lower redshift and therefore cluster members tend to be brighter, and (ii) at the faint end because our data are slightly shallower, complicating the shape measurements of very faint sources. The magnitudes m_{phot} have been corrected for Galactic extinction using the Schlafly & Finkbeiner (2011) recalibration of the Schlegel et al. (1998) infrared-based dust map.

Unlike most cluster lensing studies (e.g., Hoekstra et al. 2012; Applegate et al. 2014; Umetsu et al. 2014), we do not apply a colour cut to our source sample, since this only reduces contamination by ~ 30 per cent for $z \sim 0.2$ clusters (Hoekstra 2007). In fact, one of the advantages of using low-redshift clusters is that contamination by cluster members is significantly lower than at higher redshifts, since cluster members are spread over a larger area on the sky. Instead of applying colour cuts to reduce contamination, we follow Hoekstra et al. (2015) and correct for contamination in the source sample by applying a ‘boost factor’ to the measured lensing signal to account for the dilution by cluster members (e.g., Mandelbaum et al. 2005a). We discuss this and other corrections to the shape measurements, along with the source redshift distribution, in Section 5.4.

5.3.3 Shape measurements

To measure the galaxy-galaxy lensing signal we must accurately infer the shear field around the lenses by measuring the shapes of as many background galaxies as possible. For most of the sources this is a difficult procedure as they are faint and of sizes comparable to the image resolution, quantified by the point spread function (PSF). Blurring by the PSF and noise lead to a multiplicative bias, μ , while an anisotropic PSF introduces an additive bias, c (e.g., Heymans et al. 2006a). The measured (or observed) shear is therefore related to the true shear by

$$\gamma^{\text{obs}}(\theta_{\text{ls}}|R_{\text{sat}}) = (1 + \mu) \gamma^{\text{true}}(\theta_{\text{ls}}) \mathcal{B}^{-1}(R_{\text{sat}}) + c, \quad (5.9)$$

where θ_{ls} is the lens-source separation and μ and c are referred to simply as the multiplicative and additive biases, respectively; $\mathcal{B}(R_{\text{sat}})$ is the ‘boost factor’ that corrects for contamination by cluster members, described in Section 5.4.3. Note that μ , c and $\mathcal{B}(R_{\text{sat}})$ all depend on both the dataset and the shape measurement method.

We measure galaxy shapes by calculating the moments of galaxy images using the KSB method (Kaiser et al. 1995; Luppino & Kaiser 1997), incorporating the modifications by Hoekstra et al. (1998, 2000). The PSF is measured from the shapes of stars in the image and interpolation between stars is used to estimate the PSF for each galaxy. Hoekstra et al. (2015) used extensive image simulations to assess the performance of KSB depending on the observing conditions and background source

⁵We denote r -band magnitudes with m_{phot} in order to avoid confusion with subhalo masses, which we denote with lower case m and subscripts depending on the definition (see Section 5.5.2).

ellipticity, magnitude and size distributions. We adopt the size- and signal-to-noise-dependent multiplicative bias correction obtained by Hoekstra et al. (2015). Instead of correcting each source’s measured shape, we apply an average correction to each data point (which is an average over thousands of sources), since the latter is more robust to uncertainties in the intrinsic ellipticity distribution (Hoekstra et al. 2015). In the next section we take a detailed look at possible sources of bias in our shape measurements.

Due to lensing, sources are magnified as well as sheared, and this may alter the inferred source density, affecting the boost correction discussed in Section 5.4.3. The increase in flux boosts the number counts relative to an unlensed area of the sky, but the decrease in effective area works in the opposite direction. The net effect depends on the intrinsic distribution of source galaxies as a function of magnitude, and cancels out for a slope $d \log N_{\text{source}}/dm_{\text{phot}} = 0.40$ (Mellier 1999). In fact, this slope is 0.38–0.40 for the MegaCam r -band data (Hoekstra et al. 2015), so we can safely ignore magnification in our analysis.

In order to account for the measurement uncertainties in defining the quality of our lensing data, throughout this work we use the source *weight* density. We define the weight density, $\xi_s \equiv (1/A) \sum_i w_i$, as the sum of the shape measurement weights (Equation 5.5) per square arcminute.

5.4 Source sample and shear calibration

We now explore the impact of cluster galaxies in our analysis, as they contaminate our source sample and in some cases bias shape measurements through blending of their light with that of source galaxies. In order to assess the impact of cluster galaxies in the shear measurement pipeline, we use dedicated sets of image simulations. We extend the image simulations produced by Hoekstra et al. (2015) by introducing simulated cluster galaxies into the images of source galaxies. We create two sets of image simulations with different cluster galaxies to investigate different features of the analysis pipeline, as described in the following sections.

The image simulation pipeline of Hoekstra et al. (2015) creates mock images of the MegaCam instrument with randomly placed source galaxies. In short, these simulated galaxies have properties based on GALFIT (Peng et al. 2002) measurements of galaxies in the GEMS survey (Rix et al. 2004). The modulus of the ellipticity is drawn from a Rayleigh distribution with a width of 0.25 and truncated at 0.9, and galaxies are assigned random position angles. Figure 5.2 shows the distribution of magnitudes and sizes measured with GALFIT of MENeCS cluster galaxies (from Sifón et al. 2015b). We use these measurements to simulate lens galaxies which we add to the simulations of source galaxies. The surface brightness profiles of galaxies are drawn, assuming their light follows Sérsic (1968) radial profiles, using the GALSIM software (Rowe et al. 2015).

5.4.1 Sensitivity to background subtraction

Before discussing the impact of cluster galaxies in the source sample and shape measurements, we describe a bias pertaining to the shape measurement pipeline itself. The pipeline proceeds in two steps: the first is to detect sources using a global background estimation, while the second is to measure the shapes of these detected objects. In the second step, a local background level is determined by measuring the root mean square

brightness in an annulus with inner and outer radii of 16 and 32 pixels respectively, after masking all detected objects. This annulus is split into four quadrants. The background is modelled by fitting a plane through them, and is then subtracted from the image. This background subtraction works well in general, but when light from nearby objects is not properly accounted for, it significantly modifies the estimated magnitude of the test galaxy. Since the simulations do not have a diffuse background component, a proper background subtraction would leave the galaxy magnitude untouched. Therefore, changes in the magnitude pre- and post-background subtraction in the simulations, which we denote $\delta m_{\text{phot}} \equiv m_{\text{postbg}} - m_{\text{prebg}}$, mean that the shape measurement process is not robust for that particular galaxy. As our sources are in close proximity to bright satellite galaxies, this feature is potentially detrimental to our shear measurements. The cluster image simulations indeed contain a population of sources with large values of δm_{phot} , which is absent in the simulations without cluster galaxies. Comparing the simulations with- and without cluster galaxies we determined an empirical relation to flag any galaxies severely affected by the local background subtraction. We discard all source galaxies with

$$\delta m_{\text{phot}} < -49.04 - 7.00m_{\text{phot}} + 0.333m_{\text{phot}}^2 - 0.0053m_{\text{phot}}^3, \quad (5.10)$$

since these galaxies are outliers in the $\delta m_{\text{phot}} - m_{\text{phot}}$ plane. Inspecting the images of the galaxies thus discarded in the real data, we find that they are mostly located either near bright, saturated stars (but these galaxies would be discarded in subsequent steps by masking stellar spikes and ghosts), or close to big galaxies with resolved spiral arms or other features, that make the plane approximation of the background a bad description of the local background. We have verified that the calibration of the shape measurements by Hoekstra et al. (2015) remains unchanged when discarding these galaxies (which were included in their sample); this is essentially because Equation 5.10 is independent of galaxy shape. Typically, an additional 10–12 per cent of sources in the data are flagged by Equation 5.10.

5.4.2 Additive shear bias

In galaxy-galaxy lensing (and equivalently cluster lensing), source shapes are azimuthally averaged around the lenses. This washes out any spatial PSF anisotropy, and the additive bias c in Equation 5.9 can be neglected. (In other words, additive biases in γ_1 and γ_2 vanish when projected onto γ_t .) However, our measurements are focused on the immediate surroundings (tens of arcseconds to few arcminutes) of thousands of luminous lenses, such that galaxy light may bias the shape measurements of fainter background sources. Given that the light profile always decreases radially, the azimuthal averaging can in fact introduce an additive bias in γ_t (as opposed to $\gamma_{1,2}$) by biasing the background subtraction along the radial direction. We refer to this additive bias in γ_t as c_t hereafter.

We expect the bias to depend on the size and magnitude of cluster galaxies and therefore create image simulations to determine this relation. We selected a set of magnitudes and sizes representative of the full sample of cluster members (shown as cyan circles in Figure 5.2) and simulated lens galaxies with those properties. In order to accurately estimate c_t , we simulate large numbers of galaxies with the same magnitude and size, placed in a regular grid in the image simulations, separated by at least $60''$ to avoid overlap between the lenses. We refer to these simulations as ‘grid image simulations’. The PSF in these simulations is circular with a full width at half

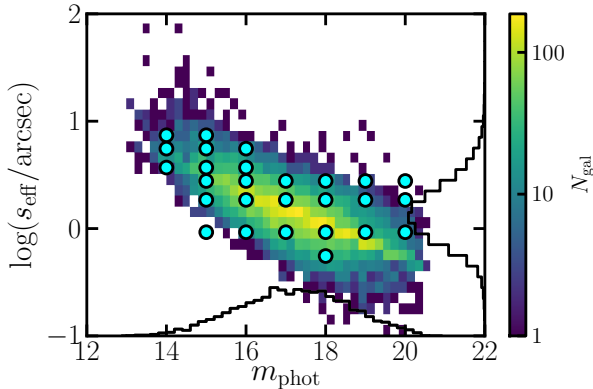


Figure 5.2: Magnitude and size distribution of satellites in the MENeACS spec+RS sample. The logarithmic color scale shows the number of galaxies per two-dimensional bin, while black histograms show the one-dimensional distributions. Cyan circles show the coordinates used in the grid image simulations used to determine the additive bias on the shape measurements in Section 5.4.2. Note that galaxies at the bottom-right corner of the distribution, not covered by the simulations, are faint and small and therefore can be safely assumed to produce no obscuration (see Section 5.A).

maximum of $0''.67$. We generate a large number of grid image simulations spanning a range of lens size and r -band magnitude and measure the average shear around these simulated lenses, which is by construction zero in the source-only image simulations.

In Section 5.A we show that we can model this (negative) bias as a function of lens-source separation, lens magnitude and size, and we correct the shear measured for each lens-source pair for this bias. For illustration, we show in Figure 5.3 the average $\Sigma_c c_t$ obtained from the image simulations after weighting the results in the simulations by the two-dimensional distribution of real galaxies in r -band magnitude and size, when binning MENeACS galaxies into five stellar mass bins (see Section 5.6). As expected, the correction is larger for more massive galaxies, which are on average larger. At $R \sim 20$ kpc (i.e., the smallest scales probed), the correction is 20–30 per cent and is typically negligible at $R \sim 50$ kpc. We find that on average c_t is approximately independent of cluster-centric distance, because there is no strong luminosity segregation of galaxies in clusters as massive as those in MENeACS (e.g., Roberts et al. 2015). For reference, a fraction of order 10^{-6} lens-source pairs have $|c_t| > 0.01$, which corresponds to the typical shear produced by massive cluster galaxies in our sample. We remove these lens-source pairs from our analysis, since such corrections are most of the time larger than the signal itself, although such a small fraction of lens-source pairs has no effect on our results. We find that lens galaxies with $s_{\text{eff}} < 1''$ produce no noticeable obscuration at the scales of interest (see Section 5.A), and we therefore did not produce simulations for the smallest lenses (Figure 5.2).

5.4.3 Contamination by cluster members

In addition to the additive bias discussed above, lens galaxies affect the source density in their vicinity for two reasons: big lenses act as masks on the background source population, while small ones enter the source sample. We refer to these effects as obscuration and contamination, respectively.

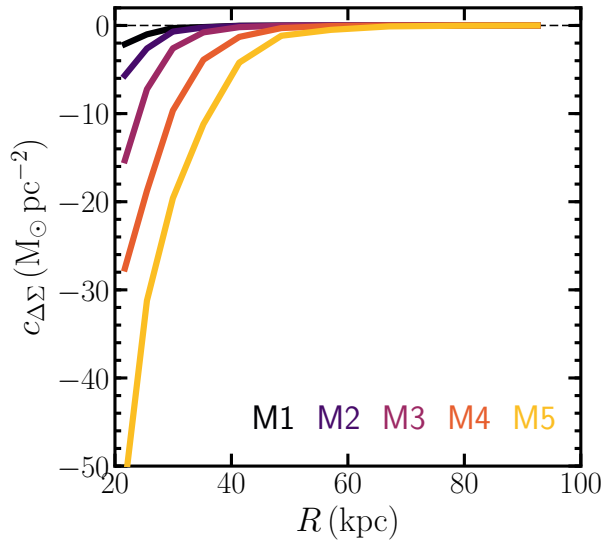


Figure 5.3: Average tangential additive bias, $c_{\Delta\Sigma} \equiv \Sigma_c c_t$, for the five stellar mass bins studied in Section 5.6, from low (M1) to high (M5) stellar mass (see Table 5.1). Note the smaller extent of the horizontal axis compared to other similar figures.

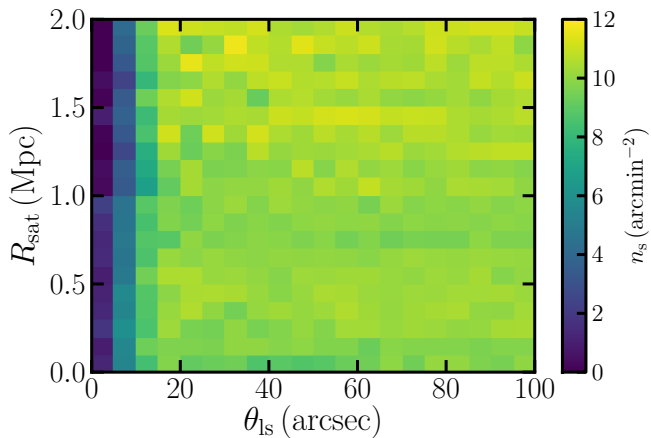


Figure 5.4: Observed number density of background sources as a function of lens-source separation, θ_{ls} , and distance from the lens to the cluster centre, R_{sat} , for all MENeACS clusters, after applying all the cuts described in Section 5.3.

Since on average cluster galaxies are randomly oriented (Sifón et al. 2015b), contamination by cluster members biases the (positive) lensing signal low; the correction for this effect is usually referred to as the ‘boost factor’ (e.g., Mandelbaum et al. 2005a). Obscuration, in turn, has two effects: it reduces the statistical power of small-scale measurements, and it complicates the determination of the contamination correction, since the observed source density is affected by obscuration. Figure 5.4 shows the number density of sources as a function of lens-source separation and cluster-centric distance. The obscuration of source galaxies is evident: the source density decreases rapidly at $\theta_{ls} \lesssim 10''$, while it remains essentially constant over the rest of the $\theta_{ls} - R_{\text{sat}}$ plane. The effect of contamination is not so readily seen (i.e., the source density is approximately constant for varying R_{sat}), because of the low redshift of our clusters: cluster galaxies are sufficiently separated on the sky that they do not appreciably boost the source density if obscuration is not accounted for.

Correcting the observed source density profile for obscuration

To measure the obscuration by cluster members, we generate a new set of image simulations, in which the spatial distribution of lens galaxies in the observations is reproduced and each lens galaxy is simulated with its measured properties. In this way a realistic simulation of each observed MENeACS cluster is created. We refer to these image simulations as ‘cluster image simulations’.

The cluster image simulations were designed to mimic the data as closely as possible to accurately capture the obscuration produced by MENeACS cluster members. We used the image simulation pipeline of Hoekstra et al. (2015) to create images of the source population with the same seeing and noise level measured from the data for each cluster. We then created images with the same properties, including a foreground cluster. Where available, we used the GALFIT measurements of Sifón et al. (2015b) to create surface brightness profiles for cluster members. For cluster members without reliable GALFIT measurements (which constitute approximately 10 per cent of the simulated cluster galaxies, and are mostly on the faint end of the population) we draw random values following the distribution of morphological parameters for galaxies with similar magnitude and redshift. Although individual galaxies may not be accurately represented in the simulations, the average obscuration should be well captured. We include all spectroscopic and red-sequence member galaxies down to an apparent magnitude $m_{\text{phot}} = 23$ and to $R_{\text{sat}} = 3$ Mpc. As shown by Sifón et al. (2015b), the red sequence is severely contaminated at such large distances. As we show below, this ‘interloper’ population can be easily accounted for, since the density of interlopers is not a function of cluster-centric distance.

We use the cluster image simulations to calculate the average obscuration produced by cluster galaxies by measuring the source density as a function of cluster-centric radius. Because in these simulations we reproduce the spatial distribution of cluster galaxies, we can account for the radial dependence of the obscuration, given the number density profile of cluster galaxies. We show in Figure 5.5 the average obscuration profile, defined as

$$\mathcal{F}_{\text{obsc}}(R_{\text{BCG}}) \equiv \frac{\xi_{\text{s,cluster}}(R_{\text{BCG}})}{\xi_{\text{s,background}}}, \quad (5.11)$$

where ξ_{s} is the source weight density, and the subscripts “cluster” and “background” refer to the image simulations with and without the cluster galaxies, respectively.

In fact, the obscuration at large cluster-centric distance is not exactly zero, but reaches a constant value $\hat{\mathcal{F}}(R_{\text{BCG}} > 1.5\text{Mpc}) \simeq 0.06$ (where the hat symbol simply de-

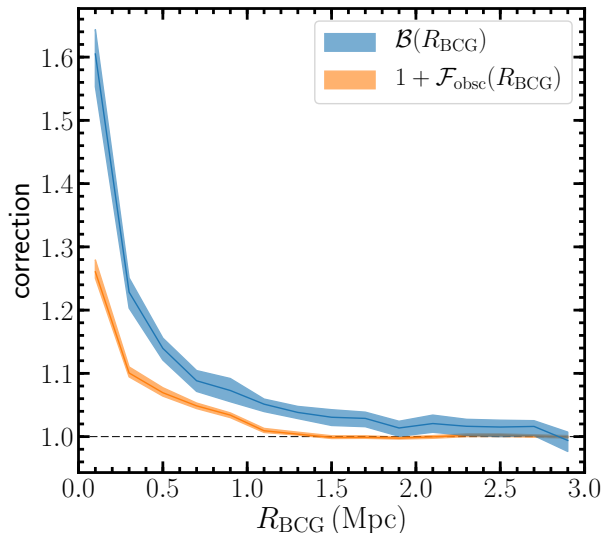


Figure 5.5: Obscuration correction (orange) and obscuration-corrected contamination correction (i.e., boost factor, blue) as a function of cluster-centric distance. Both quantities are averages over all clusters. The width of each curve shows the uncertainty on the mean correction.

notes a biased measurement of the true $\mathcal{F}(R)$). This is because, to ensure completeness, the image simulations include all red sequence galaxies, which inevitably includes a contaminating population of galaxies that are in fact not part of the cluster, especially at large R_{BCG} (Sifón et al. 2015b). We account for this excess obscuration by contaminating galaxies by simply subtracting the large-scale value of $\hat{\mathcal{F}}(R_{\text{BCG}})$, which results in the curve shown in Figure 5.5.

Boost correction

Because the source sample is both obscured and contaminated by cluster galaxies, we need an external measurement of the reference source density. Furthermore, because the bulk of our sample is at $z < 0.1$, the Megacam field of view is not enough to estimate cluster-free source densities—our images only reach $R_{\text{sat}} \sim 3$ Mpc at $z = 0.1$. Therefore, we retrieved data for 41 blank fields from the Megacam archive (Gwyn 2008), which provides an area of approximately 33 sq. deg. after manual masking. These blank fields contain no galaxy clusters and have noise and seeing properties at least as good as the MENeACS data. We construct the source sample and shape catalogue exactly as described above, after degrading the blank field observations to the typical noise level of MENeACS data (see Herbonnet et al. in prep.).

As described in Herbonnet et al. (in prep.), we fit the the blank field source weight densities, $\xi_{\text{s,blank}}$, as a linear combination of the image quality (quantified by the average half-light radius of stars, $\langle r_{\text{h}}^{\star} \rangle$), the background noise level, ζ , and the Galactic extinction in the r -band, A_r ,

$$\langle \xi_{\text{s,blank}} \rangle (\langle r_{\text{h}}^{\star} \rangle, \zeta, A_r) = p_1 \langle r_{\text{h}}^{\star} \rangle + p_2 \zeta + p_3 A_r + p_4, \quad (5.12)$$

where ζ , $\langle r_h^* \rangle$ and A_r are in units of counts per pixel, pixels, and magnitudes respectively, and $p_i = (-68.4, -40.6, -122.8, 364.2)$ are the best-fit parameters. The blank field measurements are well described by a normal distribution around Equation 5.12, with a constant scatter of 12 weight-units per sq. arcmin, as shown in Figure 5.6. We adopt the noise-, extinction-, and seeing-dependent source density measured in the blank fields as the background level for each of the MENeCS clusters. We have checked that at the high redshift end of our sample, the source densities at the outskirts of clusters ($R_{\text{BCG}} \gtrsim 3$ Mpc) are consistent with the expectations from the blank fields. The limiting factor to the precision of the blank field source density prediction is the number of blank fields. For the available 41 fields, the relative uncertainty in the blank field prediction is 1.0%, which is precise enough for our analysis.

Having computed the obscuration from the image simulations and the contamination by comparing with blank fields, we now calculate the boost correction appropriate to our dataset. Given a source's R_{BCG} , we calculate its corrected (or 'true') shear through Equation 5.9, where the boost correction is

$$\mathcal{B}(R_{\text{BCG}}) = \frac{\langle \xi_{s,\text{data}}(R_{\text{BCG}}) \rangle}{\langle \xi_{s,\text{blank}} \rangle} [1 - \mathcal{F}_{\text{obsc}}(R_{\text{BCG}})]^{-1}, \quad (5.13)$$

where all quantities are averaged over all clusters. Equation 5.13 assumes that faint cluster galaxies (which enter the source sample) do not cluster strongly with the bright cluster members constituting our lens sample; this small-scale cluster would introduce a dependence of \mathcal{B} on θ_{ls} . For reference, Fang et al. (2016) showed that there is an excess of galaxies in the vicinity of cluster members, but at the level of a few galaxies per cluster, which would have no impact on our results. In fact, we find no evidence of small-scale clustering in our sample of red sequence galaxies.

5.4.4 More details on obscuration by cluster members

In the previous section we calculated the average obscuration produced by cluster members as a function of cluster-centric distance, R_{BCG} , in order to properly estimate the boost correction. In this section, we look closer at the obscuration by cluster galaxies individually rather than collectively as part of the cluster.

We calculate obscuration profiles around galaxies, $\mathcal{F}_{\text{obsc}}(\theta_{\text{ls}})$, in bins of cluster centric distance, R_{BCG} ; we show these profiles in Figure 5.7. Because of the high *lens* density at small R_{BCG} , the obscuration drops only down to roughly 0.45 up to $\theta_{\text{ls}} \sim 50''$, decreasing slowly at larger separations. However, the effect of neighbouring lenses is negligible at $R_{\text{BCG}} \gtrsim 200$ kpc. Note that these obscuration profiles do not affect the calculation of the boost factor, because as mentioned in the preceding section the density of cluster galaxies does not depend on θ_{ls} . (Integration of this set of curves over θ_{ls} gives rise to Equation 5.13). Instead, the steep rise in the obscuration below $\theta_{\text{ls}} \approx 20$ arcsec fundamentally limits the scales accessible in this study. Pushing to smaller scales would require subtraction of the light profiles of lens galaxies, an avenue we will explore in future work.

5.4.5 Source redshift distribution

The measurement of the ESD is averaged over each lens source pair in the source population so that redshifts for individual sources are required. However, we lack the deep colour information to estimate reliable photometric redshifts for individual source galaxies. Instead, we can use an average lensing efficiency $\langle \beta \rangle = \langle \max[0, D_{\text{ls}}/D_s] \rangle$ for

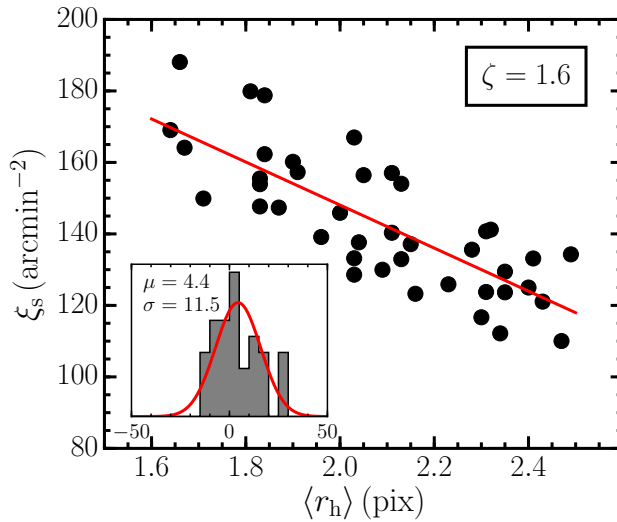


Figure 5.6: Total source weight density as a function of half-light radius of stars in each of the blank Megacam fields, with the background level artificially increased to 1.6 counts per pixel (the mean noise level in MENEaCS) and assuming no Galactic extinction for illustration. The red solid line shows the best-fitting function described by Equation 5.12. The inset shows a histogram of the residuals in ξ_s about the best-fit, with the best-fit Gaussian distribution in red, and the legend reports the mean (μ) and standard deviation (σ) of this distribution.

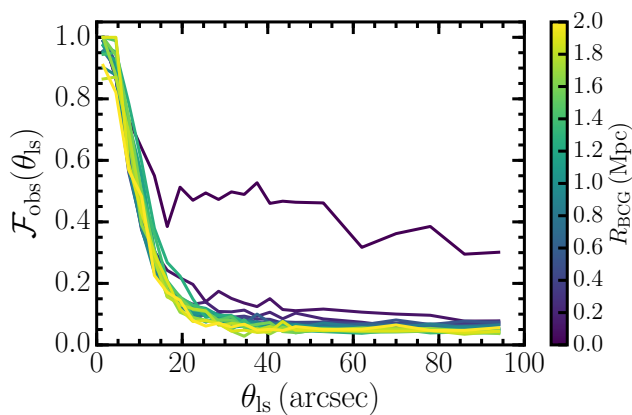


Figure 5.7: Obscuration profile measured in the image simulations as a function of lens-source separation, θ_{ls} , in bins of cluster-centric distance, R_{BCG} , averaged over all MENEaCS clusters.

the entire source population, which can be inferred from a representative field with a reliable redshift distribution, as a proxy for the cluster background (see, e.g., Hoekstra et al. 2015).

We take as a reference the COSMOS2015 catalogue (Laigle et al. 2016), which contains photometric redshift estimates for galaxies in the 2 square degrees COSMOS field. This catalogue is deep enough to cover our magnitude range and contains near infrared measurements that help break degeneracies in photometric redshift estimation. The COSMOS field was also targeted by a deep observation with the CFHT, from which there exists a lensing catalogue. The matched lensing-photometric redshift catalogue allows us to apply the same quality cuts on the redshift distribution as have been applied to the lensing data, which could otherwise bias the results (Gruen & Brimiouille 2016). The overlapping area is only 1 square degree, which raises concerns that it might be unrepresentative for our cluster fields. However, we have used additional photometric data of different areas on the sky to confirm that the uncertainties on our mean lensing efficiency, $\langle\beta\rangle$, including cosmic variance, are less than 2 per cent. Such precision is sufficient for our analysis. For more details see Herbonnet et al. (in prep.).

The assumption of using only the average value $\langle\beta\rangle$ and ignoring the width of the distribution introduces a bias into our measurement of $\Delta\Sigma$ (Hoekstra et al. 2000). However, for our low redshift clusters the effect is expected to be very small. With our photometric redshift catalogue and Equation 7 from Hoekstra et al. (2000) we estimate that this bias is at most $1+0.06\kappa$ (where κ is the lensing convergence, and $\kappa \ll 1$ in the weak lensing regime). This is a negligible bias compared to other sources of uncertainty and we therefore ignore it in the rest of our analysis.

5.4.6 Resulting lensing signal

Figure 5.8 shows the resulting lensing signal from satellites in MENeCS clusters, corrected for both $c_1(\theta_s)$ and $\mathcal{B}(R)$. We make the distinction in the arguments of both corrections because the former is applied to each lens-source pair, while the latter is applied as an average correction after stacking all lenses in each bin. We compare the ESDs of the five bins in satellite stellar mass for the spec and spec+RS samples. There are two differences in the signal measured for both samples. Firstly, the signal from the spec+RS sample is slightly lower than the signal from the spec sample at the smallest scales. This is expected, as in general the more massive galaxies have been targeted in the spectroscopic observations; this is reflected also in the average stellar masses listed in Table 5.1. Secondly, the spec+RS signal is larger at intermediate scales, which is a reflection of the fact that spectroscopic observations tend to be incomplete at the dense centres of clusters, so the average cluster-centric distance of the spec+RS sample is lower. We base our analysis on the spec+RS sample, which is a more complete sample of lenses.

At intermediate scales, $0.3 \lesssim R/\text{Mpc} \lesssim 2$, the two samples produce different signals. In particular, the signal from the spec+RS sample is higher. This is a consequence of the fact that we only include red sequence galaxies out to 1 Mpc, so the spec+RS sample is on average closer to the cluster centre than the spec sample. Therefore, the peak of the host cluster signal happens at smaller R . Beyond the peak the two signals are consistent, because all galaxies come from the same clusters. See Figure 3 of Sifón et al. (2015a) for a graphical representation. We account for the measured radial distribution of satellites in our modelling below.

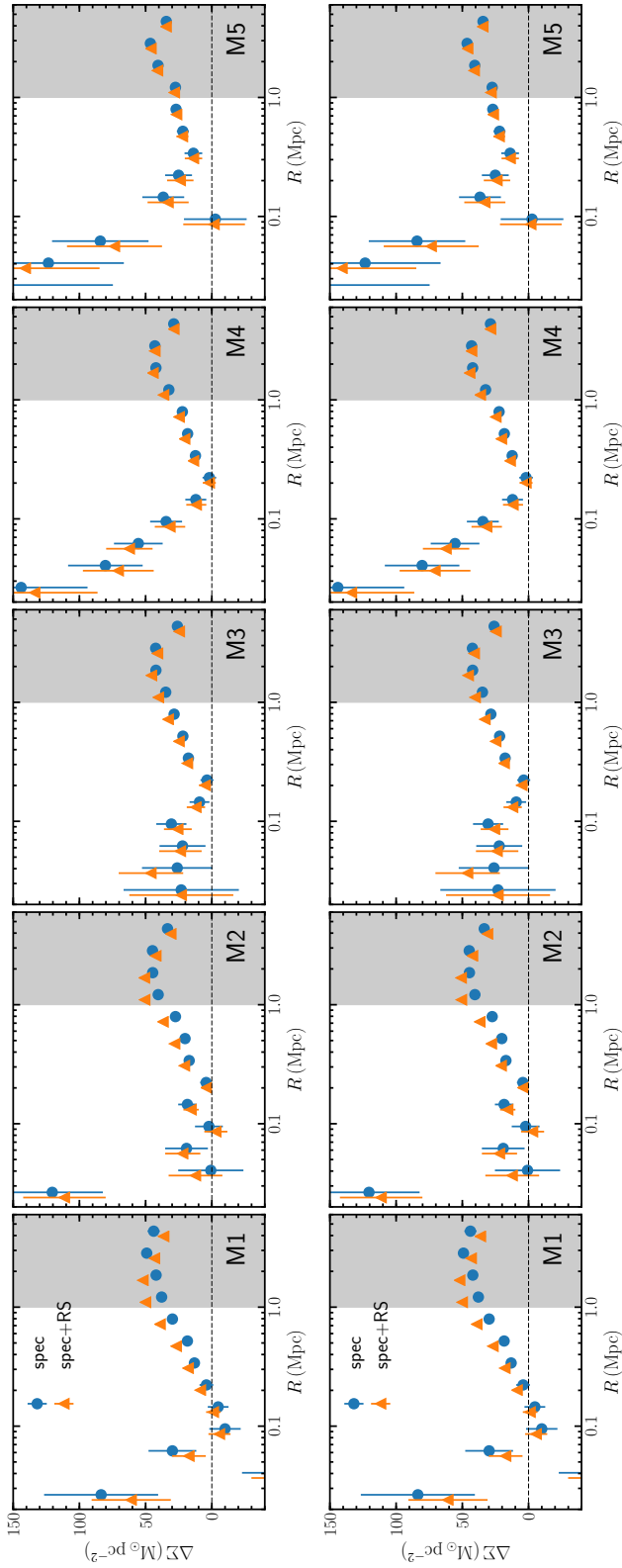


Figure 5.8: Excess surface mass density (ESD) of satellite galaxies binned by stellar mass. Blue circles and orange triangles show the ESD of the spectroscopic and spectroscopic-plus-red sequence samples, respectively. Errorbars are the square roots of the diagonal terms of the covariance matrices. The dashed horizontal line shows $\Delta\Sigma = 0$ for reference. In our analysis we only use data points up to 1 Mpc, shown over a white background.

5.5 Satellite galaxy-galaxy lensing model

We interpret the galaxy-galaxy lensing signal produced by subhaloes following the formalism introduced by Yang et al. (2006, see also Li et al. 2013a), and applied to observations by Li et al. (2014, 2016); Sifón et al. (2015a, 2017) and Niemiec et al. (2017). This formalism assumes that measurements are averages over a large number of satellites *and* clusters, such that the stacked cluster is (to a sufficient approximation) point-symmetric around its centre and well-described by a given parametrization of the density profile. A similar method was introduced by Pastor Mira et al. (2011), which however does not rely on such parametrization by virtue of subtracting the signal at the opposite point in the host cluster. A different approach is to perform a maximum likelihood reconstruction of the lensing potential of cluster galaxies accounting for the cluster potential, which must be well known a priori (e.g., Natarajan & Kneib 1997; Geiger & Schneider 1998) or modelled simultaneously with the cluster galaxies (Limousin et al. 2005). This method has been applied in several observational studies (e.g., Natarajan et al. 1998, 2009; Limousin et al. 2007). We discuss results from the literature using either method after presenting our analysis, in Section 5.7. In the following we describe our modelling of the satellite galaxy-galaxy lensing signal.

The ESD measured around a satellite galaxy is a combination of the contributions from the subhalo (including the galaxy itself) at small scales, and that from the host halo at larger scales,

$$\Delta\Sigma_{\text{sat}}(R) = \Delta\Sigma_{\star}(R|m_{\star}) + \Delta\Sigma_{\text{sub}}(R|m_{\text{bg}}, c_{\text{sub}}) + \Delta\Sigma_{\text{host}}(R|M_{\text{h}}, c_{\text{h}}), \quad (5.14)$$

where $\Delta\Sigma_{\star}$ represents the contribution from baryons in the satellite galaxy, which we model as a point source contribution throughout, such that

$$\Delta\Sigma_{\star}(R|m_{\star}) = \frac{m_{\star}}{\pi R^2}. \quad (5.15)$$

Here, we take m_{\star} to be the median stellar mass of all satellites in the corresponding sample (e.g., a given bin in satellite luminosity). In Equation 5.14, R refers to the lens-source separation in physical units; m_{bg} is the average subhalo mass (see below) and c_{sub} its concentration; and M_{h} and c_{h} are the average mass and concentration of the host clusters. In the remainder of this section we describe the other two components in Equation 5.14. Detailed, graphical descriptions of these components can be found in Yang et al. (2006), Li et al. (2013a) and Sifón et al. (2015a).

5.5.1 Host cluster contribution

In numerical simulations, the density profiles of dark matter haloes are well described by a Navarro-Frenk-White (NFW, Navarro et al. 1995) profile,

$$\rho_{\text{NFW}}(r) = \frac{\delta_c \rho_m}{r/r_s (1 + r/r_s)^2}, \quad (5.16)$$

where $\rho_m(z) = 3H_0^2(1+z)^3\Omega_m/(8\pi G)$ is the mean density of the Universe at redshift z and

$$\delta_c = \frac{200}{3} \frac{c^3}{\ln(1+c) - c/(1+c)}. \quad (5.17)$$

The two free parameters, r_s and $c \equiv r_{200}/r_s$, are the scale radius and concentration of the profile, respectively. Stacked weak lensing measurements have shown that this

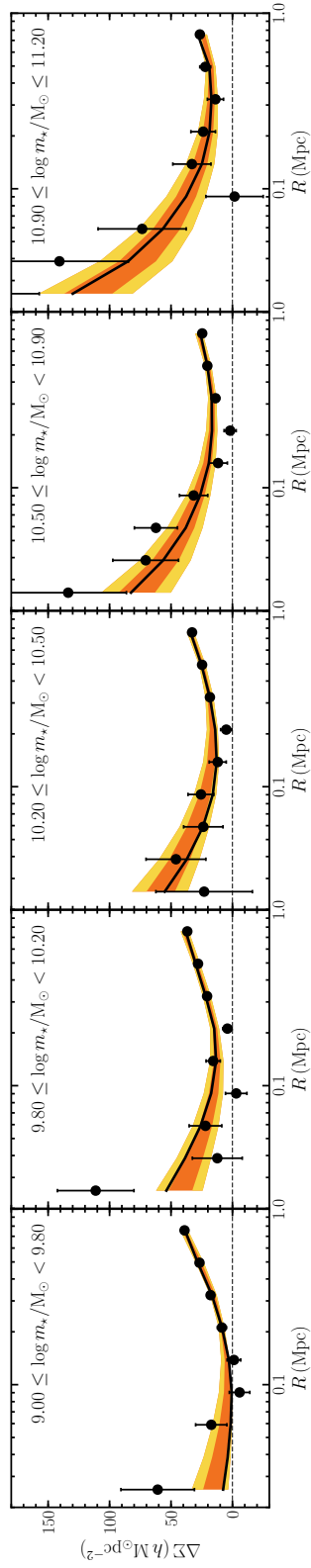


Figure 5.9: Excess surface mass density of the spec+RS sample, binned by stellar mass as shown in the legends (same as the magenta triangles in Figure 5.8). The black line shows the best-fitting model from the MCMC and the orange and yellow regions outline the 68 and 95 per cent credible intervals.

theoretical profile is a good description, on average, of real galaxy clusters as well (Oguri et al. 2012; Umetsu et al. 2016). We therefore adopt this parametrization for the density profile of the host clusters.

The concentration parameter is typically anti-correlated with mass. This relation, referred to as $c(M)$ hereafter, has been the subject of several studies (e.g., Bullock et al. 2001; Duffy et al. 2008; Macciò et al. 2008; Prada et al. 2012; Dutton & Macciò 2014). Most of these studies parametrize the $c(M)$ relation as a power law with mass (and some with redshift as well), with the mass dependence being typically very weak. Since our sample covers relatively narrow ranges in both quantities (i.e., cluster mass and redshift), the exact function adopted is of relatively little importance. We therefore parametrize the mass-concentration relation as a power law with mass,

$$c_h(M_{200,h}) = a_c \left(\frac{M_{200,h}}{10^{15} M_\odot} \right)^{b_c} \quad (5.18)$$

where $M_{200,h}$ is the host halo mass within $r_{200,h}$, and a_c and b_c are free parameters. We follow Sifón et al. (2015a) and account for the observed separations between the satellites and the cluster centre (which we assume to coincide with the BCG) in each observable bin to model the total host halo contribution to Equation 5.14.

5.5.2 Subhalo contribution

Although in numerical simulations satellite galaxies are heavily stripped by their host cluster, the effect on their density profile is not well established. For instance, Hayashi et al. (2003) found that, although tidal stripping removes mass in an outside-in fashion, tidal heating causes the subhalo to expand, and the resulting density profile is similar in shape to that of a central galaxy (which has not been subject to tidal stripping). Similarly, Pastor Mira et al. (2011) found that the NFW profile is a better fit than truncated profiles for subhaloes in the Millenium Simulation (Springel et al. 2005), and that the reduction in mass produced by tidal stripping is simply reflected as a change in the NFW concentration of subhaloes, which have roughly a factor 2–3 higher concentration than host haloes, consistent with the mass-concentration relation for subhaloes derived by Moliné et al. (2017) from N-body simulations. Moliné et al. (2017) further showed that the subhalo concentration depends on cluster-centric distance, with subhaloes closer in having a larger concentration as a result of the stronger stripping.

We therefore assume that the density profile of subhaloes can also be described by an NFW profile. We adopt the subhalo mass-concentration relation derived by Moliné et al. (2017), which depends on both the subhalo mass and its halo-centric distance,

$$c_{\text{sub}}(m_{200}, x) = c_0 \left(1 + \sum_{i=1}^3 \left[a_i \log \left(\frac{m_{200}}{10^8 h^{-1} M_\odot} \right) \right]^i \right) \times [1 + b \log x], \quad (5.19)$$

where $x \equiv r_{\text{sat}}/r_{h,200}$ (defined in three-dimensional space), $c_0 = 19.9$, $a_i = \{-0.195, 0.089, 0.089\}$ and $b = -0.54$.

Note that the quantity m_{200} is used for mathematical convenience only, but is not well defined physically. Instead, we report subhalo masses within the radius at which the subhalo density matches the background density of the cluster at the distance of the subhalo in question (which we denote r_{bg}), and refer to this mass as m_{bg} . This

radius r_{bg} scales roughly with cluster-centric distance as $r_{\text{bg}} \propto (R_{\text{sat}}/r_{200,\text{h}})^{2/3}$ (see also Natarajan et al. 2007, for a comparison between m_{bg} and m_{200}). The reported subhalo masses are therefore similar to those that would be measured by a subhalo finder based on local overdensities such as SUBFIND (Springel et al. 2001), which allows us to compare our results with numerical simulations consistently.

Because the density profile is a steep function of cluster-centric distance, we take the most probable three-dimensional cluster-centric distance, $\langle r_{\text{sat}} \rangle$, to be equal to the weighted average of the histogram of two-dimensional distances, R_{sat} :

$$\langle r_{\text{sat}} \rangle = \frac{\sum_i n(R_{\text{sat},i}) R_{\text{sat},i}}{\sum_i n(R_{\text{sat},i})}, \quad (5.20)$$

where the index i runs over bins of width $\Delta R_{\text{sat}} = 0.1$ Mpc (see Figure 5.1). We use this $\langle r_{\text{sat}} \rangle$ in Equation 5.19.

5.5.3 Fitting procedure

We fit the model presented above to the data using the affine-invariant Markov Chain Monte Carlo (MCMC) ensemble sampler EMCEE (Foreman-Mackey et al. 2013). This sampler uses a number of walkers (set here to 5000) which move through parameter space depending on the position of all other walkers at a particular step, using a Metropolis Hastings acceptance criterion (see Goodman & Weare 2010, for a detailed description). The loss function to be maximized is defined as

$$\mathcal{L} = \frac{1}{(2\pi)^{k^2/2}} \prod_{m=1}^k \prod_{n=1}^k \frac{1}{\sqrt{\det(\mathbf{C}_{mn})}} \times \exp \left[-\frac{1}{2} (\mathbf{O} - \mathbf{E})_m^T \mathbf{C}_{mn}^{-1} (\mathbf{O} - \mathbf{E})_n \right], \quad (5.21)$$

where $k = 5$ is the number of bins into which the sample is split (in stellar mass or cluster-centric distance bins); \mathbf{O} and \mathbf{E} are the observational data vector and the corresponding model predictions, respectively; \mathbf{C} is the covariance matrix; $\det(\cdot)$ is the determinant operator; and the index pair (i, j) runs over data points in each bin (m, n) . As implied by Equation 5.21, we account for the full covariance matrix, including elements both within and between observable bins.

We quote the prior ranges and marginalized posterior central values and 68 per cent uncertainties for all free parameters in our model in Table 5.2, both when binning by stellar mass and by cluster-centric distance (each discussed in Sections 5.6 and 5.7, respectively). Although we quote parameters of host clusters, we treat them as nuisance parameters throughout. Note that priors are defined in real space, and are only quoted as logarithmic quantities in Table 5.2 for convenience. As a result, when poorly constrained by the data, posterior host cluster masses are unrealistically high. For guidance, the values in Table 5.2 can be compared to dynamical masses and weak lensing masses reported for the same clusters by Sifón et al. (2015b) and Herbonnet et al. (in prep.), which suggest an average cluster mass $M_{200,\text{h}} \sim 6 \times 10^{14} M_{\odot}$.

5.6 The subhalo-to-stellar mass relation

We first bin the sample by stellar mass, as shown in the top-left panel of Figure 5.1. The ESD of the five bins, along with the best-fit model, are shown in Figure 5.9. The

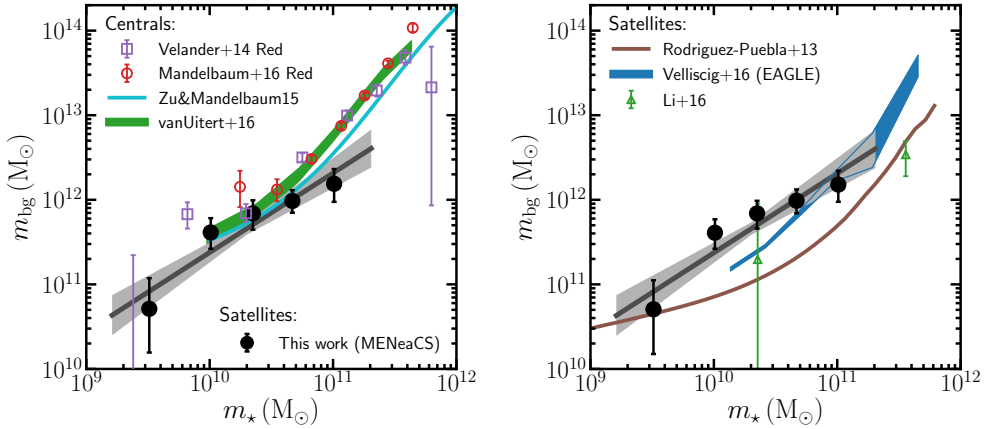


Figure 5.10: Stellar-to-subhalo mass relation. Big black circles in both panels correspond to the best-fit subhalo masses of spectroscopic plus red sequence satellites, assuming the subhalo mass-concentration relation of Moliné et al. (2017). The grey line and shaded regions show the best-fit linear relation using the BCES $X_2|X_1$ estimator and the 68 per cent confidence interval on the fit, respectively. Subhalo masses refer to the mass within r_{bg} (see Section 5.5.2). The left panel shows for comparison the stellar-to-halo mass relations of central galaxies (where halo mass refers to $M_{200,h}$) from galaxy-galaxy lensing measurements by van Uitert et al. (2016) and specifically of *red* central galaxies by Velander et al. (2014) and Mandelbaum et al. (2016), plus the relation from combined lensing and clustering measurements by Zu & Mandelbaum (2015). The right panel shows measurements of subhalo masses as a function of stellar mass by Li et al. (2016) (green triangles), in addition to the subhalo-to-stellar mass relations for satellite galaxies in the EAGLE simulation (Velliscig et al. 2016), and from abundance matching applied to galaxy clustering measurements by Rodríguez-Puebla et al. (2013) as a cyan band (with the width corresponding to the error on the mean) and a brown line, respectively.

best-fit masses resulting from this model are shown in both panels of Figure 5.10. We fit a power law relation⁶ between subhalo and stellar masses using the BCES $X_2|X_1$ estimator, an extension of least squares linear regression which accounts for measurement uncertainties on both variables (although here we neglect uncertainties on the average stellar masses) and intrinsic scatter (Akritas & Bershady 1996), and find an approximately linear relation,

$$\frac{m_{\text{bg}}}{M_{\odot}} = 10^{11.66 \pm 0.07} \left(\frac{m_{\star}}{2 \times 10^{10} M_{\odot}} \right)^{0.94 \pm 0.15}. \quad (5.22)$$

We remind the reader that this relation applies to the subhalo mass, m_{bg} , within the radius r_{bg} where the subhalo density equals the host halo background density. If we replace m_{bg} with m_{200} , the normalization increases by a factor 2.8, while the best-fit slope is 0.96 ± 0.14 , indistinguishable from that reported in Equation 5.22. We also tested that varying the concentration of subhaloes by 20 per cent does not change the slope of the SHSMR.

5.6.1 The SHSMR in the context of the total-to-stellar mass relation of central galaxies

We also show in the left panel of Figure 5.10 various determinations of the relation between total and stellar mass of central galaxies from the literature (Velandar et al. 2014; Zu & Mandelbaum 2015; Mandelbaum et al. 2016; van Uitert et al. 2016), where halo mass refers to M_{200} .⁷ These have all been determined with weak lensing measurements (in combination with measurements of galaxy clustering and the stellar mass function in the cases of Zu & Mandelbaum 2015; van Uitert et al. 2016, respectively), and are broadly consistent with each other. Both Velandar et al. (2014) and Mandelbaum et al. (2016) divided their samples into red and blue centrals, and we only show their results for red galaxies since MENeACS satellites are in their great majority red as well. Indeed, small differences between some of these determinations are probably driven by the different galaxy samples used in each study, and are not relevant for the present discussion. In particular, Mandelbaum et al. (2016) found good agreement with the model of Zu & Mandelbaum (2015) once the galaxy samples are matched between the two studies.

The comparison between the central total-to-stellar mass relation and the satellite SHSMR is however not straightforward. In principle, we may consider in the case of central galaxies that $M_{\text{bg}} = M_{200}$, so at least the mass definitions can be regarded consistent. However, identifying the progenitors of present-day satellites is not an easy task, as there is evidence that most satellites in massive clusters today were part of smaller groups long before entering their current hosts. In the context of the decreased star formation of satellite galaxies, this is usually referred to as ‘pre-processing’ (e.g., McGee et al. 2009; Gabor & Davé 2015; Haines et al. 2015). The impact of this pre-processing on the total mass content of present-day satellites is unknown. Nevertheless, we can make a phenomenological comparison. We find that at $\log m_{\text{bg}} \lesssim 10.3$ the shape of the SHSMR coincides with that of the analogous relation for central galaxies, consistent with the prediction that galaxies lose dark matter more easily than stellar matter (e.g., Chang et al. 2013). We also cautiously note an increased difference in

⁶Our choice of a single power law to model the SHSMR is motivated only by our limited statistics.

⁷We scale all these relations to both the value of H_0 and the definition of halo mass—that is, M_{200} defined with respect to the average Universal density—adopted in this paper.

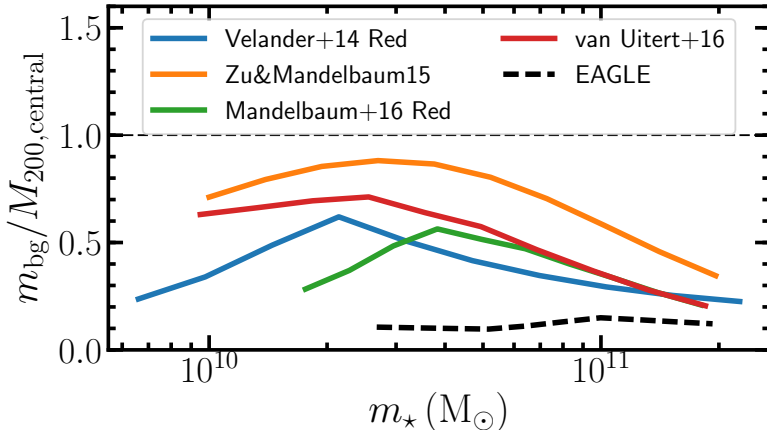


Figure 5.11: Ratio of subhalo masses measured in this work to central halo masses from different observational studies (solid lines) and measured in the EAGLE simulation by Velliscig et al. (2016) (black dashed line), at fixed stellar mass. The ratios for Velander et al. (2014) and Mandelbaum et al. (2016) are linearly interpolated from the data points. Note that statistical uncertainties dominate measurements at $m_\star \lesssim 2 \times 10^{10} M_\odot$. See Figure 5.10 for a more detailed description of the different measurements.

the SHSMR with the relation for centrals at high masses, $m \gtrsim 10^{11} M_\odot$, which may suggest that massive satellites lose dark matter more easily than lower-mass satellites, compared to their stellar mass loss. This may for instance be related to the fact that dynamical friction pulls more massive satellites toward the centre more efficiently (compared to lower mass galaxies), where tidal forces are stronger. If stellar mass is more resistant to tidal stripping then we would expect satellite galaxies with larger stellar masses to have a lower total-to-stellar mass ratio, as observed.

Velander et al. (2014) and Mandelbaum et al. (2016) constrained the total-to-stellar mass relation of central galaxies making use of a mixture of central and satellite galaxies, assuming that subhaloes have lost approximately half of their mass since being accreted onto the clusters. They achieve this by truncating the NFW density profiles of subhaloes at $r_t = 0.4r_{200}$ (Mandelbaum et al. 2005b), but they do not fit for any parameter relating to the subhalo contribution to their signal except for the fraction of satellite galaxies (which are hosted by subhaloes). We show in Figure 5.11 the ratio between our subhalo masses (more precisely, of the power-law fit to them) and the central halo masses shown in the left panel of Figure 5.10. For red galaxies (Velander et al. 2014; Mandelbaum et al. 2016) the ratio reaches a maximum of 0.5-0.6 at $m_\star = (2 - 4) \times 10^{10} M_\odot$, but quickly drops for both lower and higher stellar masses, reaching ~ 20 per cent at the low- and high-stellar mass ends probed here, although we note that uncertainties are significant at the low mass end (cf. Figure 5.10). If we interpret this as the fraction of mass that has not been lost to tidal stripping (that is, we assume we can directly compare present-day satellites to present-day centrals), then this impacts the total-to-stellar mass relation of central galaxies when centrals and satellites are not separated a priori, as done by van Uiter et al. (2016). As discussed by Velander et al. (2014), the effect of this increased fraction of stripped material on the inferred halo masses is to reduce halo masses, by an amount that depends on both the level of stripping and the satellite fraction in the sample. Since the satellite

Table 5.2: Prior ranges and marginalized posterior estimates of best-fitting parameters of the satellite lensing model. Masses are in units of M_{\odot} . Uncertainties correspond to 68 per cent credible intervals. All parameters have flat priors in the quoted ranges. Note that priors are defined in linear, rather than logarithmic, space. The binning schemes are summarized in Table 5.1.

Parameter	Prior range	m_{\star} bins (Section 5.6)	R_{sat} bins (Section 5.7)
$\log\langle m_{\text{bg},1}\rangle$	[7, 14]	$10.64^{+0.35}_{-0.60}$	$10.88^{+0.27}_{-0.37}$
$\log\langle m_{\text{bg},2}\rangle$	[7, 14]	$11.58^{+0.16}_{-0.18}$	$11.60^{+0.12}_{-0.13}$
$\log\langle m_{\text{bg},3}\rangle$	[7, 14]	$11.80^{+0.15}_{-0.17}$	$11.60^{+0.18}_{-0.19}$
$\log\langle m_{\text{bg},4}\rangle$	[7, 14]	$11.98^{+0.12}_{-0.13}$	$11.19^{+0.30}_{-0.44}$
$\log\langle m_{\text{bg},5}\rangle$	[7, 14]	$12.14^{+0.16}_{-0.22}$...
$a_{c,h}$	[0, 10]	$6.2^{+0.6}_{-0.7}$	$5.1^{+1.1}_{-1.3}$
$b_{c,h}$	[-1, 1]	$-0.82^{+0.15}_{-0.11}$	$-0.48^{+0.25}_{-0.22}$
$\log\langle M_{h,1}\rangle$	[13, 16]	$15.19^{+0.29}_{-0.18}$	$15.10^{+0.52}_{-0.34}$
$\log\langle M_{h,2}\rangle$	[13, 16]	$15.00^{+0.14}_{-0.11}$	$15.62^{+0.19}_{-0.17}$
$\log\langle M_{h,3}\rangle$	[13, 16]	$15.23^{+0.28}_{-0.17}$	$15.68^{+0.24}_{-0.60}$
$\log\langle M_{h,4}\rangle$	[13, 16]	$15.03^{+0.32}_{-0.13}$	$15.68^{+0.23}_{-0.42}$
$\log\langle M_{h,5}\rangle$	[13, 16]	$15.69^{+0.24}_{-0.30}$...

fraction can be high at low stellar masses (e.g., > 40 per cent at $m_{\star} < 10^{10} M_{\odot}$ in Velander et al. 2014), a stripping of 80 per cent of the mass of subhaloes could have an appreciable effect. Exactly how much of an effect that is will also depend on the effect the stripping has on the density profile, however, and is not easy to quantify in advance. At the very least, our results should inform systematic uncertainty budgets for estimations of the total-to-stellar mass relation of central galaxies when the sample of lenses includes satellite galaxies as well.

Irrespective of whether present-day satellites are at all comparable to present-day centrals, the solid lines shown in Figure 5.11 represent a direct, quantitative prediction for hydrodynamical simulations. As shown in Figure 5.11, the same ratio measured in the EAGLE simulation yields significantly different results, with a ratio $m_{\text{bg}}/M_{200,\text{central}} < 0.1$ for all stellar masses $m_{\star} < 2 \times 10^{11} M_{\odot}$. This suggests that subhalo masses are low in EAGLE. We discuss possible origins for this discrepancy in the next section.

5.6.2 Comparison to other subhalo measurements and predictions

In the right panel of Figure 5.10, we compare our measurements to a previous measurement of subhalo mass as a function of stellar mass by Li et al. (2016). We also compare to determinations of the subhalo-to-stellar mass relation in low-mass galaxy clusters ($M_h \lesssim 10^{14} M_{\odot}$) from measurements in the EAGLE simulation (Crain et al. 2015; Schaye et al. 2015) devised to match the satellite sample of Sifón et al. (2015a) by Velliscig et al. (2016), and from a combination of galaxy clustering measurements and abundance matching predictions by Rodríguez-Puebla et al. (2013)⁸.

⁸Rodríguez-Puebla et al. (2013) used their measurements to fit for $m_{\star}(m)$, which we invert by Monte Carlo-sampling their relation accounting for the subhalo mass function at the time of infall

The measurements by Li et al. (2016) are consistent with our results within their comparatively large errorbars. It is worth noting that the high-stellar mass measurement by Li et al. (2016) supports our tentative detection of a difference between the SHSMR and the relation for central galaxies at the high mass end of subhaloes, $m \gtrsim 10^{11} M_{\odot}$. Their measurement is in fact consistent with a simple extrapolation of Equation 5.22. However, this comparison should be taken with care, as both the adopted density profiles and the mass definitions are different between us and Li et al. (2016).

Combining our measurements with those of Li et al. (2016), we find that satellite galaxies in EAGLE have a steeper SHSMR than the observations require. This may be due to the efficiency of tidal stripping implemented in the simulations, but it is likely that this is also influenced by technical differences such as assumptions about the density profiles and mass definitions. The EAGLE simulation was calibrated to measurements of the stellar mass function at $z = 0.1$ assuming the same Chabrier (2003) IMF, and subhaloes and galaxies are identified, and their masses calculated, using SUBFIND, so in terms of definitions the comparison with our work seems consistent. However, Knebe et al. (2011) have shown that the accuracy of subhalo masses estimated by SUBFIND depends significantly on halo-centric distance. Given that more massive satellites are generally located further out than less massive satellites (conversely, the average stellar mass at $R_{\text{sat}} \sim 0.2 \text{ Mpc}$ is roughly 60 per cent that at $R_{\text{sat}} \sim 1.5 \text{ Mpc}$ cf. Table 5.1), this bias in SUBFIND might exacerbate real differences at low masses somewhat. This should not, however, strongly affect the high mass end. We also note that Velliscig et al. (2016) have shown that, on average, the excess surface density around subhaloes in EAGLE is consistent with lensing measurements of observed satellite galaxies, but the *number* of subhaloes per host halo in EAGLE can appear inconsistent with the observations if the selection function is not carefully accounted for. It is plausible that a combination of biases in SUBFIND and inconsistent satellite fractions might explain the observed difference.

Similarly, the abundance matching-based measurement of Rodríguez-Puebla et al. (2013) is significantly lower than our measurements at stellar masses $m_{\star} \gtrsim 5 \times 10^9 M_{\odot}$. At face value, this suggests that subhalo abundance matching does not capture the relation between total and stellar mass properly. The existence of *assembly bias*—the hypothesis that the correspondence between halo mass and stellar mass depends on halo formation history (Gao et al. 2005)—would indeed mean that this is the case, but the extent of this problem is not well determined. Much work is devoted these days to the characterization and modelling of assembly bias (e.g., Hearin et al. 2016), and future work may be able to determine the role of assembly bias, or other effects, in this comparison.

5.7 Subhalo mass segregation

In this section we explore the dependence of subhalo mass on the distance to the cluster centre. van den Bosch et al. (2016) have shown using N-body simulations that, although subhalo mass segregation is seen more strongly in three dimensions, the projected halo-centric distance still preserves some of the correlation of subhalo physical parameters with the binding energy, which is closely related to the time a subhalo has spent bound to the host halo. However, after multiple orbits the correlation is significantly reduced because at any particular (projected) distance from the halo centre

from van den Bosch et al. (2016), including intrinsic scatter, and binning the data points by m_{\star} .

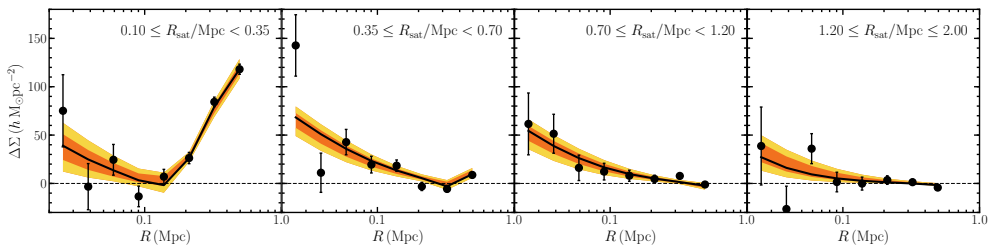


Figure 5.12: Excess surface mass density (black points with errorbars) and best-fit NFW model (black line) of satellites binned by cluster-centric distance. Blue lines show best-fit models while orange and yellow regions show 68 and 95 per cent credible intervals. The model for the host clusters is not flexible enough because the small field of view (in physical units at $z \sim 0.05$) biases our large-scale ($R \gtrsim 0.5$ Mpc) measurements, but this has no impact on the modelling of the subhalo signal. See Section 5.7 for details.

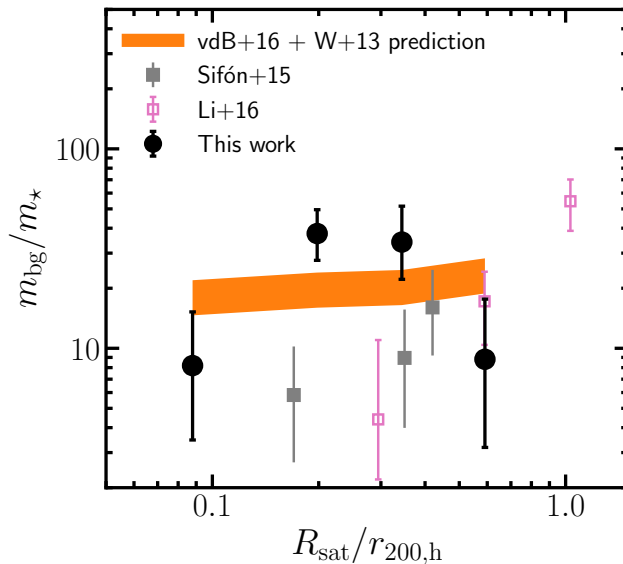


Figure 5.13: Best-fit subhalo-to-stellar mass ratio as a function of projected distance to the cluster centre, in units of r_{200} of the host cluster assuming $M_{200,h} = 6 \times 10^{14} M_{\odot}$ (see text). Uncertainties show 68 per cent credible intervals. As in Figure 5.10, subhalo masses refer to the mass within r_{bg} . Also shown are previous measurements from Sifón et al. (2015a) and Li et al. (2016). The orange band shows a prediction for the total-to-stellar mass ratio as a function of R_{sat} from numerical simulations from van den Bosch et al. (2016), which give $m/m_{\text{acc}}(R_{\text{sat}})$ (where m_{acc} is the subhalo mass at the time at accretion), combined with the semi-analytic $m_{\text{acc}}(m_{\star})$ predictions from Wang et al. (2013) for the median stellar masses in the five R_{sat} bins (cf. Table 5.1). For illustration purposes, the orange band has a width of 20 per cent.

there are subhaloes with a wide range of infall times. We might therefore expect satellites at similar R_{sat} to have been part of similar halo-subhalo interactions on average, but with a large scatter in their individual histories.

Figure 5.12 shows the measured ESD and best-fit model when we split the satellite sample into four R_{sat} bins. Because of the finite field of view of our observations, we cannot average galaxy shapes in full annuli with large lens-source separations around most lenses, so additive biases do not cancel out. For this reason, in this section we only use measurements out to lens-source separations $R = 0.6$ Mpc. At larger separations the signal is dominated by the host clusters, with little to no contribution from the subhaloes, and we have verified that subhalo masses are not affected by this cut.

5.7.1 Comparison to theoretical predictions

van den Bosch et al. (2016) have shown that the parameter that correlates most strongly with both binding energy and halo-centric distance is the ratio m/m_{acc} , where m_{acc} is the mass of the subhalo at the time of its accretion onto the main halo. This is because of the average relation between time a subhalo has spent in the host halo (or the accretion redshift, for a given redshift of observation) and the subhalo's distance to the halo centre, combined with the strong dependence of the mass ratio to the time since accretion as a result of tidal stripping.

We show the posterior subhalo masses, normalized by the median stellar mass in each bin, in Figure 5.13. In order to compare with literature measurements and predictions, we normalize cluster-centric distances by cluster radius $r_{200,h}$. Following the discussion in Section 5.5.3, we adopt a mean cluster mass $M_{200,h} = 6 \times 10^{14} M_{\odot}$ suggested by galaxy velocity dispersions and weak lensing measurements by Sifón et al. (2015b) and Herbonnet et al. (in prep.), respectively, instead of the posterior masses reported in Table 5.2. We find no statistically significant evidence for subhalo mass segregation, with an overall average ratio $\langle m_{\text{bg}}/m_{\star} \rangle = 21.5^{+6.3}_{-5.5}$.

We also show in Figure 5.13 a prediction obtained by combining numerical simulations and a semi-analytic model, as follows. We use the average m/m_{acc} (that is, the ratio between present mass and mass at the time of accretion) as a function of projected distance measured by van den Bosch et al. (2016) for subhaloes in a set of N-body simulations. We combine these predictions with $m_{\text{acc}}(m_{\star})$ estimated by Wang et al. (2013)⁹ by fitting predictions from semi-analytical models to the stellar mass function and the clustering of SDSS galaxies, adopting the median stellar masses for each cluster-centric bin, as quoted in Table 5.1. These predictions are in good agreement with our measurements, and show that we do not expect to see a dependence of the subhalo-to-stellar mass ratio with cluster-centric distance with the current uncertainty levels.

Note that the normalization of the predicted subhalo-to-stellar mass ratio is fixed by the $m_{\star}(m_{\text{acc}})$ relation, and has not been adjusted to match our results, except for the use of the stellar masses of Table 5.1 as input to the $m_{\text{acc}}(m_{\star})$ relation. The fact that the predicted normalization of m/m_{\star} is consistent with our data lends credence to our definition of subhalo mass as m_{bg} or, at the very least, supports the idea that the comparison with theoretical predictions is internally consistent. In fact, van den Bosch et al. (2016) based their analysis on the ROCKSTAR phase-space halo finder (Behroozi et al. 2013) which has been shown to accurately recover subhalo masses

⁹Similar to the treatment of the $m_{\star}(m)$ relation of Rodríguez-Puebla et al. (2013), we Monte-Carlo sample the $m_{\star}(m_{\text{acc}})$ relation of Wang et al. (2013), convolved with the subhalo mass function and accounting for intrinsic scatter, in order to recover $m_{\text{acc}}(m_{\star})$.

at all halo-centric distances; most other subhalo finders (including SUBFIND) tend to underestimate subhalo masses closer to the halo centre (Knebe et al. 2011).

5.7.2 Previous measurements of subhalo mass segregation

Several previous observational studies have focused on the mass segregation of subhaloes. However, differences in the adopted density profiles, mass definitions, and the fact that some works did not report the masses of the host clusters (nor normalized cluster-centric distance by host cluster size), preclude a detailed comparison with our results. To contextualize our results, we nevertheless compare these studies to the present one in a qualitative sense.

Okabe et al. (2014) measured the lensing signal of galaxy- and subgroup-scale subhaloes in the Coma cluster. They found that, while subgroup-scale subhaloes (which they analyzed individually) are better fit by truncated profiles, a stack of individual luminous galaxies is well-fit by a simple NFW model like the one adopted in this work, with no discernible truncation radius. This suggests that, maybe, the stacking of subhaloes with varying truncation radii, produces an average signal in which a truncation radius is no longer discernible. However, this contrasts with the results of Natarajan et al. (1998, 2002, 2007, 2009) and Limousin et al. (2007), who found evidence for galaxy truncation by interpreting the weak lensing signal of cluster galaxies using a maximum likelihood approach. Moreover, these studies found significant evidence for smaller truncation radii (or, equivalently, more compact cores) in galaxies closer to the cluster centres. It is unclear whether the methodology itself allowed the latter set of authors to detect a truncation radius while our methodology is more limiting in this respect, or if the parametrization of the subhalo mass density profile has any influence on this discrepancy, as argued by Pastor Mira et al. (2011). Since the papers above do not show the signal from which their results are derived, it is difficult to assess the origin of the different conclusions we reach compared to theirs. Our detailed assessment of shape measurements in Section 5.4 makes it unlikely that truncation radii of order 10–20 kpc can be detected directly with weak lensing measurements using ground-based observations (as suggested by Limousin et al. 2007), unless perhaps if lens galaxies were subtracted from the images before the analysis, something we will explore in future work. On the other hand, by incorporating the spatial distribution of galaxies into the analysis, one may potentially be able to extract more information than our methodology allows. As it stands, this difference remains unresolved.

More recently, several authors have measured the stacked weak lensing signal as a function of cluster-centric distance, producing results that are more directly comparable to ours. As mentioned in the preceding section, Sifón et al. (2015a) also found no significant segregation of subhalo mass in GAMA groups using weak lensing measurements from KiDS. The results of Niemiec et al. (2017) also suggest no evidence for mass segregation¹⁰. In contrast, Li et al. (2016) found a factor 10 increase in subhalo-to-stellar mass ratio going from $R_{\text{sat}} \sim 0.3r_{200,\text{h}}$ to $R_{\text{sat}} \sim r_{200,\text{h}}$.¹¹ As shown in Figure 5.13, Li et al. (2016) probe cluster-centric distances larger than we do, and it is possible that the effect would be more apparent at larger distance. We caution, however,

¹⁰Niemiec et al. (2017) interpret their measurements as evidence of tidal stripping, given that galaxies closer to the centre have smaller total-to-stellar mass ratios than galaxies further out. However, their results show that there is only a 1σ difference between galaxies closer in and further out, and we instead choose to interpret them as showing no evidence for mass segregation.

¹¹Li et al. (2016) do not report the masses of the host clusters; we adopt a mass $M_{200,\text{h}} = 10^{14} M_{\odot}$ to estimate $r_{200,\text{h}}$.

that Li et al. (2016) used the *photometric* redMaPPer cluster catalogue (Rykoff et al. 2014) to construct their lens sample. Sifón et al. (2015b) showed that even in the case of unbiased photometric redshifts, samples of galaxies selected to be at the cluster redshift are significantly contaminated at large distances. It is therefore possible that the trend observed by Li et al. (2016) may be due at least in part to contamination by unrelated galaxies although the use by Li et al. (2016) of only high-probability cluster members (based on the redMaPPer definition) may somewhat mitigate this (see Zu et al. 2016).

5.8 Conclusions

We present the average masses of satellite galaxies in massive galaxy clusters at $0.05 < z < 0.15$ using weak galaxy-galaxy lensing measurements. We use a combination of deep, wide-field observations of galaxy clusters and extensive archival spectroscopic data (Sifón et al. 2015b). Using extensive image simulations of bright lenses in the foreground of a population of field galaxies resembling the source population in our data, we model and account for biases arising from (i) shape measurements, due to confusion of light from the lens with the faint sources, and (ii) contamination of the source sample by faint cluster members (Section 5.4).

We model the lensing signal from subhaloes using an NFW profile and the subhalo mass-concentration relation measured from N-body simulations by Moliné et al. (2017), which depends on cluster-centric distance. We split the sample in bins of stellar mass and measure the subhalo-to-stellar mass relation (SHSMR) of galaxies in massive clusters. Fitting the resulting masses with a power-law relation, we find $\log m_{\text{bg}} = (11.66 \pm 0.07) + (0.94 \pm 0.15) \log m_{\star}$ (Figure 5.10). The slope of this relation is robust to both the adopted subhalo mass-concentration relation and the subhalo mass definition. We find that at a characteristic stellar mass of $\sim 3 \times 10^{10} M_{\odot}$, the ratio between subhalo mass and host halo mass is maximal, reaching a value of approximately 0.5, and dropping to 0.2 at the high-mass end (Figure 5.11). This behaviour is likely caused by a combination of tidal stripping and dynamical friction.

We also study the masses of subhaloes at different cluster-centric distances with the aim of studying the evolution of subhaloes within clusters. We find no statistically significant evidence for mass segregation and a mean total-to-stellar mass ratio $\langle m_{\text{bg}}/m_{\star} \rangle = 21.5_{-5.5}^{+6.3}$ (Figure 5.13). Our results are consistent with predictions from a combination of numerical simulations and semi-analytic models and with some, but not all, previous results. Although direct comparison with the observational literature is complicated by the use of different definitions and conventions, our results are generally consistent within the overlapping stellar mass and cluster-centric distance ranges.

The halo model commonly employed in galaxy-galaxy lensing studies requires some assumptions about the density profiles of the subhaloes hosting satellite galaxies. At low stellar masses ($M_{\star} \lesssim 10^{10} M_{\odot}$), most red galaxies seem to be satellites (Mandelbaum et al. 2006; Velander et al. 2014). Therefore, host halo masses at low stellar masses are determined through the halo model based on observations of what are mostly satellite galaxies; degeneracies in the halo model dominate the resulting masses (Velandar et al. 2014). While we are not able to constrain the density profiles of subhaloes at present, the subhalo-to-stellar mass relation is an important ingredient that could be incorporated in future galaxy-galaxy lensing analyses to inform these choices.

Acknowledgments

We are indebted to Marcello Cacciato for valuable discussions and suggestions throughout the development of this work. We thank David Sand and Melissa Graham for their participation in the MENeCS proposal and observations. CS, RH, HH and MV acknowledge support from the European Research Council under FP7 grant number 279396. RFJvdB acknowledges support from the European Research Council under FP7 grant number 340519.

Based on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/DAPNIA, at the Canada-France-Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences de l'Univers of the Centre National de la Recherche Scientifique of France, and the University of Hawaii.

Bibliography

- Akritas M. G., Bershady M. A., 1996, *ApJ*, 470, 706
- Applegate D. E., et al., 2014, *MNRAS*, 439, 48
- Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013, *ApJ*, 762, 109
- Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, *MNRAS*, 370, 645
- Boylan-Kolchin M., Bullock J. S., Kaplinghat M., 2011, *MNRAS*, 415, L40
- Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000
- Bullock J. S., Kolatt T. S., Sigad Y., Somerville R. S., Kravtsov A. V., Klypin A. A., Primack J. R., Dekel A., 2001, *MNRAS*, 321, 559
- Cacciato M., van den Bosch F. C., More S., Li R., Mo H. J., Yang X., 2009, *MNRAS*, 394, 929
- Chabrier G., 2003, *PASP*, 115, 763
- Chang J., Macciò A. V., Kang X., 2013, *MNRAS*, 431, 3533
- Clowe D., Luppino G. A., Kaiser N., Henry J. P., Gioia I. M., 1998, *ApJL*, 497, L61
- Coupon J., et al., 2015, *MNRAS*, 449, 1352
- Crain R. A., et al., 2015, *MNRAS*, 450, 1937
- Duffy A. R., Schaye J., Kay S. T., Dalla Vecchia C., 2008, *MNRAS*, 390, L64
- Dutton A. A., Macciò A. V., 2014, *MNRAS*, 441, 3359
- Fahlman G., Kaiser N., Squires G., Woods D., 1994, *ApJ*, 437, 56
- Fang Y., Clampitt J., Dalal N., Jain B., Rozo E., Moustakas J., Rykoff E., 2016, *MNRAS*, 463, 1907
- Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, 125, 306

- Gabor J. M., Davé R., 2015, *MNRAS*, 447, 374
- Gao L., Springel V., White S. D. M., 2005, *MNRAS*, 363, L66
- Geiger B., Schneider P., 1998, *MNRAS*, 295, 497
- Ghigna S., Moore B., Governato F., Lake G., Quinn T., Stadel J., 1998, *MNRAS*, 300, 146
- Goodman J., Weare J., 2010, *Communications in Applied Mathematics and Computational Science*, 5, 65
- Gruen D., Brimiouille F., 2016, preprint, ([arXiv:1610.01160](https://arxiv.org/abs/1610.01160))
- Gwyn S. D. J., 2008, *PASP*, 120, 212
- Haines C. P., et al., 2015, *ApJ*, 806, 101
- Halkola A., Seitz S., Pannella M., 2007, *ApJ*, 656, 739
- Han J., Cole S., Frenk C. S., Jing Y., 2016, *MNRAS*, 457, 1208
- Hao J., et al., 2009, *ApJ*, 702, 745
- Hayashi E., Navarro J. F., Taylor J. E., Stadel J., Quinn T., 2003, *ApJ*, 584, 541
- Hearin A. P., Zentner A. R., van den Bosch F. C., Campbell D., Tollerud E., 2016, *MNRAS*, 460, 2552
- Heymans C., et al., 2006a, *MNRAS*, 368, 1323
- Heymans C., et al., 2006b, *MNRAS*, 371, L60
- Hoekstra H., 2007, *MNRAS*, 379, 317
- Hoekstra H., Franx M., Kuijken K., Squires G., 1998, *ApJ*, 504, 636
- Hoekstra H., Franx M., Kuijken K., 2000, *ApJ*, 532, 88
- Hoekstra H., Hsieh B. C., Yee H. K. C., Lin H., Gladders M. D., 2005, *ApJ*, 635, 73
- Hoekstra H., Mahdavi A., Babul A., Bildfell C., 2012, *MNRAS*, 427, 1298
- Hoekstra H., Herbonnet R., Muzzin A., Babul A., Mahdavi A., Viola M., Cacciato M., 2015, *MNRAS*, 449, 685
- Kaiser N., Squires G., Broadhurst T., 1995, *ApJ*, 449, 460
- Klypin A., Gottlöber S., Kravtsov A. V., Khokhlov A. M., 1999, *ApJ*, 516, 530
- Knebe A., et al., 2011, *MNRAS*, 415, 2293
- Kriek M., van Dokkum P. G., Labbé I., Franx M., Illingworth G. D., Marchesini D., Quadri R. F., 2009, *ApJ*, 700, 221
- Lacey C. G., et al., 2015, preprint, ([arXiv:1509.08473](https://arxiv.org/abs/1509.08473))
- Laigle C., et al., 2016, *ApJS*, 224, 24

- Leauthaud A., et al., 2012, *ApJ*, 744, 159
- Li R., Mo H. J., Fan Z., Yang X., Bosch F. C. v. d., 2013a, *MNRAS*, 430, 3359
- Li C., Wang L., Jing Y. P., 2013b, *ApJL*, 762, L7
- Li R., et al., 2014, *MNRAS*, 438, 2864
- Li R., et al., 2016, *MNRAS*, 458, 2573
- Limousin M., Kneib J.-P., Natarajan P., 2005, *MNRAS*, 356, 309
- Limousin M., Kneib J. P., Bardeau S., Natarajan P., Czoske O., Smail I., Ebeling H., Smith G. P., 2007, *A&A*, 461, 881
- Luppino G. A., Kaiser N., 1997, *ApJ*, 475, 20
- Macciò A. V., Dutton A. A., van den Bosch F. C., 2008, *MNRAS*, 391, 1940
- Mancone C. L., Gonzalez A. H., 2012, *PASP*, 124, 606
- Mandelbaum R., et al., 2005a, *MNRAS*, 361, 1287
- Mandelbaum R., Tasitsiomi A., Seljak U., Kravtsov A. V., Wechsler R. H., 2005b, *MNRAS*, 362, 1451
- Mandelbaum R., Seljak U., Kauffmann G., Hirata C. M., Brinkmann J., 2006, *MNRAS*, 368, 715
- Mandelbaum R., Wang W., Zu Y., White S., Henriques B., More S., 2016, *MNRAS*, 457, 3200
- McGee S. L., Balogh M. L., Bower R. G., Font A. S., McCarthy I. G., 2009, *MNRAS*, 400, 937
- Mellier Y., 1999, *ARA&A*, 37, 127
- Moliné Á., Sánchez-Conde M. A., Palomares-Ruiz S., Prada F., 2017, *MNRAS*, 466, 4974
- Moore B., Ghigna S., Governato F., Lake G., Quinn T., Stadel J., Tozzi P., 1999, *ApJL*, 524, L19
- More S., van den Bosch F. C., Cacciato M., Skibba R., Mo H. J., Yang X., 2011, *MNRAS*, 410, 210
- Natarajan P., Kneib J.-P., 1997, *MNRAS*, 287, 833
- Natarajan P., Kneib J.-P., Smail I., Ellis R. S., 1998, *ApJ*, 499, 600
- Natarajan P., Kneib J.-P., Smail I., 2002, *ApJL*, 580, L11
- Natarajan P., De Lucia G., Springel V., 2007, *MNRAS*, 376, 180
- Natarajan P., Kneib J.-P., Smail I., Treu T., Ellis R., Moran S., Limousin M., Czoske O., 2009, *ApJ*, 693, 970
- Navarro J. F., Frenk C. S., White S. D. M., 1995, *MNRAS*, 275, 720

- Niemiec A., et al., 2017, preprint, ([arXiv:1703.03348](#))
- Oguri M., Bayliss M. B., Dahle H., Sharon K., Gladders M. D., Natarajan P., Hennawi J. F., Koester B. P., 2012, *MNRAS*, 420, 3213
- Okabe N., Futamase T., Kajisawa M., Kuroshima R., 2014, *ApJ*, 784, 90
- Old L., et al., 2015, *MNRAS*, 449, 1897
- Pastor Mira E., Hilbert S., Hartlap J., Schneider P., 2011, *A&A*, 531, A169
- Peacock J. A., Smith R. E., 2000, *MNRAS*, 318, 1144
- Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2002, *AJ*, 124, 266
- Planck Collaboration 2015, preprint, ([arXiv:1502.01589](#))
- Prada F., Klypin A. A., Cuesta A. J., Betancort-Rijo J. E., Primack J., 2012, *MNRAS*, 423, 3018
- Rines K. J., Geller M. J., Diaferio A., Hwang H. S., 2016, *ApJ*, 819, 63
- Rix H.-W., et al., 2004, *ApJS*, 152, 163
- Roberts I. D., Parker L. C., Joshi G. D., Evans F. A., 2015, *MNRAS*, 448, L1
- Rodríguez-Puebla A., Drory N., Avila-Reese V., 2012, *ApJ*, 756, 2
- Rodríguez-Puebla A., Avila-Reese V., Drory N., 2013, *ApJ*, 767, 92
- Rowe B. T. P., et al., 2015, *Astronomy and Computing*, 10, 121
- Rykoff E. S., et al., 2014, *ApJ*, 785, 104
- Sand D. J., et al., 2012, *ApJ*, 746, 163
- Schaye J., et al., 2015, *MNRAS*, 446, 521
- Schechter P., 1976, *ApJ*, 203, 297
- Schlafly E. F., Finkbeiner D. P., 2011, *ApJ*, 737, 103
- Schlegel D. J., Finkbeiner D. P., Davis M., 1998, *ApJ*, 500, 525
- Schneider P., 2003, *A&A*, 408, 829
- Seljak U., 2000, *MNRAS*, 318, 203
- Sérsic J. L., 1968, *Atlas de galaxias australes*
- Sifón C., et al., 2015a, *MNRAS*, 454, 3938
- Sifón C., Hoekstra H., Cacciato M., Viola M., Köhlinger F., van der Burg R. F. J., Sand D. J., Graham M. L., 2015b, *A&A*, 575, A48
- Sifón C., van der Burg R. F. J., Hoekstra H., Muzzin A., Herbonnet R., 2017, preprint, ([arXiv:1704.07847](#))
- Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, *MNRAS*, 328, 726

- Springel V., et al., 2005, *Nature*, 435, 629
- Springel V., et al., 2008, *MNRAS*, 391, 1685
- Suyu S. H., Halkola A., 2010, *A&A*, 524, A94
- Taffoni G., Mayer L., Colpi M., Governato F., 2003, *MNRAS*, 341, 434
- Taylor J. E., Babul A., 2005, *MNRAS*, 364, 535
- Tormen G., Diaferio A., Syer D., 1998, *MNRAS*, 299, 728
- Umetsu K., et al., 2014, *ApJ*, 795, 163
- Umetsu K., Zitrin A., Gruen D., Merten J., Donahue M., Postman M., 2016, *ApJ*, 821, 116
- Velander M., et al., 2014, *MNRAS*, 437, 2111
- Velliscig M., et al., 2016, preprint, ([arXiv:1612.04825](https://arxiv.org/abs/1612.04825))
- Viola M., et al., 2015, *MNRAS*, 452, 3529
- Wang L., De Lucia G., Weinmann S. M., 2013, *MNRAS*, 431, 600
- Yang X., Mo H. J., van den Bosch F. C., Jing Y. P., Weinmann S. M., Meneghetti M., 2006, *MNRAS*, 373, 1159
- Yang X., Mo H. J., van den Bosch F. C., 2009, *ApJ*, 693, 830
- Zolotov A., et al., 2012, *ApJ*, 761, 71
- Zu Y., Mandelbaum R., 2015, *MNRAS*, 454, 1161
- Zu Y., Mandelbaum R., Simet M., Rozo E., Rykoff E. S., 2016, preprint, ([arXiv:1611.00366](https://arxiv.org/abs/1611.00366))
- van Uitert E., Hoekstra H., Velander M., Gilbank D. G., Gladders M. D., Yee H. K. C., 2011, *A&A*, 534, A14
- van Uitert E., et al., 2016, *MNRAS*,
- van den Bosch F. C., 2017, *MNRAS*, 468, 885
- van den Bosch F. C., More S., Cacciato M., Mo H., Yang X., 2013, *MNRAS*, 430, 725
- van den Bosch F. C., Jiang F., Campbell D., Behroozi P., 2016, *MNRAS*, 455, 158
- van der Burg R. F. J., et al., 2013, *A&A*, 557, A15
- van der Burg R. F. J., Hoekstra H., Muzzin A., Sifón C., Balogh M. L., McGee S. L., 2015, *A&A*, 577, A19

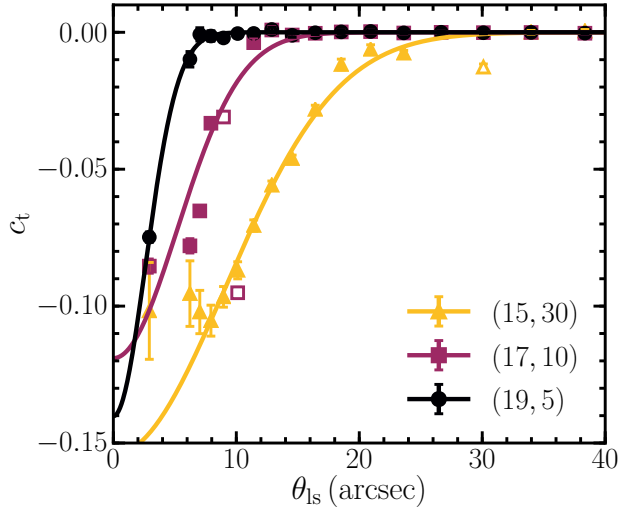


Figure 5.14: Additive tangential shear biases measured in three sets of image simulations as shown in the legend, which shows the magnitude and size (in pixels = 0.185 arcseconds) of each set. The three examples correspond to big bright (yellow triangles), average (purple squares), and small faint (black circles) simulated lenses, and illustrate the range of biases. The relevance of each set with respect to the real satellite galaxies can be seen in Figure 5.2: both extremes are very rare, while the purple set corresponds to the most common (mag,size) configuration. Data points with errorbars show measured tangential shear and solid lines show Gaussian fits to each set of simulations. Empty points are biased because they are adjacent the chosen truncation radius of the lenses, and are excluded from the fits.

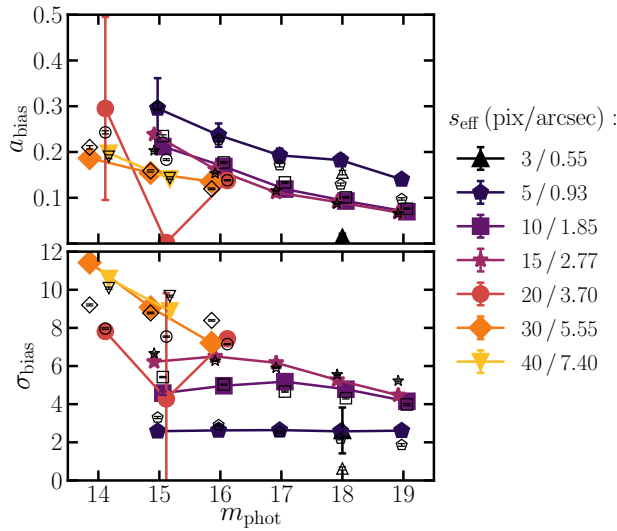


Figure 5.15: Amplitude and width of Gaussian fits to the additive bias c_t (solid symbols), and the results from an overall fit to each panel given in Equation 5.24 (empty symbols). Solid lines connect simulation sets with the same half-light radius as shown in the legend.

5.A Lens-induced bias on the shape measurements

Extended light from bright lens galaxies affects measurements of sources, such that their shapes are estimated to be more radially elongated than they really are. This induces a negative additive bias *in the coordinate frame of the lens galaxy*, which we label c_t .

In order to account for this bias we measure the shapes of galaxies in the image simulations of Hoekstra et al. (2015), after adding bright lens galaxies in a grid pattern (separated by 1 arcmin from each other). These injected lenses are modelled as a circular Sérsic (1968) profile (i.e., $I(r) \propto r^{1/n}$) using GALSIM, with a power-law index $n = 4$. A Sérsic profile with a high index has very extended wings and to avoid the surface brightness profiles of different lenses to overlap we truncate the lens profiles at $5 s_{\text{eff}}$, where s_{eff} is the effective, or half-light, radius of the Sérsic profile. The source galaxies in the image simulations have a constant shear applied to them, which cancels out when we average over an isotropic grid of shears. Therefore any measured shear in the tangential frame can be attributed to a bias induced by extended light from the lenses. The lenses we inject into the simulations span the ranges $14 \leq m_{\text{phot}} \leq 20$ and $3 \leq s_{\text{eff}}/\text{pix} \leq 40$ (corresponding to $0''.55 \leq s_{\text{eff}} \leq 7''.40$), and are compared to the magnitude and size distribution (as measured by GALFIT, see Sifón et al. 2015b) in the MENeCS data in Figure 5.2.

We show the measured c_t for three sample sets of simulations in Figure 5.14. We find that the bias profiles can be well modelled in each bin as a Gaussian centred at $\theta_{\text{ls}} = 0$,

$$c_t(\theta_{\text{ls}}) = a_{\text{bias}} \exp\left[\frac{-\theta^2}{2\sigma_{\text{bias}}^2}\right]. \quad (5.23)$$

We then fit the best-fit parameters a_{bias} and σ_{bias} as functions of lens magnitude and size,

$$\begin{aligned} a_{\text{bias}} &= -0.81 - 1.22(m_{\text{phot}} - 16) - 0.36 \log(s_{\text{eff}}/15 \text{ pix}), \\ \sigma_{\text{bias}} &= 6.27 - 14.01 \log(m_{\text{phot}}/16) + 7.04 \log(s_{\text{eff}}/15 \text{ pix}). \end{aligned} \quad (5.24)$$

Figure 5.15 shows the best-fit individual values of a_{bias} and σ_{bias} and the values predicted by Equation 5.24. While at face value Equation 5.24 is not a good description of the measurements in the simulations for the full $(m_{\text{phot}}, s_{\text{eff}})$ space (and especially for σ_{bias}), the discrepancy is limited to the extremes of this space. One notable discrepancy is roughly a 25 per cent difference in the prediction of σ_{bias} for $(m_{\text{phot}}, s_{\text{eff}}) = (14, 30)$ (here, sizes are given in pixels). However, as shown in Figure 5.2, this combination of magnitude and size accounts for much less than 1 per cent of the lenses in our sample. The other notable difference happens at $(m_{\text{phot}}, s_{\text{eff}}) = (18, 3)$, but the bias introduced by such small, faint galaxies is negligible to start with. Moreover, as can be seen in Figure 5.15, the difference arises because of the degeneracy between the amplitude and width of the Gaussian, such that the predicted bias is negligible as well.

Nederlandse Samenvatting

6.1 De duistere kant van het Universum

De nachtelijke hemel wordt verlicht door duizenden zichtbare sterren en ontelbaar meer die niet helder genoeg zijn om met het oog waar te nemen. Al deze sterren vormen echter maar een klein deel van alle materie in het Universum. Enorme gaswolken bevatten het merendeel van deeltjes die in het hele licht spectrum zijn te observeren. Maar verreweg het meest voorkomend materiaal in het Universum is in de vorm van de mysterieuze donkere materie. Dit fenomeen is verantwoordelijk voor de grote hoeveelheid zwaartekracht die nodig is om astronomische objecten, zoals sterrenstelsels en clusters van sterrenstelsels, bijeen te houden. Hoewel de zwaartekracht van donkere materie reeds in de jaren dertig was opgemerkt, is donkere materie zelf (nog) niet waargenomen omdat het geen (bekende) interactie heeft met licht. Toch is ook donkere materie niet het meest dominante ingrediënt van het Universum: ongeveer 75% van het energie budget zit in donkere energie. Donkere energie is een beschrijving voor de waargenomen versnelde uitdijning van het Universum, terwijl een Universum gevuld met massa onder de invloed van zwaartekracht steeds langzamer zou moeten expanderen. De benaming van deze fysische verschijnselen is gekozen omdat men nog in het duister tast over de theoretische verklaringen ervoor.

Er zijn veel theoretische modellen ontwikkeld om het bestaan van donkere materie en donkere energie te verklaren, maar tot nu toe heeft geen enkel model een bevredigend antwoord gegeven voor alle kosmologische waarnemingen. Het zoeken van een fysische oorzaak wordt bemoeilijkt door het feit dat de donkere componenten van het Universum niet direct waargenomen kunnen worden. Er zijn wel experimenten die proberen deeltjes, die geproduceerd worden door donkere materie, te detecteren, maar daar is nog geen positief resultaat uit gekomen. In plaats daarvan worden de gevolgen van donkere materie en donkere energie bestudeerd en vergeleken met voorspellingen van theoretische modellen om de fysische oorsprong te achterhalen. Veel kosmologische onderzoeken richten zich op de totale hoeveelheid van donkere materie en energie in het Universum. Aan de hand daarvan kunnen fysische eigenschappen worden bepaald, zoals een mogelijke evolutie van donkere energie met de tijd.

Een van de mogelijkheden voor kosmologisch onderzoek is het in kaart brengen van de structuur in het Universum. Materie op kosmische schalen is niet willekeurig verdeeld, maar het klontert samen onder de invloed van zwaartekracht. Waarnemingen laten zien dat materie verdeeld is in een soort kosmisch web, waarbij sterrenstelsels zich bevinden in lange draden. Waar de draden elkaar kruisen ontstaan de grootste verzamelingen van massa in het Universum: clusters van sterrenstelsels. Simulaties laten ook zien dat de donkere materie een soortgelijke verdeling heeft en dat de zwaartekracht van de donkere materie de belangrijkste factor is voor het ontstaan van deze structuur. Daarentegen wordt de samenklontering tegengewerkt door de uitdijning van

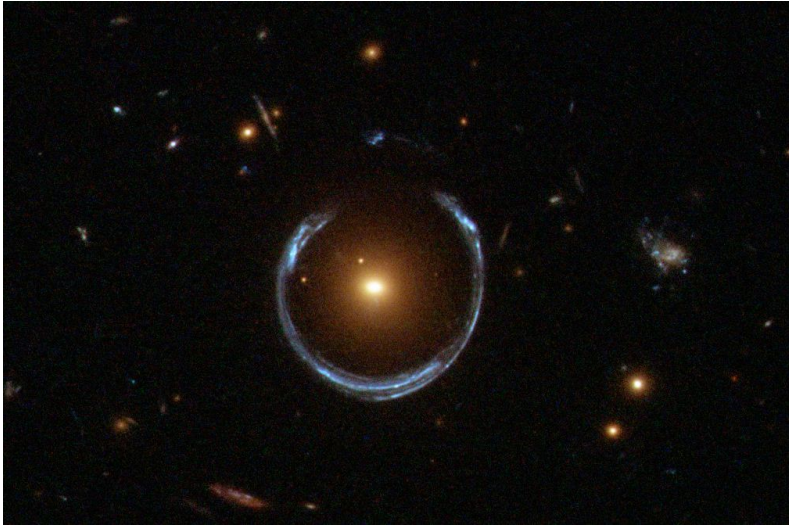
het Universum. Uit simulaties van hypothetische donkere materie deeltjes blijkt dat de aanwezigheid van donkere energie er voor zorgt dat er minder van de meest massieve objecten (ook wel *halo's* van donkere materie genoemd) zijn en dat de gemiddelde massa van halo's lager is. Door de massa's en verdeling van halo's in het Universum te meten en te vergelijken met simulaties kan de hoeveelheid donkere materie en donkere energie worden bepaald.

Het bepalen van de verdeling van massa in het Universum vereist een methode om donkere materie halo's in kaart te brengen. De grootste verzamelingen van lichtgevende deeltjes bevinden zich in massieve wolken van heet gas en in sterrenstelsels en deze objecten worden door de zwaartekracht van de donkere materie halo's aangetrokken. De locaties van de heldere objecten, die direct met telescopen waar te nemen zijn, onthullen dus de verdeling van donkere materie. Echter is de relatie tussen de hoeveelheid donkere- en lichte materie niet volledig bekend. Dit komt omdat de verscheidenheid aan astrofysische processen, die ondergaan worden door sterrenstelsels en gaswolken, nog niet volledig begrepen is. Er is dus een directe manier nodig om de massa van halo's van donkere materie te meten en daarmee empirisch de relatie tot heldere objecten te bepalen of de gehele verdeling van donkere materie in kaart te brengen. Het enige wat bekend is van donkere materie is dat het zwaartekracht uitoefent en het bestuderen van donkere materie kan dus alleen door middel van de zwaartekracht.

6.2 Zwaartekrachtslenzen

De algemene relativiteitstheorie van Albert Einstein beschrijft zwaartekracht als de invloed van een object met massa op de ruimte-tijd eromheen. De omliggende ruimte-tijd wordt gekromd en door deze kromming wordt het pad van het licht, uitgezonden door een helder hemellichaam achter het massieve object, verbogen. Het licht van het achterliggende hemellichaam zal anders worden waargenomen, dan dat het zou worden zonder het massieve object in de voorgrond. De zwaartekracht werkt dus als een optische lens en het massieve object wordt dan een *zwaartekrachtslens* genoemd. Voor kosmologische onderzoeken zijn de zwaartekrachtslenzen voornamelijk grote wolken van donkere materie en zijn de heldere achtergrond objecten verafgelegen sterrenstelsels. Sterrenstelsels zijn grote objecten en licht uit verschillende delen van het stelsel wordt anders beïnvloed door een zwaartekrachtslens, waardoor het waargenomen beeld van sterrenstelsels wordt vervormd. Het Hoefijzer sterrenstelsel in Figuur 6.1 is een van de bekendste voorbeelden van lenswerking door zwaartekracht en laat de vervorming duidelijk zien. De blauwe band is een enkel helder sterrenstelsel direct achter het centrale gele sterrenstelsel dat door de bijbehorende donkere materie halo is vervormd. Deze vervorming is een herkenbaar teken van een zwaartekrachtslens en kan gebruikt worden om de massa van de lens te bepalen. Dit geval van *sterke lenswerking* gebeurt alleen bij een sterrenstelsel direct achter een zwaartekracht lens. Zulke configuraties zijn zeldzaam.

In het meest voorkomende geval van *zwakke lenswerking door zwaartekracht* ligt een sterrenstelsel niet direct achter de zwaartekracht lens en is het effect een minieme uitrekking (*shear*) van het waargenomen licht profiel van het sterrenstelsel. In tegenstelling tot de duidelijke vervorming van het Hoefijzer stelsel, is deze shear veel kleiner dan de intrinsieke elliptische vorm van het sterrenstelsel. Het is dus onmogelijk om het effect van de zwaartekracht lens te bepalen voor een enkel stelsel. In plaats daarvan kan de shear met de vormen van een grote hoeveelheid sterrenstelsels statistisch



Figuur 6.1: Waarneming met de Hubble Space Telescope van het Hoefijzer sterrenstelsel. Het blauwe licht wordt uitgezonden door een enkel sterrenstelsel wat door lenswerking door zwaartekracht wordt vervormd. Hierdoor nemen wij het licht waar in de vorm van een hoefijzer. Het Hoefijzer sterrenstelsel is een bekend voorbeeld van sterke lenswerking door zwaartekracht en de massa van de zwaartekrachtslens kan bepaald worden aan de hand het hoefijzervormige licht. In dit geval is het gelige sterrenstelsel in het midden van het hoefijzer bij lange na niet zwaar genoeg om het hoefijzer te veroorzaken en dus is er een halo van donkere materie rond dat sterrenstelsel.

worden gemeten, onder de aanname dat sterrenstelsels gemiddeld cirkelvormig zijn. Bij een intrinsiek cirkelvormig sterrenstelsel is de enige uitrekking afkomstig van de zwaartekrachtslens en dus kan de shear geïsoleerd worden voor een verzameling van sterrenstelsels. De aanname vereist een zo groot mogelijke hoeveelheid waargenomen sterrenstelsels om de statistische fout door de intrinsieke vormen te verminderen. Experimenten met zwakke lenswerking (door zwaartekracht) gebruiken dus een zo groot mogelijk oppervlak aan de hemel.

Een belangrijk onderdeel van onderzoek naar zwaartekrachtslenzen is het precies meten van de vormen van sterrenstelsels. Aangezien sterrenstelsels ruwweg kunnen worden beschreven als ellipsen, kan uit de ratio van de lengte van de assen de ellipticiteit worden bepaald, die wiskundig te relateren is aan de shear. Het bepalen van de ellipticiteit van een sterrenstelsel wordt bemoeilijkt door de andere processen die licht ondergaat als het zich een weg baant door de atmosfeer en de telescoop apparatuur. Dit zorgt voor versmering van het lichtprofiel van het sterrenstelsel met een puntspreidingsfunctie (PSF) en daarnaast is er altijd ruis in de observaties. Beide effecten kunnen de ellipticiteit van het waargenomen lichtprofiel veranderen en dus een meetfout in de shear veroorzaken.

Hoewel veel verschillende technieken zijn ontwikkeld om het effect van de PSF ongedaan te maken, is het effect van ruis pas in de laatste jaren meer bestudeerd. Door de verbeterde technologie, worden de experimenten met zwakke lenswerking steeds beter en moeten meetfouten ook beter onder controle worden gehouden. De enige (beproeft) manier om meetfouten van methoden die vormen meten te bepalen, is door de lichtprofielen van sterrenstelsels te simuleren na toevoeging van alle oorzaken

van vervormingen. Doordat de toegevoegde shear bekend is, kan de meetfout precies worden gekwantificeerd voor specifieke processen die het lichtprofiel beïnvloeden. Deze simulaties van telescoop opnames vormen de basis voor elk experiment met zwakke lenswerking door zwaartekracht.

6.3 Dit proefschrift

Dit proefschrift begint met het verbeteren van methodes om shear te meten. In het tweede deel van het proefschrift verleggen we de aandacht naar metingen van zwakke lenswerking door zwaartekracht van de grootste structuren in het heelal, met als uiteindelijk doel limieten te stellen aan de mogelijke hoeveelheid donkere energie en materie in het Universum.

In **Hoofdstuk 2** ontwikkelen wij een nieuw algoritme om de vormen van sterrenstelsels te meten. De nieuwe methode is speciaal ontworpen voor zwakke stelsels waarvoor het meten van vormen moeilijk is. In tegenstelling tot andere vergelijkbare methoden, is de deconvolutie van de PSF een analytisch proces, waardoor in principe ook kleine stelsels kunnen worden bestudeerd. De meetfout door de aanwezigheid van ruis is klein door het gebruik van een van tevoren vastgestelde gewichtsfunctie en het gebruik van momenten in plaats van ellipticiteit om de shear te bepalen. Deze meetfout kan verder worden verminderd door een extra analytisch proces en wij vinden dat ruis een meetfout van minder dan een procent introduceert. Dit resultaat is goed genoeg voor de huidige experimenten, al moet de methode nog worden getest voor andere oorzaken van meetfouten.

Een van de grootste experimenten voor zwakke lenswerking op dit moment is de Kilo Degree Survey (KiDS). Dit experiment zal uiteindelijk 1500 vierkante graden van de hemel bestuderen met behulp van zwakke lenswerking door zwaartekracht en de verdeling van donkere materie in kaart brengen, om daarmee limieten te zetten op kosmologische parameters. Het algoritme dat het KiDS team gebruikt om vormen te meten, moest gekalibreerd worden om de statistiek van de grote dataset volledig uit te buiten. In **Hoofdstuk 3** beschrijf ik de tests die wij hebben gedaan om een betrouwbare kalibratie te verzekeren. Met enorme simulaties van de KiDS data, waarbij de sterrenstelsels gesimuleerd zijn aan de hand van echte data, hebben wij de meetfout gekwantificeerd. Onrealistische simulaties kunnen een verkeerde meetfout opleveren en dus heb ik veel aandacht besteed om de eigenschappen van de gesimuleerde sterrenstelsels zoveel mogelijk op de echte data te laten lijken. Voor de overgebleven verschillen heb ik geverifieerd dat de kalibratie onveranderd bleef als de simulaties wel op de data lijken. De kalibratie die wij hebben verzorgd, is nauwkeurig tot op 1% en is gebruikt in, onder andere, de kosmologische analyse.

Een vergelijkbaar onderzoek als dat beschreven in **Hoofdstuk 3** heb ik tijdens mijn promotie traject verricht voor een andere studie, die buiten dit proefschrift beschreven is. Met dat onderzoek hebben wij een praktische methode om vormen te meten, zeer nauwkeurig gecorrigeerd voor meetfouten. Deze methode is toegepast op twee grote collecties van waarnemingen van clusters van sterrenstelsels om de massa's ervan te bepalen. Het werk voor een van deze collecties, die de zwaarste clusters bevat met een roodverschuiving lager dan 0.15, is beschreven in **Hoofdstuk 4**. De afstanden tussen de Aarde, de waargenomen clusters, en de vervormde achtergrond stelsels zijn bepaald met behulp van diepe fotometrische data. Met nog meer simulaties van telescoop opnames is bepaald hoe clusters de populatie van achtergrond stelsels beïnvloeden en dus het gemeten shear signaal veranderen. Na correctie voor dit laatste fenomeen,

heb ik de massa's gemeten van de clusters en deze vergeleken met andere schattingen van de massa. De massa gemeten aan de hand van het hete gas is een van de meest gebruikte indicatoren massa, omdat dit van veel clusters bekend is. Met onze directe metingen van de verdeling van donkere materie tonen we aan dat de relatie tussen de twee massa indicatoren afhangt van de massa van het cluster. Met deze relatie kunnen bestaande catalogi van massa metingen aan de hand van het gas beter gebruikt worden voor kosmologische analyses.

In **Hoofdstuk 5** gebruiken we dezelfde methode als in **Hoofdstuk 4** op dezelfde clusters en zijn we geïnteresseerd in de structuur binnen de clusters zelf. Hiertoe meten wij de massa's rond satelliet sterrenstelsels in het cluster. Satelliet stelsels zijn sterrenstelsels, die zich niet in het centrum van het cluster bevinden, omdat ze geïsoleerde sterrenstelsels waren die het cluster zijn ingevallen. Deze accumulatie van satelliet stelsels is een van de belangrijkste processen waarmee het cluster meer massa vergaart en toch is er nog maar weinig bekend over de donkere materie halo's rond deze satelliet stelsels (*subhalo's*). Om de massa van subhalo's te bepalen, moet de shear van sterrenstelsels, die zich in projectie dicht bij satelliet stelsels bevinden, gemeten worden. Het licht van de satelliet stelsels zorgt voor een meetfout en ik heb daarom nauwkeurig het effect van de nabijheid van deze objecten op de vorm metingen gekwantificeerd. We hebben de relatie tussen de massa's van subhalo's en van massa's in de sterren in satelliet sterrenstelsels bepaald en vinden overeenstemming met eerder werk. In tegenstelling tot ander werk vinden wij niet verschillende subhalo massa's als functie van nabijheid tot het centrum van het cluster. De onzekerheid in de metingen is vooralsnog te groot om een duidelijke conclusie te trekken over hoeveel massa van de subhalo verloren gaat als een satelliet stelsels het cluster invalt. Meer theoretisch en observationeel onderzoek in dit veld is nodig om de groei van clusters te begrijpen.

Curriculum Vitae

I saw first light on the 3rd of May 1988 in Rotterdam in the Netherlands. I attended the Stedelijk Gymnasium in Breda as my high school, specialising in the natural sciences and following extra classes in classical Greek, which enabled me to pronounce the weird symbols used in physics. The choice to study astronomy came in the last few months of the final year, when I joined a friend for an open day at the University of Leiden. Not sure yet what I wanted do for the rest of my life, I decided to combine the two studies of astronomy and physics.

In those days I was one of the few students of both physics and astronomy and for our bachelor's project we were relegated to choose something with spectroscopy, which was deemed to be sufficient for both fields. I spent several months under the supervision of dr. Peter Gast working with extremely powerful magnets to investigate hydrogen bonds of nitroxide spin-labeled protein sites in organic molecules. I obtained a bachelor's degree in both physics and astronomy in 2009, after which I spent one year as treasurer in the board of directors for student society Catena.

I pursued a master's programme in astronomy and for my first project I investigated the performance of shape measurement methods under the supervision of prof. Konrad Kuijken. This project was continued in the next year by (now dr.) Axel Buddendiek and together our projects have laid the ground work for the first chapter of this thesis. In my second year as a master student I joined the group of prof. Huub Röttgering and studied protoclusters: galaxy clusters before they can be called galaxy clusters. In a simulation of galaxies, the Millennium Simulation, I looked at the detectability of protoclusters at redshifts 2-3 with submillimeter wavelengths, specifically those of the SPIRE instrument on the Herschel space satellite. I found that some, but not all protoclusters could be identified with the submillimeter measurements. Although I did not pursue this project, collaborators published their work on protoclusters after I had finished the master's project. And, eventually I did return to galaxy clusters for the second half of this thesis.

In 2012 I started as a PhD student under the supervision of, then dr. and now, prof. Henk Hoekstra and prof. Konrad Kuijken. The emphasis of the PhD would be the improvement of shape measurement methods, so that weak gravitational lensing experiments are able to produce reliable results. Having some experience with calibration simulations I was drafted as an active member of the Kilo Degree Survey (KiDS) collaboration, whose goal it is to set constraints on cosmological parameters using cosmic shear. In a dedicated team we derived the calibration for the KiDS shape measurement algorithm using the simulated telescope images I created. Unsatisfied with just fake data, I used the observations of galaxy clusters to measure the dark matter structures with weak gravitational lensing. For this I also spent several nights observing galaxy clusters at the Isaac Newton Telescope on La Palma. This was my second time at that telescope, with which I had already observed in the second year of my bachelor's programme. I have participated in a summer school in Cargese and I have attended conferences at several (inter)national venues: Leiden, Madrid, Rome, Marseille, Philadelphia. I have been invited to Bonn twice, but for different projects, as well as to Malta. At some of these occasions I have also presented my own work, and also at the regular KiDS collaboration meetings in Leiden, Edinburgh and Oxford.

This autumn I will continue to study galaxy clusters with weak gravitational lensing

in the group of dr. Anja Von der Linden at the Stony Brook University in the United States of America.

Acknowledgements

After 11 years at the Leiden Observatory, a page or two cannot suffice to acknowledge all who have in some way contributed to the completion of this thesis. First of all my thanks goes out to the entire institute, which, as an independent reviewer has said, has a very comfortable atmosphere and displays easy interaction between members of different groups, who work on vastly different subjects. I wholeheartedly agree with this statement and I have enjoyed meeting lots of people in the long 11 years I have spent here.

Specific thanks to those who have enjoyed (or endured) my company most. Koen, although we collaborated mostly just for the first half of my thesis, besides our two-weekly telecons, working with you has always been fun and enlightening. In contrast, the frequency of my interactions with professor Hoekstra grew over time and correlated with the decreasing distance between my office and yours, to an almost daily get-together. Given our Dutch inclination to get to the point, these meetings were always productive, but enjoyable too when the conversation diverted to astropolitics. I have grown much as a researcher under your tutelage and thank you for the importance you attach to your role as supervisor.

Jeroen, after embarking on the journey of astronomy together 11 years ago, there is too much to mention about you. Although I (and others with me) found you scary at first, you've been a great friend and a great help during my PhD. Thanks for sticking with me through club power and going for walks in the fresh air until you replaced me with someone better.

Ian, you've shared the KiDS image simulations adventure with me. You've given me the best working week of my PhD; it was both fun and productive. And you have taught me that a Mediterranean style of conversation is sometimes preferable to a Dutch style.

The greatest victim of my club power agenda has been Fabian, who made it clear that he was definitely victim. However, you stuck with me through the suffering and, for me at least, our dinners together, where we discussed the horror of club power, PhD life and we made some fun of general Dutch and German stereotypes, made the whole thing an enjoyable experience.

I could not have been more fortunate to share my time as a PhD with Massimo. You were always available for a pleasant talk, for helpful discussions, to dispense general wisdom and sometimes to give me some insights in astrosociology. My PhD could not have been done without you. Although I wish you the best of times in the outside world, you will be sorely missed.

Cristobal, after working closely together, I cannot but have respect for your achievements as a scientist, given your demanding private life. Thanks for all your invaluable help for this thesis.

Axel, your help was indispensable in getting the start of this thesis finished.

Being a PhD student is about soaking up knowledge and expertise to grow as a researcher and I have been fortunate with the cosmology group in Leiden. It has been instructive to work with you and many thanks for your patient explanations. In a wider sense the KiDS team has provided me with a professional environment in which to work and learn, as well as a great group of people to get to know. Although there

was a lot of work for a relatively small group of people, I'm very proud to have been part of what we accomplished. Thanks to all those who made the Sterrewacht the fun place it has been for me: David, Henriette, Emmanuele, Clement, Leah, Eva, Ann-Sofie, Edwin, Niels, Vincent, Kirstin, Heather, Paula, Andrej, Francois, Allison, Mason, Chris, Michael, Nico, Geert-Jan, Mieke and anyone else who I might have forgotten, but has been there with me in these last 11 years.

A Dutch saying states that a good neighbour is better than a distant friend and this sentiment is applicable to office mates as well. Thanja, talking with you always made me feel better, whatever we discussed; Andrew, although I had to wait until 12.25 we had great fun in the office and I hope I will find as good a friend in my next office; Tim, your presence always seemed to fill up the room; Caroline; Umut; Merel; Joe; Christian, for sharing the fresh air; and Emmanuele, because you have been in my office at least as often as I have been and you made sure there was never a dull moment. Jens, being TA with you made grading silly mistakes an enjoyable part of my PhD. Jens, Marijke, and Margot, thanks for sharing all the weird and wonderous questions the public could throw at us. Xander, even though our meetings were not something I looked forward to, you have been a great help in getting this thesis finished and putting me at ease about it. Lastly, it should never be forgotten that a big part of life at the Sterrewacht is thanks to the computer group and the secretaries.

Although the PhD can be an all-consuming project, there is still a life outside it. I want to thank the M(W)BOA crew for a staple source of food and fun on Wednesdays and Huib for the same on Tuesdays. And even though I wasn't around as often as I would have wanted to, and wasn't as fit as you probably would have wanted me to be, I have to thank everyone at Argonauta for making it a fantastic club. Special thanks of course to my team, with which we've accomplished so much in the same time span as my PhD. Robert, together we have had some awesome vacations in exotic locations and discussions with you about life as a PhD and as a post-doc have helped me choose my path. I will see you soon in the USA. Luc, you have always been a great friend for the past 17 years and you got me started with astronomy in the first place. This thesis would definitely not exist without you.

Home is where the heart is, as well as the sarcastic remarks and the barely acceptable humour. Michiel and Elanor, thanks for a home and hours of video entertainment, some of which was actually not too bad. Mother, you're maybe not completely a tiger mother, but thanks for pushing me beyond my normal lazy state of being. Thanks for the unwavering support in every aspect of my life. Now it is time for the next step after my PhD, and I could not have been happier to share it with you, Mette. Thank you for all the support and I look forward to our life together in the USA.