Universiteit
Leiden
The Netherlands

Light weighed: on the statistics and systematics of weak gravitational lensing

Smit, D.M.

Citation

Smit, D. M. (2021, December 8). *Light weighed: on the statistics and systematics of weak gravitational lensing*. Retrieved from https://hdl.handle.net/1887/3245825

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/3245825

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

Light Weighed:
On the Statistics and Systematics of Weak Gravitational Lensing
by
Galaxy Groups and Clusters

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof. dr. ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op woensdag 8 december 2021
klokke 13:45 uur

door

Daniël Merijn Smit-van Leusden

geboren te Haarlem
in 1977

Promotiecommissie

Promotores:	Prof. dr. Koenraad Kuijken	
	Prof. dr. Henk Hoekstra	
Overige leden:	Prof. dr. Huub Röttgering	
	Prof. dr. Tim de Zeeuw	
	Prof. dr. Catherine Heymans	(University of Edinburgh,)
	Prof. dr. Hendrik Hildebrandt	(Ruhr-Universität Bochum)
	Dr. Mario Radovich	(INAF - Osservatorio di Padova)

ISBN 978-94-6419-389-3
©2021 by Merijn Smit

"The dreams of childhood – its airy fables,
its graceful, beautiful, humane, impossible adornments of the world beyond;
so good to be believed in once, so good to be remembered when outgrown."
– Charles Dickens, *Hard Times*

Voor Marije,
voor alle dromen die we samen werkelijkheid maken

Voor Annelie,
dat jouw dromen en verbeelding eindeloos mogen zijn

In herinnering aan mijn vader George,
als dromer, lezer en denker in grootse verhalen

"Why do you go away? So that you can come back.
So that you can see the place you came from with new eyes and extra colors.
And the people there see you differently, too.
Coming back to where you started is not the same as never leaving."
– Terry Pratchett, *A Hat Full of Sky*

Printed by: Gildeprint – The Netherlands

Cover:
The foreground is an engraving of a "portable" camera obscura in Athanasius Kircher's *Ars Magna Lucis Et Umbrae* (1645).
The background is the strong lensing cluster Abell S1063, as observed by the NASA/ESA Hubble Space Telescope as part of the Frontier Fields programme. *Credit: NASA, ESA, and J. Lotz (STScI).*

Contents

1 Introduction **1**
- 1.1 Now you see me... 2
  - 1.1.1 Observations and bias 2
  - 1.1.2 Dark matter 4
- 1.2 Weak gravitational lensing 7
  - 1.2.1 The basics of gravitational lensing 7
  - 1.2.2 The measurement of shapes 12
  - 1.2.3 The inference of weak shear 16
  - 1.2.4 Results and interpretation 17
- 1.3 Thesis outline 18
  - 1.3.1 Supergroup SG1120–1202 18
  - 1.3.2 Low redshift galaxy groups in ZENS 19
  - 1.3.3 Optimal statistics for weak shear 19
  - 1.3.4 AMICO clusters in KiDS-450 20

2 Supergroup SG1120–1202 **21**
- 2.1 Introduction 22
- 2.2 Weak lensing framework 23
- 2.3 Data 24
  - 2.3.1 HST imaging 24
  - 2.3.2 Spectroscopy 25
  - 2.3.3 Subgroups 26
- 2.4 Analysis 27
  - 2.4.1 KSB+ shape measurements 27
  - 2.4.2 Redshift distribution 29
  - 2.4.3 Lensing analysis 33
- 2.5 Results 34
  - 2.5.1 Matter distribution 34
  - 2.5.2 Individual groups 35
- 2.6 Summary 38
- 2.A Quality and selection criteria for background sources 39

3 Low redshift galaxy groups in ZENS **41**
- 3.1 Introduction 42
- 3.2 Data 43
- 3.3 Methods 44
  - 3.3.1 Weak gravitational lensing terminology 45
  - 3.3.2 Shapelets 46
  - 3.3.3 KSB+ 47
- 3.4 Shape Measurements 47
  - 3.4.1 Robustness 48

3.4.2 Possible systematics . . . 51
3.4.3 Background redshift distribution . . . 54
3.5 Gravitational shear signal and mass estimates . . . 57
3.5.1 Convex Hull Peeling . . . 58
3.5.2 Group centre of mass . . . 59
3.5.3 Velocity dispersion . . . 61
3.5.4 Mass and concentration . . . 62
3.5.5 The heaviest groups . . . 63
3.6 Conclusions . . . 64
3.6.1 Shape measurements . . . 64
3.6.2 Measured signal . . . 65
3.6.3 Profile shapes . . . 65
3.6.4 General conclusion . . . 65
3.6.5 Future work . . . 66

4 Optimal statistics for weak shear **69**
4.1 Introduction . . . 70
4.2 Weak lensing . . . 72
4.3 Statistical framework . . . 75
4.3.1 Bias, efficiency and robustness . . . 75
4.3.2 Estimators . . . 77
4.3.3 Fourier mode fitting . . . 80
4.4 Simulations and data . . . 81
4.4.1 Simulated ellipticity distributions . . . 82
4.4.2 Data: CFHTLenS . . . 84
4.4.3 Simulated noise . . . 85
4.5 Results . . . 87
4.5.1 Central value estimation . . . 87
4.5.2 Fourier mode fitting . . . 92
4.6 Conclusions and summary . . . 94
4.6.1 Optimal estimators . . . 94
4.6.2 Direct Fourier mode fitting . . . 94
4.6.3 Future considerations and possible applications . . . 95
4.A Bias estimations . . . 96
4.B Efficiency estimations . . . 96
4.C Estimations from Fourier mode fitting . . . 96

5 AMICO galaxy clusters in KiDS-450 **103**
5.1 Introduction . . . 104
5.2 Weak gravitational lensing statistics . . . 106
5.2.1 Principles of weak lensing . . . 106
5.2.2 Estimation of the surface density profile . . . 110
5.2.3 Statistical framework . . . 111
5.2.4 Halo model . . . 113
5.3 Data and analysis . . . 114
5.3.1 Lenses . . . 115

5.3.2 Sources . . . 115
5.3.3 Implementation . . . 118
5.4 Results . . . 120
5.4.1 ESD profiles . . . 120
5.4.2 Bias and efficiency . . . 120
5.4.3 Halo masses . . . 122
5.4.4 $L_{200} - M_{200}$ scaling relation . . . 123
5.5 Summary and conclusions . . . 124
5.5.1 $L_{200} - M_{200}$ relation . . . 125
5.5.2 Optimal estimators . . . 127
5.A Tests for systematics . . . 128
5.A.1 Photometric redshift . . . 128
5.A.2 Contamination of the background sample by cluster galaxies . . . 129
5.A.3 Individual bin ESD profiles and cross signals . . . 129
5.A.4 Tile bootstrap . . . 130
5.B Analysis of dependence on outer data points . . . 130

6 Nederlandse samenvatting **133**
6.1 "De hemel bij nacht" . . . 134
6.1.1 Waarnemingen en vertekening . . . 134
6.1.2 Donkere materie . . . 135
6.2 Gravitatielenzen . . . 136
6.2.1 Kromming van ruimtetijd . . . 136
6.2.2 Metingen en interpretatie . . . 137
6.3 Dit proefschrift . . . 139

Bibliography **143**

Publications **153**

Curriculum Vitae **155**

Acknowledgements **159**

1

Introduction

In his dialogue *Parmenides* (129B), Plato lets his master Socrates state:

"... nor, again, if a person were to show that all is one by partaking of one, and at the same time many by partaking of many, would that be very astonishing. But if he were to show me that the absolute one was many, or the absolute many one, I should be truly amazed."

In Plato's Theory of Forms, it is argued that our empirical, sensible observations can only be of the many possible different reflections of an insensible, unique universal Form, an absolute and unchanging concept, outside the limits of our physical space-time.

If any branch of the natural sciences is restricted, in its quest for a truthful description of reality, by the glimpses granted by the Universe, it is astronomy. Astronomers may improve their instruments and statistical techniques, but cannot set up a controlled experiment on most celestial objects, change their position in our Milky Way or follow the billion year evolution of a singular object beyond 'momentary' observations lasting mere centuries or decades. In Plato's well known metaphor, we can only see the shadows on the wall of the cave and hope to discover the true reality outside it.

Plato's contemplations were of a metaphysical nature, and the comparison with astronomy would end here, but it should be noted that it is no coincidence that astronomy was part of the *quadrivium*, four arts (the other three being arithmetic, geometry, and music) required for admittance to his *Akademia*.

This thesis modestly attempts to combine scientific research, by which we mean the aforementioned description of our universe, with theoretical considerations of the statistical methods used for that research, or how we can derive that description from the reflections that we see. In this work, we focus on the matter distribution in groups and clusters of galaxies, and consider the intricacies of the method of weak gravitational lensing that we use, respectively.

This introduction starts with a conceptual overview in section 1.1, after which we give a more mathematical and technical summary of these subjects in section 1.2. Section 1.3 then gives a short outline of the scientific chapters.

1.1 Now you see me...

Astronomy is one of the oldest natural sciences. The mysteries of the night sky have captured the imaginations of ancient civilizations since prehistoric times[1].

One of the oldest stories of observations and interpretations is that of the Pleiades, "Seven Sisters" in Greek mythology, an open star cluster designated as M45, which may have inspired humans for 100,000 years (Norris & Norris 2021). The cluster consists of more than 1000 confirmed members, of which formally[2] ten can be seen by the naked eye (Kyselka 1993, Norris & Norris 2021), but in practice, only six[3] are visible to most people with good visual acuity in a dark night. Examples are Galileo's depiction of the Pleiades in his *Sidereus Nuncius* (1610), where he indicates six visible stars among 36 observable through his telescope, and the Greek Aratus of Soli in the third century BC, who reported that "only six [sisters] are visible to the eyes" (Krupp 1991). In Greek mythology, this is explained by one of the sisters (Merope) hiding from their pursuer Orion, the hunter. It is very likely that the cluster was observed and named, before the story was associated with it (Hard 2004), so why tell a story about seven sisters, when only six are seen?

Might this be an example of *confirmation bias*, that is, an observer interpreting what is seen in a way that 'fits the story' best, thereby confirming preexisting theories or beliefs? Is this mythological story adjusted, so it could fit the observations? This explanation is contradicted by the strong similarity, suggesting a common origin, of stories on a "lost sister" or "daughter"[4] among many cultures around the globe, even those that had not been in contact for 100,000 years (e.g. Aboriginal Australian cultures predating European contact, Burnham 1978, Gibson 2017).

An alternative hypothesis is posed by Norris & Norris (2021), who calculate from data of the Hipparcos satellite (Van Leeuwen 2009) that, because of the proper motion of Pleione, that star was 100,000 years in the past 8.4 arcminutes away from the much brighter Atlas, an angular distance that is more than a quarter or a full moon. Nowadays, the two are so close, that the glare from Atlas prevents Pleione from being seen by the naked eye under most circumstances.

If this hypothesis is true, then there is no mystery to the story of a "lost sister". That 'mystery' is supposed by us, because the cultural stories surrounding M45 do not match *our* observations, but they matched observations at the time the *root story* might have originated in Africa. This would mean that such an *observer's bias* would be ours.

1.1.1 Observations and bias

This thesis focuses on systematic effects, statistics and subsequent interpretation. In the context of the latter, we have already used the terms *confirmation bias*, a predisposition to prior beliefs, and *observer's bias*, an indication of limitations in the available information or perspective, and therefore categorized under *information bias*. In our interpretation of the general term, a bias is a discrepancy between an observation or interpretation, and the 'truth'.

[1] At least, that is our current *interpretation* of some artifacts, sites and their orientations.

[2] Using a criterion of an apparent 'visual magnitude' brighter than six ($m_V < 6$).

[3] In fact, the *Greek* interpretation includes Atlas, the father of the sisters, among those six, reducing the number of visible sisters to five in that cultural version.

[4] Or occasionally indeed two.

Every student has heard of the Copernican Revolution, the changing of our view of the universe from geocentric, revolving around Earth, to heliocentric, revolving around the Sun. The Ptolemaic geocentric model, in basis a set of circles with the Earth as center, upon which moved the Sun and planets, had known problems in its description of reality. Some of these, the observed change of speed and retrograde motions of objects such as planets, where the direction of motion on the sky seems to reverse, could not be explained without adding adjustments that compromised the philosophical elegance of the basic system. See Figure 1.1.

One such an adjustment was the addition of epicycles, little circles that themselves moved over the larger, basis circles. Another was the introduction of equants, extra points beside Earth, around which objects moved at constant angular speed, while moving around other points, the deferents, with constant distance (i.e. circles).

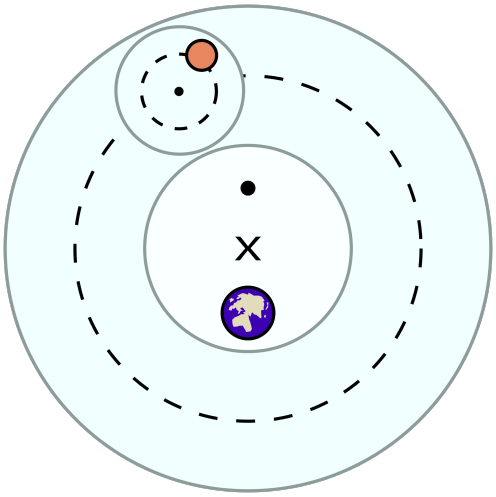


Figure 1.1: Representation of a planetary motion in the Ptolemaic geocentric model. The planet moves in a small epicycle. The center of that epicycle, shown as a small dot $\cdot$, moves in turn in a circle around the deferent, shown as $\times$, and with constant angular speed around the equant, shown as a bullet $\bullet$. This image has been acquired in the public domain.

Copernicus presented the heliocentric model in 1543 as a mathematically much more elegant model and a better description of reality, explaining for instance retrograde motion as a natural consequence of the Earth's motion. However, an empirical model needs not only describe observations, but also predict them, and a major source of criticism for Copernicus's model was the need of many more epicycles to match new observations. The reason for this was the continued use of circles as a basis model. It was not until Kepler's laws of planetary motions in 1609, that circular orbits where replaced by elliptical ones, removing the need for epicycles and equants altogether.

As such, Copernicus's original heliocentric model is one of the most famous examples of *model bias*, a discrepancy that arises because the model is intrinsically not suited to describe the situation. Model bias is a recurring point of interest throughout this work. In chapter 2, we assess the assumption that we can model groups of galaxies that lie close to each other as separate objects, ignoring the influence they have on each other. In chapters 2, 3, and 5,

we discuss how we can model the center of groups of galaxies, as mis-centering can have a severe impact on our results.

In astronomy, one of the best known biases is the Malmquist bias, the fact that intrinsically brighter objects are easier to detect and therefore can be seen to a greater distance (Malmquist 1925). Looking back through our Universe, greater distance also means earlier times, so the Malmquist bias can create a false impression of evolution. If one doesn't account for this effect, the sample is said to suffer from *selection bias*. We consider a variant of this effect in chapter 2, where we estimate how many observed sources in the background of a supergroup of galaxies might actually be faint, instead of far away, and belong to the structure, throwing off our calculations.

In chapter 4 and 5, we discuss *statistical bias*, where it is not limitations in information or physical models that causes a biased result, but our *interpretation* of calculations. A *statistic* is nothing more than the result of an algorithm, usually a calculation, performed on a sample of values. The arithmetic mean is the best known example. We then interpret the meaning of this result. This a form of descriptive statistics, that aims to describe features of a population or a sample of that population, such as an 'average value' that best represents the sample.

As a well known example, when discussing salaries of the 'average' working citizen in a country, we have to take into account that there is a strict minimum wage (even if that minimum is zero), but not a maximum, with executives of large concerns or soccer players earning millions or even tens of millions. When the average income of a country is discussed, the median income is much more descriptive of the population and the mean can give a seriously biased interpretation of the people's prosperity.

The calculations performed are correct in themselves. The results just don't give an accurate description of reality. In chapters 4 and 5, we compare several descriptive statistics, or *estimators*, on samples of galaxy ellipticities. While these methods are all mathematically correct, they may not be equally appropriate to be used in a description of the matter distribution.

1.1.2 Dark matter

Astronomy studies many objects that cannot be observed directly, not even with a telescope, or not yet. The dark patches one sees when looking at the band of the Milky Way at night, known as The Great Rift, used to be thought of as emptiness, until observations (Barnard 1906) proved that they were actually obscuring dust clouds, that can nowadays be observed directly in infrared. In contrast, black holes, by their nature, can never be observed directly[5], only indirectly by energetic phenomena just outside their event horizon.

Before this century, planets outside our solar system were too small to be observed directly, until the observation of Fomalhaut b (Kalas et al. 2008) by the Hubble Space Telescope (HST). Before that, their existence could be deduced indirectly, e.g. from the regular dimming of the star they orbit when they pass in front of it, or the 'wobble' in the star's motion from their gravitational interaction. This requires the exoplanet to be sufficiently large or massive and our view of its orbit to be more or less from the side. This biased early samples of exoplanet candidates to massive planets, orbiting their star very closely.

In the same manner, one of the biggest mysteries in the universe to date (and a major subject of this thesis) was discovered: *dark matter*. The first indication of unseen matter came

[5]Which is why they are often called black hole *candidates*, as their nature cannot be confirmed by direct imaging.

from observations of stellar motions in our own Milky Way, from which Lord Kelvin deduced in 1884 that there must be many more stars than could be seen, so that "many of our supposed thousand million stars, perhaps a great majority of them, may be dark bodies" (Kelvin 1904). Kapteyn (1922) first coined the term "dark matter" to explain the distribution of stars and velocities, as seen from the solar neighborhood. Although it is sometimes mentioned that the work of Oort (1932) confirmed the hypothesis of Kapteyn, it was shown that these and similar works suffered from one or both of selection or model bias and that these works did not prove the need for dark matter (Kuijken & Gilmore 1989), but the concept was born.

Arguably the first real evidence for dark matter came from the work of Zwicky (1933), who studied the orbital motions of galaxies in the Coma cluster. Zwicky concluded from their large velocities that the visible mass, the known matter of stars, gas and dust, also known as *baryonic matter*, was not enough by a factor of more than 400 to keep them in place, instead of flying off. (Schwarzschild 1954) also found the mass-luminosity ration of the Coma cluster to be "bewilderingly high".

Freeman (1970) and Bosma (1981) found a similar effect when studying the rotation curves of spiral galaxies. If there was no mass beyond the disk, one would expect the rotational velocity to fall off the further one observed from the central mass distribution. However, the observed rotation curves remained flat far beyond the visible disks. Rubin (1983) made the connection to Zwicky's dark matter and discussed the implications for the geometry of the universe.

So far, the arguments for the existence of this unknown dark matter came from dynamical considerations. A next fundamental discovery came from observations of the Cosmic Microwave Background (CMB), first observed by Penzias & Wilson (1965) and mapped by the COBE satellite (Mather 1982). The CMB can be thought of as an afterglow of the Big Bang, created when the hot ionized plasma in the universe cooled down sufficiently to recombine into the matter we see today. From observations by the WMAP satellite, Spergel et al. (2003) and Hinshaw et al. (2007) showed that of the total matter in the universe, less than 20% was in the forms of baryons, and the rest was dark matter.

In fact baryonic and dark matter together only provide $\sim 30\%$ of the matter-energy content of the universe, while $\sim 70\%$ of the content of the universe is energy[6]. In this thesis, we use cosmological values consistent with results from the Planck satellite (Planck Collaboration et al. 2014), with 4.9% ordinary matter, 26.6% dark matter and 68.5% energy. So far, the nature of and observations of dark matter elude us, as elementary particle physics have no conclusive theoretical explanation and our instruments have not yet been able to detect it. We know it's there and we know it interacts through gravity, but so far there doesn't seem to be any interaction with electromagnetic radiation or baryonic matter.

The curvature of space-time

As dark matter can only be observed indirectly by its gravitational interactions, it follows that we rely on indirect methods to study the total distribution of all matter. Our method to study this distribution makes use of gravitational lensing, an effect of the curvature of spacetime, most accurately[7] predicted by Einstein's theory of general relativity (e.g. Eddington 1920,

[6] In fact, the nature of this energy is also 'dark', i.e. unknown and unobserved, possibly in the form of a cosmological constant Λ, a term first used by Einstein (1917)

[7] Classical theory, already by Newton in 1704, also predicts the bending of light rays, but is off by 50%.

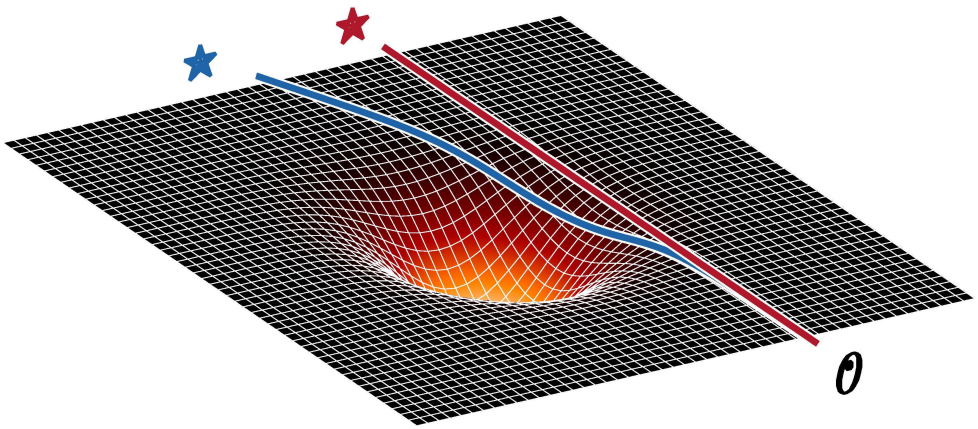


Figure 1.2: The curvature of spacetime causes the light rays from a source in the background (blue) to deflect from a straight line. An observer perceives an image (red) displaced from the position of the source.

Chwolson 1924).

In the general relativistic description, the geometry of spacetime is determined by the presence and state of energy and matter. More simply put, spacetime is bent due to a concentration of mass, an effect that we experience as gravity. Because light rays travel along the shortest path available, this path is no longer straight in a curved geometry (see Fig. 1.2). This makes gravitational lensing a direct probe of the geometry of the universe and the distribution of *all* matter, without the need to make assumptions about the astrophysical or dynamical state of observed phenomena.

The first observations of gravitational lensing were made by Dyson et al. (1920), during the solar eclipse on the 29th of May, 1919. Around this time, it was discovered that our Milky Way was just one of hundreds of billions (Lauer et al. 2021) of galaxies, that form the building blocks of the universe (Slipher 1915, Curtis 1917). Zwicky (1937) suggested that galaxies would be massive enough[8] lenses to make this effect easier to observe, and numerous enough to be likely candidates to be lensed. Measurement of this variant, aptly called galaxy-galaxy lensing, was first attempted by Tyson et al. (1984) and finally detected by Brainerd et al. (1996).

This thesis focuses on lensing by more massive structures, like groups and clusters of galaxies, first detected by Tyson et al. (1990). Perhaps one of the most powerful examples of gravitational lensing as a way to probe the matter distribution independent of astrophysical assumptions is given by the bullet cluster (Clowe et al. 2006, see Fig. 1.3).

The bullet cluster system actually consists of two clusters of galaxies that have passed through each other, the smaller cluster moving at higher velocity to the right being considered the ‘bullet’. The groups of galaxies of these clusters can pass collisionlessly through each other and appear unaffected, but the intracluster gas, where most of the known, baryonic mass in these clusters resides (Clowe et al. 2006), lags behind. Their reconstruction of the

[8] Typical luminosities of bright galaxies like our own Milky Way are 100 billion times the luminosity of the Sun, $L_\star \sim 10^{11} L_\odot$.

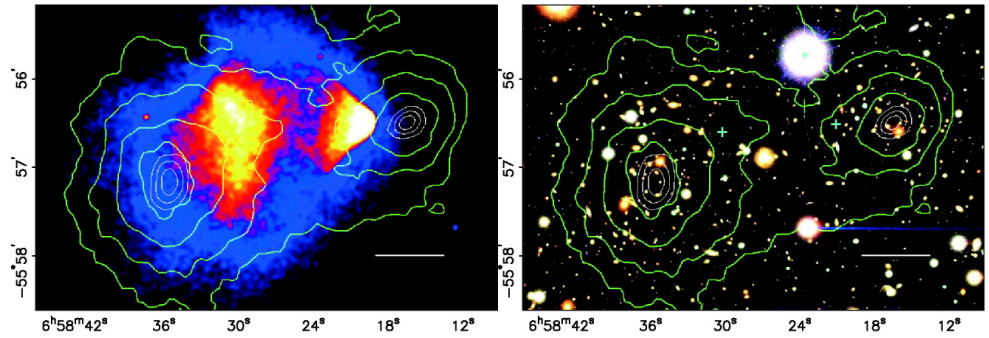


Figure 1.3: *Left.* A reconstruction of the matter distribution (green contours) using gravitational lensing compared to the X-ray emission of the hot gas (color scale), where most of the visible, baryonic matter of this system resides. *Right.* The same reconstruction, now compared to the location of the galaxies that make up the two subclusters. These images show that most of the total matter does not reside where most of the baryonic matter is observed, a clear indication that the major part is 'dark' and collisionless. Originally published in Clowe et al. (2006).

matter distribution using weak gravitational lensing clearly show that most of the total matter does not reside where most of the baryonic matter is observed, a clear indication that the major part is 'dark'.

1.2 Weak gravitational lensing

In this section, we introduce the mathematical framework of gravitational lensing, as well as the principles of weak lensing we have used in our work. We also present the central concepts to our statistical approach and our tests for systematic effects. We introduce terminology and notation conventions used in this thesis.

1.2.1 The basics of gravitational lensing

We start with an analytic derivation of the framework of weak lensing, the main observables and concepts that then form the basis for the research in this thesis. We refer the reader to excellent reviews such as Bartelmann & Schneider (2001), Schneider (2006), Hoekstra & Jain (2008), Bartelmann & Maturi (2017), for more in-depth approaches.

As the lensing effect is caused by rays of light being deflected by the curvature of space-time due to mass inhomogeneities along their path, we consider how a mass overdensity acts on the light rays from distant sources behind it. Figure 1.4 shows a simple representation of a gravitational lens system.

For the purposes of this work, the extent of the lensing mass along the line of sight, compared to the distances from observer to lens and from lens to background source, is negligible. In this so-called thin-lens approximation, we can describe the path of light by straight line segments. In this representation, we use angular-diameter distances D_l from observer to lens, D_ls from lens to background source[9], and D_s from observer to background

[9]More accurately, the distance to the source *plane* perpendicular to the line of sight from observer to lens, but in

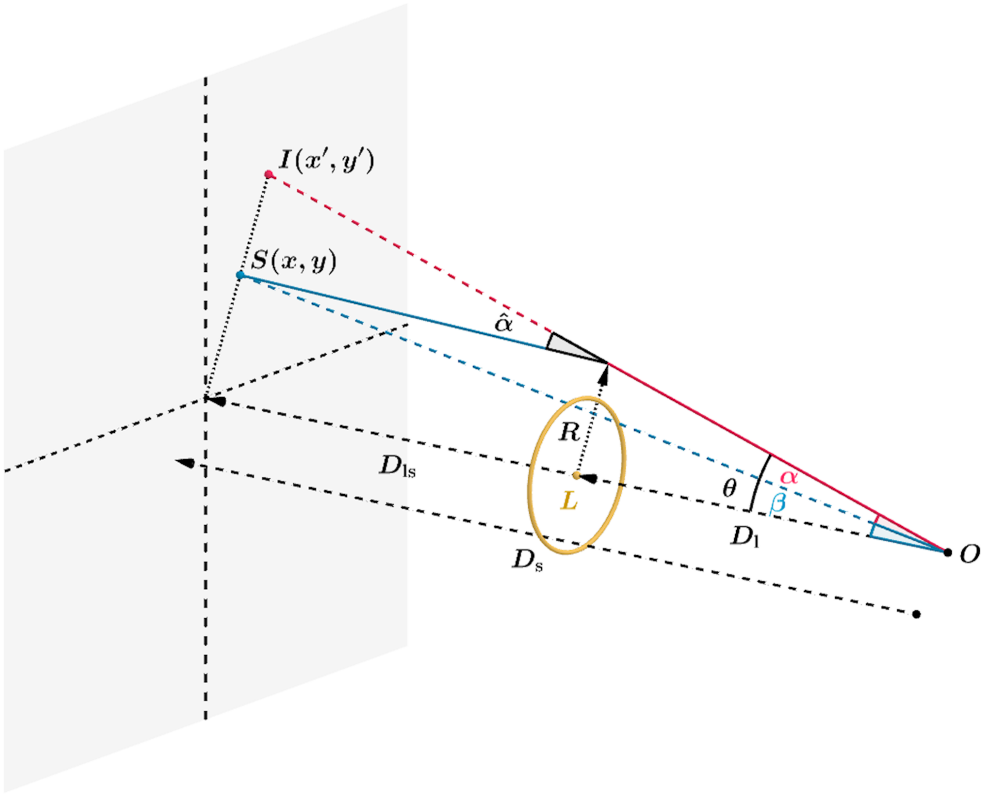


Figure 1.4: Representation of a gravitational lens system, showing the displacement of a source at position $S(x, y)$ to an image at position $I(x', y')$, where we take the origin of the source plane collinear with the positions of the lens L and the observer O. Based on Smit et al. (2021).

source.

The deflection angle $\hat{\vec{\alpha}}$ is determined by the gradient of the gravitational potential Φ perpendicular to the path of light, integrated along that path:

$$\hat{\vec{\alpha}} = -\frac{2}{c^2}\frac{1}{D_l}\int \vec{\nabla}_\theta \Phi \, dl \,, \tag{1.1}$$

where $\vec{\nabla}_\theta$ is the two-dimensional gradient in angular coordinates perpendicular to the line of sight and the angular-diameter distance factor of D_l^{-1} arises from the conversion of physical to angular coordinates in the case of small angles.

This leads to an angular displacement, as seen from the observer and again using the small-angle approximation (see Fig. 1.4)

$$\vec{\alpha} = -\frac{D_{ls}}{D_s}\hat{\vec{\alpha}} \,, \tag{1.2}$$

this approximation, these are the same.

also called the reduced deflection angle, relating the observed position $\vec{\theta}$ of a distant point source to its unlensed position $\vec{\beta}$ by the lens equation

$$\vec{\beta} = \vec{\theta} - \vec{\alpha}. \tag{1.3}$$

Under the thin-lens approximation, the deflection of light rays by the lensing mass is then described by $\vec{\alpha} = \vec{\nabla}_\theta \psi$, where

$$\psi = \frac{2}{c^2} \frac{D_{\rm ls}}{D_{\rm l} D_{\rm s}} \int \Phi \, dl \tag{1.4}$$

is the two-dimensional, dimensionless, lensing potential. In this way, the deflection angle can be related to the density of the lensing mass and to observable distortion of sources in the background, as follows, starting with the latter.

The differential effect of the deflection of light on the images $I(x, y)$ of background galaxies, which are extended sources, can to first order be described as a coordinate transformation, by taking the derivatives in the lens equation (1.3) of the original position β with respect to the observed position θ. Substituting $\vec{\nabla}_\theta \psi$ for $\vec{\alpha}$, we obtain the Jacobian of the lens mapping,

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 - \psi_{11} & -\psi_{12} \\ -\psi_{21} & 1 - \psi_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \tag{1.5}$$

with

$$\psi_{ij} = \frac{\partial^2 \psi}{\partial \theta_i \partial \theta_j}, \tag{1.6}$$

resulting in the lensed image $I(x', y')$, which is the key observable in our work.

Critical surface mass density and convergence

To interpret the effect on the source image, we note that the linear, symmetric coordinate transformation in eq. 1.5 can be decomposed in three parts, namely the identity I and two perturbations, consisting of an isotropic part describing a magnification, and an anisotropic, traceless part, describing a shearing of the image:

$$I - \frac{1}{2}(\psi_{11} + \psi_{22}) I + \begin{pmatrix} -\frac{1}{2}(\psi_{11} - \psi_{22}) & -\psi_{12} \\ -\psi_{21} & \frac{1}{2}(\psi_{11} - \psi_{22}) \end{pmatrix} \tag{1.7}$$

To relate ψ_{ij} to the density of the lensing mass, we start with the isotropic term, which is half the Laplacian of the lensing potential: $\frac{1}{2}(\psi_{11} + \psi_{22}) = \frac{1}{2}\nabla^2_\theta \psi$. Using equation 1.4 and the thin-lens approximation, introducing a factor of $D_{\rm l}^2$ due to conversion between angular and physical coordinates, we obtain

$$\frac{1}{2}\nabla^2_\theta \psi = \frac{1}{c^2} \frac{D_{\rm l} D_{\rm ls}}{D_{\rm s}} \int 4\pi G \rho \, dl \,, \tag{1.8}$$

which is a dimensionless quantity. Defining the surface mass density as

$$\Sigma \equiv \int \rho \, dl \tag{1.9}$$

and gathering the remainder of the right-hand side into

$$\frac{4\pi G}{c^2}\frac{D_\mathrm{l}D_\mathrm{ls}}{D_\mathrm{s}} \equiv \Sigma_\mathrm{cr}^{-1}\,, \tag{1.10}$$

with Σ_cr called the critical surface mass density, we find that the isotropic term can be written as

$$\kappa \equiv \frac{1}{2}\nabla_\theta^2\psi = \frac{\Sigma}{\Sigma_\mathrm{cr}}\,, \tag{1.11}$$

where we recognize κ as a normalized dimensionless surface mass density. Recognizing that $\nabla_\theta^2\psi = \vec{\nabla}\cdot\vec{\alpha}$ is the divergence of the deflection of the light rays, or the manner in which those light rays converge due to the lensing effect, κ is simply called the convergence.

Shear and intrinsic ellipticity

The shear matrix in eq. 1.7 has two independent components $\gamma_1 = \frac{1}{2}(\psi_{11} - \psi_{22})$ and $\gamma_2 = \psi_{12} = \psi_{21}$, with $\gamma \equiv \gamma_1 + i\gamma_2$ called the complex shear. Eq. 1.5 then becomes

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}. \tag{1.12}$$

This transformation leads to magnification and distortion of the light distribution of background sources. Weak lensing magnification analyses (e.g. Hildebrandt et al. 2009, Van Waerbeke et al. 2010, Hildebrandt et al. 2011) require the intrinsic (distribution of) source sizes or magnitudes. In weak shear analyses, the focus lies on the net distortion or reduced shear $g = g_1 + ig_2 \equiv (\gamma_1 + i\gamma_2)/(1-\kappa)$:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = (1-\kappa)\begin{pmatrix} 1-g_1 & -g_2 \\ -g_2 & 1+g_1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}, \tag{1.13}$$

where the transformation is written as a multiplication of $(1-\kappa)$ and a distortion matrix describing the alignment of lensed sources in the foreground potential.

The effect on a circular source is a shearing into an ellipse with axis ratio $q = \frac{b}{a}$, where

$$q = \frac{1-|g|}{1+|g|} \quad \Leftrightarrow \quad |g| = \frac{1-q}{1+q} = \frac{a-b}{a+b}\,, \tag{1.14}$$

and position angle φ via

$$g = |g|\,(\cos 2\varphi + i\sin 2\varphi)\,. \tag{1.15}$$

See Fig. 1.5.

This gravitational distortion cannot be measured directly in practice. Galaxies that are used as background sources, have an intrinsic shape distribution and we can only measure the combined effect of their intrinsic shape and a weak lensing distortion. While galaxies often have complex morphologies, it is adequate to describe images by their quadrupole brightness moments or their ellipticities, and their respective response to weak shear distortions. A common definition[10] of the shape of an image with elliptical isophotes is the ellipticity ϵ =

[10] An alternative definition of ellipticity is often denoted as $|\chi| = \frac{1-q^2}{1+q^2}$, related to the geometrical eccentricity, and called polarization (e.g. Seitz & Schneider 1995, Viola et al. 2014, and section 1.2.2).

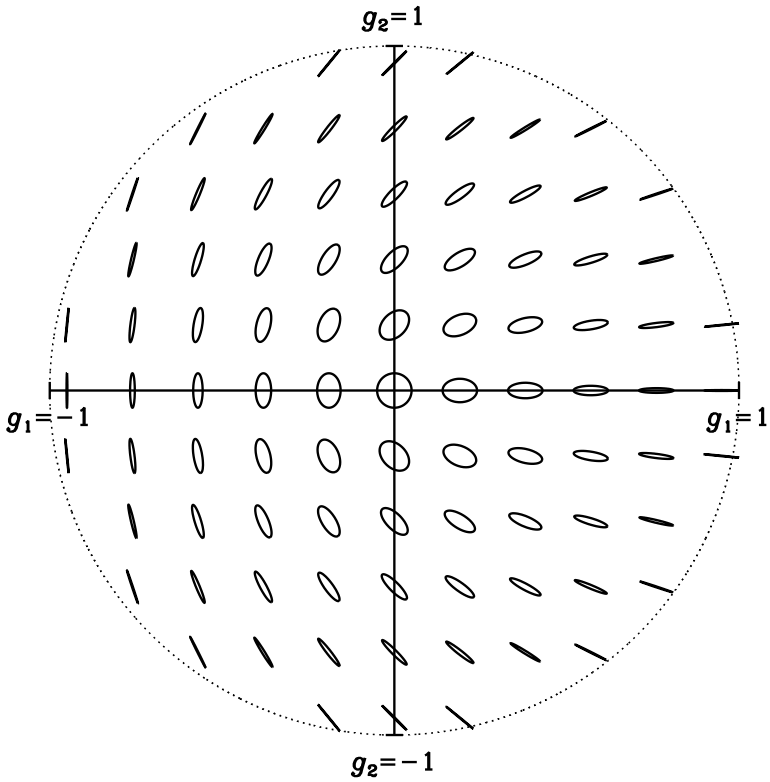


Figure 1.5: The effect of shear on an intrinsically round source. The g_1 component stretches the image horizontally or vertically. The g_2 component stretches the image diagonally.

$\epsilon_1 + i\epsilon_2$, defined as the reduced shear needed to create this image from an image with circular isophotes (Bernstein & Jarvis 2002, Kuijken 2006).

The complex notation gives a most straightforward formulation of the ellipticity ϵ that results after transforming an image with intrinsic ellipticity ϵ^I with a distortion g. As shown by Seitz & Schneider (1997),

$$\epsilon = \frac{\epsilon^I + g}{1 + g^* \epsilon^I} \quad \text{for} \quad |g| \leq 1\,, \tag{1.16}$$

with g^* the complex conjugate of g.

The intrinsic shape distribution is called shape noise and assuming no preferred direction on the sky, should average to zero: $\left\langle \epsilon^I \right\rangle = 0$. Seitz & Schneider (1997) showed that, for eq. 1.16, the mean[11] $\langle \epsilon \rangle$ does not depend on the intrinsic shape distribution $P(\epsilon^I)$. Each background shape measurement ϵ is then an unbiased, albeit very noisy, estimate of the reduced shear g, but only in the absence of further sources of noise, that would alter the observed distribution of ϵ.

[11] Or, in fact, any $n-$th moment $\langle \epsilon \rangle^n$.

In practice, there are always sources of error that manifest as 'noise' in the observed ellipticity values, such as pixel noise of the detector, systematic distortions in the optical system and limitations in modeling algorithms.

1.2.2 The measurement of shapes

While the basis of weak gravitational lensing is given in analytic detail, in practice one has to deal with many systematic effects, when measuring the main observable, the shapes of background sources. As many of these effects depend on specific telescopes, instruments, pipelines and surveys, the most detailed and technical descriptions are given in the corresponding (series of) papers.

Here we aim to give the reader, as a frame of reference for this thesis, a qualitative overview of a selection of effects, that will be analyzed in more detail in chapters 2 and 3. For a more detailed treatise, we point the reader to the references given there and in the following sections.

Images and distortions

At the observer's end in the schematic of Fig. 1.4, the light rays pass through the atmosphere, in case of ground-based observations, and through the telescope optics, before being registered by the detector of the instrument. The response of this total optical 'system' to the received signal is called the point spread function (PSF), its Fourier transform being the optical transfer function. It describes the image of a point source.

The PSF causes a pattern of ellipticity distortions that varies over the field of view, and can be modeled using the images of stars (which are unresolved and can therefore be treated as point sources). The correction is then a deconvolution. This can be done on a single image, if there are enough stars in the field of view to properly sample the PSF pattern.

Observations which have relatively few stars, typically at high galactic latitude, can be corrected using dense stellar fields as reference. Hoekstra (2004) modeled the time-invariant spatial pattern of the PSF by averaging many fields. To properly take the temporal variations into account, a principal component analysis (PCA) can be used (e.g. Jarvis & Jain 2004, Schrabback et al. 2007, 2010). We employ this technique in the weak lensing analysis of imaging data taken with the Advanced Camera for Surveys (ACS) on the HST in chapter 2.

In this way, the term PSF is used as a catchall, and the resulting distortions are oftentimes corrected for in a similar fashion: a correction for the combined PSF pattern, without the need to identify individual causes. As pointed out by Jarvis et al. (2008), a physical model for a known contribution to the PSF may be more accurate than an average or a PCA derived from noisy data. In this context it is worth noting, that both Jarvis & Jain (2004) and Schrabback et al. (2007) interpret their first principal component to indicate telescope focus.

The turbulence of the atmosphere causes a blurring of the image called seeing. If exposure times are long enough, the direct effect on the ellipticity or the PSF is negligible (Heymans et al. 2012a). It is usually the dominant factor in the size of the PSF for ground based observations and therefore effectively sets the limit on the angular resolution of the observations, which in turn limits the number statistics of resolved background sources that can be used for lensing. To give some examples: the seeing of the Wide Field Imager (WFI) data used in chapter 3 varies between $0.75''$ and $1.35''$ with a median of $1.0''$, whereas the

median seeing conditions for the Canada-France-Hawaii Lensing Survey (CFHTLenS, Heymans et al. 2012b) and the Kilo-Degree Survey (KiDS, de Jong et al. 2013) Data Release 3 (KiDS-450, de Jong et al. 2017) used in chapters 4 and 5 are 0.72″ (Erben et al. 2013) and 0.66″ (Hildebrandt et al. 2017), respectively.

Observations from space are not limited in resolution by seeing, but are diffraction-limited, the fundamental physical limit due to diffraction determined by the telescope aperture size and the wavelength observed. The resulting diffraction pattern makes the PSF more complex. An important type of artifact seen in space-based observations is caused by the deterioration of the instrument CCDs, due to constant exposure to cosmic rays, outside the protective atmosphere. This causes a charge-transfer inefficiency (CTI), leading to trails in the CCD readout direction. These trails will affect the measured PSF and shear patterns and need to be corrected for, as is done in e.g. Rhodes et al. (2007), Massey et al. (2010) and in chapter 2.

A final aspect we mention here concerns the translation from instrument (CCD) response to image. The dominant source of noise in the image is pixel noise, mainly due to sky background Poisson noise and CCD readout noise. This increases uncertainties in shape measurements and causes measurement bias in the derived shapes, as the dependence of ellipticity on pixel values is non-linear (see e.g. Refregier et al. 2012, Melchior & Viola 2012, Kacprzak et al. 2012).

Pixel noise can be compensated by increasing exposure time. As a CCD image is quantized on a rectangular grid, an observed field of view is built from a set of dithered exposures. The final image is obtained by stacking these exposures, thus obtaining a longer total exposure time and a higher signal-to-noise ratio. A disadvantage of using such a stacked image, is that the individual exposures have their own PSF patterns, which are then also stacked. Especially in areas of the final, stacked image that are not covered by all individual exposures, due to edges or gaps in the CCD mosaic, the resulting PSF pattern may be discontinuous[12]. We assess this effect in chapter 3. Shape measurement methods such as the *lens*fit pipeline presented in Miller et al. (2013) model the full set of single exposures instead.

Measurement and bias

Given the necessary control of systematics, and the fact that the modeling of these systematics and the intended source ellipticity measurements are based on noisy images, the development of accurate and robust shape measurement methods has been and still remains a major investment in the field of weak lensing (see e.g. Mandelbaum 2018). Community-driven projects, such as the Shear Testing Programme (STEP, Heymans et al. 2006, Massey et al. 2007), and the GRavitational lEnsing Accuracy Testing challenges (GREAT, Bridle et al. 2010, Kitching et al. 2012, Mandelbaum et al. 2015), have led to a decrease in measurement bias and variances and a better understanding of remaining systematic effects and biases.

In this thesis, we explore three particular methods, based on surface brightness moments or model fitting. In chapter 2 and 3, we make use of the KSB method (Kaiser et al. 1995). In chapter 3, we compare KSB and the Shapelets method (Refregier 2003, Refregier & Bacon 2003). In chapter 4 and 5, we make use of the shear catalogs of CFHTLenS and KiDS-450, both derived with *lens*fit (Miller et al. 2007, Kitching et al. 2008).

[12]This problem is confounded when the single exposure times become short enough for the atmosphere to contribute to the PSF ellipticity pattern.

KSB describes images by their second order brightness moments in angular coordinates

$$Q_{ij} = \frac{\int \theta_i \theta_j W(\theta) I(\theta) d^2\theta}{\int W(\theta) I(\theta) d^2\theta} \tag{1.17}$$

with I the surface brightness and W is a certain weight or window function. This gives three independent quadrupole moments, Q_{11}, Q_{22}, and $Q_{12} \equiv Q_{21}$. KSB determines the complex polarization

$$\chi = \chi_1 + i\chi_2 = \frac{Q_{11} - Q_{22} + 2iQ_{12}}{Q_{11} + Q_{22}}. \tag{1.18}$$

This polarization is related to the eccentricity and is clearly zero for a circular source. However, it differs from the definition of the ellipticity ϵ given in section 1.2.1, which can be defined in terms of the second order brightness moments as

$$\chi = \chi_1 + i\chi_2 = \frac{Q_{11} - Q_{22} + 2iQ_{12}}{Q_{11} + Q_{22} + 2\sqrt{Q_{11}Q_{22} - Q_{12}^2}}, \tag{1.19}$$

and the two definitions are related through

$$\chi = \frac{2\epsilon}{1 + |\epsilon|^2} \tag{1.20}$$

This means that $\langle\chi\rangle$ is a biased estimator of g, as it depends explicitly on the distribution of intrinsic shapes χ^I (Schneider & Seitz 1995). However, in the limit of weak shear, where $\kappa \ll 1$ and $\gamma \ll 1$, we have $\gamma \approx g \approx \langle\epsilon\rangle \approx \frac{1}{2}\langle\chi\rangle$. Besides this statistical bias, one has to consider the model[13] bias from approximations in the PSF correction and noise bias due to the non-linear combination of noisy estimates of Q_{ij}.

In the Shapelets formalism, the light distribution of a source is expanded in the orthonormal basis set of Gauss-Hermite functions. This allows for a flexible model and has the advantage that the behavior of these basis functions under simple transformations (such as an applied shear or smearing by a PSF) is well understood. A PSF model $\boldsymbol{P}$ can then be constructed from the shapelet expansions of bright stars in the image. In the implementation of Kuijken (2006), which we use in chapter 3, sources are described as intrinsically circular, with an expansion

$$\boldsymbol{C} \equiv c_0 C^0 + c_2 C^2 + c_4 C^4 + \dots, \tag{1.21}$$

with C^n circular shapelet basis functions (with n always even) and c_n free parameters. This circular model is then transformed by a distortion

$$\boldsymbol{D} \equiv 1 + \epsilon_1 \boldsymbol{S}_1 + \epsilon_2 \boldsymbol{S}_2 + \delta_1 \boldsymbol{T}_1 + \delta_2 \boldsymbol{T}_2, \tag{1.22}$$

where $\boldsymbol{S}_i$ and $\boldsymbol{T}_i$ are the first-order shear and translation operators, as given by Refregier & Bacon (2003). The free translation δ is needed to ensure an optimal centroid fitting. (We remind the reader that for a method using surface brightness moments, like KSB, the centroid is derived from the first order brightness moments.) This model is then convolved by the PSF model and the resulting model $\boldsymbol{M} = \boldsymbol{P} \cdot \boldsymbol{D} \cdot \boldsymbol{C}$ is then fitted to the observed sources.

[13] In this form also called method bias, see Viola et al. (2014).

This forward convolution with $\boldsymbol{P}$ is numerically more stable than deconvolving noisy source images and allows for a propagation of pixel noise, to ensure accurate uncertainty estimates σ_{ϵ_i}. Velander et al. (2011) showed the advantage of the shapelet flexibility, when measuring higher order distortions in diffraction limited HST images. Besides the noise bias due to fitting a (non-linear) model to noisy images, one has to consider model bias, as the shapelet basis functions are chosen for their elegant transformation properties, and are not realistic representations of galaxy shapes. This mismatch is diminished if one allows the expansion to go to higher order, but in reality the flexibility of the shapelets becomes a weakness, as we start fitting noise. In practice, we use a cutoff in the shapelet expansion, which in itself may introduce another (mild) model bias.

*Lens*fit is a Bayesian model-fitting method, using a galaxy model consisting of Sérsic (1963, 1968) bulge and disc components. Besides using the same or similar free parameters like the aforementioned methods, including the ellipticity ϵ, galaxy size and flux, and galaxy centroid, this pipeline also fits the ratio of bulge to disc, partly discriminating between late and early type galaxies (Miller et al. 2013, Fenech Conti et al. 2017). The log-likelihood (or goodness-of-fit) then has the form

$$\log \mathcal{L} = -\sum_i \frac{\left(y_i - S\left[f_\mathrm{B} b_i + (1 - f_\mathrm{B}) d_i\right]\right)^2}{2\sigma_i^2} \tag{1.23}$$

with S the galaxy flux, b_i and d_i the bulge and disc components and f_B the bulge fraction. It uses pixel-based models of the PSF, which allows for discontinuities between CCDs in the CCD mosaic and uses the full information of all single exposures containing the source. The posterior likelihood is then marginalized over the parameters that are not of interest, using assumed prior distributions, leaving a likelihood surface as function of ϵ_1 and ϵ_2. The aim is to alleviate possible model bias by using an adequate set of priors for the parameters to be marginalized over. To avoid a poorly constrained likelihood, in particular a cutoff in the ellipticity is needed, $\epsilon_\mathrm{max} = 0.804$ for CFHTLenS (Miller et al. 2013) and $\epsilon \approx 1$ for KiDS-450 (Fenech Conti et al. 2017), which introduces a bias due to truncation that needs to be calibrated and can lead to visible signatures that affect the most elliptical ϵ (e.g. Smit & Kuijken 2018, chapter 4).

In conclusion, each pipeline has its own strengths and biases, that are addressed per individual cause, just as is recommended by Jarvis et al. (2008) for the modeling of the PSF. The remaining discrepancies with the 'true' reduced shear g, are modeled as a similar catchall, first order approximation $\epsilon_\mathrm{est} = (1 + m)\epsilon_\mathrm{true} + c$, with ϵ_est the estimated ellipticity, c a constant additive bias, and m a multiplicative measurement bias. c may in general be readily corrected for, if it's not position dependent, which is most likely the result of an undersampled PSF (Van Uitert & Schneider 2016). Usually, m is determined using simulations and is dependent on observed properties, such as source brightness and size. These correlations are weak, making estimates of m on a source-by-source basis very noisy. The correction is then done on a sample basis. For instance, if source weights for the measured ellipticities, based on the estimated uncertainties by the pipeline, are w_i and the estimated multiplicative biases are m_i, one generally corrects the estimated lensing signal by a factor $(1 + K)^{-1}$, with K of the form

$$K = \frac{\sum_i w_i m_i}{\sum_i w_i}\,, \tag{1.24}$$

using the same sample of sources (e.g. Viola et al. 2015, Dvornik et al. 2017).

1.2.3 The inference of weak shear

Given a catalog of robust, reliable shapes ϵ_i, the final measured distribution $P(\epsilon_i)$ is modified from the theoretical form $P(\epsilon)$ (eq. 1.16), due to measurement noise, even if the individual measurements are unbiased. At this point, it is instructive to review our definition of bias in a statistical context. An estimate or measurement is said to be unbiased, if the estimator or measurement algorithm is *expected* to yield the 'true' value.

The measurements or estimates can still be noisy, i.e. *individual* shape measurements ϵ_i may have residual discrepancies with the true values ϵ (as a function of $\left(\epsilon^I, g\right)$, eq. 1.16). The *distribution* of estimates (realizations, measurements), however, should be 'centered', in some statistical sense, around that true value, for the measurement method to be called unbiased.

Usually, this definition of center is taken to be the mean or expected value, and the estimator (or measurement algorithm) is said to be mean-unbiased, if for an increasing number[14] of estimations, the mean estimate converges toward the value of interest of the underlying population. Note that we discuss here the distribution $P(\hat{\epsilon})$ of an estimator $\hat{\epsilon}$ of the ellipticity ϵ, that has its own distribution $P(\epsilon)$.

This unavoidable alteration of the sheared ellipticity distribution $P(\epsilon)$ means that the mean, the first moment $\langle\epsilon_i\rangle$, is no longer an unbiased tracer of g. This can partly be understood, by considering that the unbiasedness of $\langle\epsilon\rangle$ stems from the fact that the ellipticity distribution *before* the effect of gravitational shear is irrelevant, while the effect of noise, from whatever source other than intrinsic shape noise, is applied *after* the shear. Since that noise distribution will be (roughly) centered around $(0, 0)$, whether it be from noisy measurement or natural fluctuations in an isotropic universe, the observed distribution will be skewed low (see Fig. 1.6).

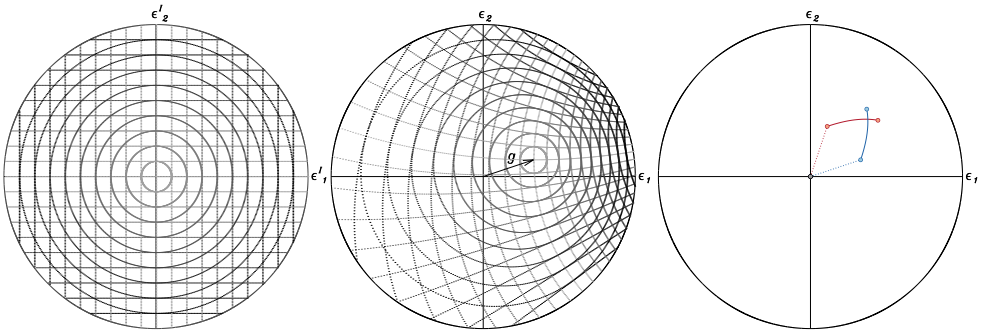


Figure 1.6: *Left and middle*: the non-linear mapping of ellipticities (with $|\epsilon \leq 1|$) by a reduced shear of $g = 0.33 + 0.11i$. *Right*: the noncommutative geometry of ϵ space. In blue, the resulting ellipticity $\epsilon_{a\to b}$ when applying a reduced shear g_b to an intrinsic ellipticity ϵ^I_a. In red, the resulting ellipticity $\epsilon_{b\to a}$ when applying a reduced shear g_a to an intrinsic ellipticity ϵ^I_b. Based on Smit et al. (2015).

At the same time, given that the intrinsic galaxy ellipticity distribution is confined to $\left|\epsilon^I\right| < 1$ and that any $g < 1$ will not cause a $\epsilon \geq 1$, the ellipticity 'space' we consider is a

[14] Since we technically take the limit to infinity, the central limit theorem justifies the use of the mean as measure of center of the *estimator* distribution, and for the use of the shorthand 'unbiased' for mean-unbiased.

highly non-Euclidean, bound but infinite[15] manifold (Bernstein & Jarvis 2002, see Fig. 1.6). One should therefore not speak of a linear difference $\epsilon_b - \epsilon_a$ between two ellipticity values ϵ_a and ϵ_b, but the *excess shearing* (defined by eq. 1.16) needed to transform ϵ_a into ϵ_b. Yet measurement pipelines such as the KSB and shapelets implementations used in chapters 2 and 3, allow for values $\epsilon \gg 1$, and the *lens*fit implementation in chapters 4 and 5 effectively truncates ellipticity space – again, even if the measurements are noisy but unbiased.

In chapters 4 and 5, we investigate the consequences for the inference of the underlying shear signal g.

Estimators and bias

Both observations (e.g. Lambas et al. 1992, Rodríguez & Padilla 2013, Smit & Kuijken 2018) and realistic theoretical models (see e.g. chapter 4) show that the intrinsic distribution of galaxy ellipticities is strongly non-Gaussian, but sharply peaked. For any realistic noise distribution, the shape of $P(\epsilon)$ is changed, but the location of this peak is still an unbiased tracer of the underlying shear. This is a powerful option, if we can accurately determine this location.

We considered several estimators as alternatives to the mean, less sensitive to outliers and/or more sensitive to a centrally peaked distribution. While the principle of the 'right' estimator for a distribution is an appealing one, in practice no estimator is generally unbiased, as noise distributions vary and are dependent on instrument and shape measurement pipeline.

1.2.4 Results and interpretation

And then, when one has a catalog of unbiased ellipticity measurements and an estimator that gives an unbiased (but still noisy) value for g, begins our interpretation of the reality we have partly inferred from the shadows on the wall.

To relate the lensing signal to the matter distribution, we rely on accurate estimates of e.g. the distances D involved (see Fig. 1.4) or the center of the structure of interest. But let's assume here for a moment, to regain some lightweightedness in this introduction, that such other necessities are of little consequence, i.e. the uncertainties and biases caused are of subdominant[16] effect: in our interpretation, we will then inevitably compare our results to an expectation, a preconceived model of the universe, in order to understand it. So, what are we actually looking at?

For instance, we have considered a single lensing event by a group or cluster of galaxies, where in practice, light rays travel to a constantly changing potential, due to the large scale structure of the universe. This leaves a statistical imprint, called cosmic shear, first detected by (Bacon et al. 2000, Van Waerbeke et al. 2000, Wittman et al. 2000). This field of gravitational lensing in itself provides a complement to cosmological parameter determinations derived from observations of e.g. the CMB (see e.g. Kilbinger 2015, for a review). For our study of lensing by groups or clusters of galaxies, this cosmic shear signal is an additional form of noise, i.e. we detect the signal from the cluster and the cosmic shear together (Hoekstra et al. 2011). In some cases, such as the supergroup observed in Smit et al. (2015, chapter

[15] By infinite, we mean that we can apply any $g < 1$ to any $\epsilon^I < 1$ an infinite, discrete amount of times, and still retain $\epsilon < 1$.

[16] But they are most certainly not, as will be discussed in chapters 2, 3, and 5.

2), there are even known matter overdensities in the background, that need to be taken into consideration.

In fact, background galaxy shapes can also be coherently affected by a mechanism known as intrinsic alignment (IA) that is not due to lensing at all. This is because background galaxies that reside in the same potential, have their intrinsic orientations influenced in the same way, for instance due to tidal forces or angular momentum (e.g. Crittenden et al. 2002).

This effect is then further confused by the lensing effect of such a potential on even more distant background galaxies, introducing a (negative) correlation between the intrinsic shapes of galaxies and the shear experienced by galaxies (e.g. Hirata & Seljak 2004).

And, coming full circle to the start of this introduction, we study the shapes of the components of background galaxies that are *visible* to us, even though the vast majority of matter is dark and could be misaligned with the visible shapes. For our work, these effects are assumed to average out, only slightly increasing uncertainties, but we mention them here in reflection on the bigger picture.

This thesis is therefore not a road map to climb out of the cave. It does intend to provide a possible foothold, when contemplating that long climb.

1.3 Thesis outline

This thesis covers four studies into weak gravitational lensing, consisting of a theoretical consideration of the statistics of weak shear inference, and three separate studies of weak lensing by galaxy groups or clusters, exploring various aspects that obscure direct observations of gravitational shear.

1.3.1 Mass distribution in an assembling super galaxy group at $z = 0.37$

Chapter 2 is based on Smit et al. (2015) and presents a weak gravitational lensing analysis of supergroup SG1120−1202 (Gonzalez et al. 2005), consisting of four distinct X-ray-luminous groups that will merge to form a cluster comparable in mass to Coma at $z = 0$. This supergroup was discovered in the Las Campanas Distant Cluster Survey (Gonzalez et al. 2001) and has been studied in a series of papers (Gonzalez et al. 2005, Tran et al. 2008, 2009, Kautsch et al. 2008, Freeland et al. 2011, Just et al. 2011, Smit et al. 2015, Monroe et al. 2017).

The member groups lie within a projected separation of 1 to 4 Mpc and within $\Delta v = 550$ km s^{-1}, which is comparable to distances in the bullet cluster system (Tucker et al. 1995, Markevitch et al. 2002, Clowe et al. 2004), and as such, form a unique protocluster to study the matter distribution in a coalescing system.

We studied the weak gravitational distortion of background galaxy images by the matter distribution in the supergroup with high-resolution HST/ACS imaging. To robustly assess the systematic image distortions and artifacts present in these images, we used the shape measurement methodology for HST/ACS imaging outlined in (Schrabback et al. 2010), based on KSB+ (Erben et al. 2001).

We compared the reconstructed projected density field with the distribution of galaxies and hot X-ray emitting gas in the system and show that the projected mass distribution closely follows the locations of the X-ray peaks and associated brightest group galaxies. Since the

groups show no visible signs of interaction, our findings support the hypothesis that we observe the groups before they merge into a cluster.

We derived halo parameters for the individual density peaks, finding velocity dispersions between 355^{+55}_{-70} and 530^{+45}_{-55} km s^{-1} and masses between $0.8^{+0.4}_{-0.3} \times 10^{14}$ and $1.6^{+0.5}_{-0.4} \times 10^{14}$ h^{-1} $M_{\odot}$, consistent with independent measurements.

1.3.2 Weak lensing by very low redshift groups: analysis of systematics and robust shape measurements

In **chapter 3**, we studied the weak gravitational lensing signal from 79 low redshift groups ($0.05 < z < 0.0585$) from the 2dF Percolation-Inferred Galaxy Group catalog (2PIGG, Eke et al. 2004), based on the Two-degree-Field Galaxy Redshift Survey (2dFGRS, Colless et al. 2001), and observed with the Wide Field Imager (WFI) as part of the Zürich Environmental Survey (ZENS, Carollo et al. 2013). These groups cover two orders or magnitude in mass, with velocity dispersion $38 \lesssim \sigma \lesssim 691$ km s^{-1}, inferred mass $10^{12} \lesssim M \lesssim 10^{14}$ h^{-1} $M_{\odot}$, and b_J luminosity $1.1 \times 10^{10} \lesssim L \lesssim 5.9 \times 10^{11}$ h^{-2} $L_{\odot}$. The aims of this work were twofold.

Firstly, these groups covered a mass range that had not been studied extensively yet at the time of this research, filling a niche in the mass spectrum. Most weak lensing studies had been into large scales, up to those of clusters and superclusters, and into small scales, the lensing by individual galaxies. On intermediate scales, there were results in the high mass group regime (see e.g. Mandelbaum et al. 2006, Leauthaud et al. 2010) and Hoekstra et al. (2001) reported the first measurements of light galaxy groups in the CNOC2.

Secondly, this research could be a suitable pathfinder for the Kilo-Degree Survey (de Jong et al. 2013), which has an extensive overlap with the 2dFGRS and this sample. Given the low redshift and low median mass of these groups, which translate to a weak gravitational lensing signal, this work focused extensively on understanding possible systematics in wide-field imaging data, shape measurements methods and the statistical effect of outliers.

And important part of the analysis of systematics was estimating the effect of stacking of exposures on PSF correction and shape measurements, comparing the shapes of stars and sources detected in varying number of exposures and sources detected on multiple parts of the CCD array, due to being present in overlapping fields of view.

To asses the robustness of shape measurements, we compared two different pipelines, namely KSB+ (Erben et al. 2001), based on surface brightness moments, and shapelets (Kuijken 2006), considered a model fitting method. The results were statistically consistent.

We analyzed the effect of outliers in shape measurements of the inference of the weak shear signal, by comparing the commonly used weighted mean of a sample of ellipticities to the estimated obtained by convex hull peeling (CHP). The results were comparable in this respect as well.

Finally, we derived estimates for the velocity dispersion by fitting the lens profile for a singular isothermal sphere (SIS), and estimates for the mass and concentration of the groups, by fitting a Navarro, Frenk & White profile (NFW, Navarro et al. 1996), in good agreement with the dynamical estimates from 2PIGG.

1.3.3 Optimal statistics for weak shear analyses

Chapter 4 is based on Smit & Kuijken (2018), the first of a set of papers (with chapter

5 describing the second), and continues the analysis of optimal methods for inferring the gravitational shear from a sample of measured ellipticities of background galaxies.

This way of determining the shear signal is fundamentally limited by the intrinsic distribution of shapes that galaxies exhibit. It is well known that the distribution of galaxy ellipticities is non-Gaussian (e.g. Lambas et al. 1992, Rodríguez & Padilla 2013), and traditional estimation methods, explicitly or implicitly assuming Gaussianity, lead to noise biases (e.g. Kacprzak et al. 2012, Melchior & Viola 2012), possibly of the order of a few percent. This makes them comparable or even dominant to other sources of uncertainty in the process, such as biases in shape measurement, uncertainty in the lensing geometry introduced by photometric redshift probability distributions, or selection biases.

An optimal estimator is, from a principled point of view, more objective and better suited than corrections to an approach that is known to mismatch the sample distribution. In this work, we refine our method of CHP and complement that method and the canonical weighted mean (or weighted least squares or L^2 estimator) by the least absolute deviations (LAD or L^1) estimator and the biweight estimator (Beaton & Tukey 1974). We also allowed for a range of possible ellipticity distributions, comparing a Gaussian distribution with a flat axis ratio distribution, an ellipticity distribution from projected ellipsoids and the ellipticity distribution in the CFHTLenS shear catalog(Heymans et al. 2012b).

We analyzed the biases, relative efficiencies and robustness of these estimators. We conclude that the LAD estimator is the most robust when applied to our simulations, reducing noise bias by more than $\sim 30\%$, while increasing efficiency by a factor of 5 in the ideal case, and a factor of 1.2 when applied to CFHTLenS data.

We applied these methods to fitting of Fourier modes to the pattern of ellipticities in a simulated image, as a proof of concept.

1.3.4 AMICO galaxy clusters in KiDS-DR3: the impact of estimator statistics on the luminosity-mass scaling relation

In **chapter 5**, based on Smit et al. (2021), we apply the findings of the analysis described in chapter 4 on data from the third KiDS data release (KiDS-450, de Jong et al. 2017). We use the shape measurements of background sources around 6925 clusters found in KiDS-450 using the Adaptive Matched Identifier of Clustered Objects (AMICO, Bellagamba et al. 2011, Radovich et al. 2017, Bellagamba et al. 2018) and compare the results obtained with the weighted LAD and mean estimators.

The high signal-to-noise ratio of the shear signal allows us to study the scaling relation between the r-band cluster luminosity L_{200}, and the derived lensing mass M_{200}. We show the results of the scaling relations derived in 13 bins in L_{200}, with a tightly constrained power law slope of $\sim 1.24 \pm 0.08$.

We observe a small, but significant relative bias of a few percent in the results of the two regression methods, which is in excellent agreement with our findings in chapter 4. The efficiency of LAD is at least that of the weighted mean, relatively increasing with higher signal-to-noise shape measurements, a further confirmation of our previous results. As such, LAD regression provides a robust consistency check for shear inference, which has been and still remains a major investment in the field of weak lensing, while increased computation times remain feasible.

2

Mass distribution in an assembling super galaxy group at $z = 0.37$

We present a weak gravitational lensing analysis of supergroup SG1120−1202, consisting of four distinct X-ray-luminous groups that will merge to form a cluster comparable in mass to Coma at $z = 0$. These groups lie within a projected separation of 1 to 4 Mpc and within $\Delta v = 550$ km s^{-1} and form a unique protocluster to study the matter distribution in a coalescing system.

Using high-resolution *HST*/ACS imaging, combined with an extensive spectroscopic and imaging data set, we studied the weak gravitational distortion of background galaxy images by the matter distribution in the supergroup. We compared the reconstructed projected density field with the distribution of galaxies and hot X-ray emitting gas in the system and derived halo parameters for the individual density peaks.

We show that the projected mass distribution closely follows the locations of the X-ray peaks and associated brightest group galaxies. One of the groups that lies at slightly lower redshift ($z \approx 0.35$) than the other three groups ($z \approx 0.37$) is X-ray luminous, but is barely detected in the gravitational lensing signal. The other three groups show a significant detection (up to 5σ in mass), with velocity dispersions between 355^{+55}_{-70} and 530^{+45}_{-55} km s^{-1} and masses between $0.8^{+0.4}_{-0.3}\times10^{14}$ and $1.6^{+0.5}_{-0.4}\times10^{14}h^{-1}M_{\odot}$, consistent with independent measurements. These groups are associated with peaks in the galaxy and gas density in a relatively straightforward manner. Since the groups show no visible signs of interaction, this supports the hypothesis that we observe the groups before they merge into a cluster.

M. Smit, T. Schrabback, M. Velander, K. Kuijken, A. H. Gonzalez, J. Moustakas, and K.-V. H. Tran
Astronomy & Astrophysics, **Volume 582**, A82 (2015)

2.1 Introduction

In the framework of hierarchical structure formation (Peebles 1970), matter overdensities grow through merging and accretion from the scales of galaxies up to those of large-scale structure (LSS). In the concordance ΛCDM cosmology, the large-scale structure of the Universe is driven by the density fluctuations of dark matter, which provide the initial framework for subsequent structure formation. As such, the mass distribution in the Universe is the driving force behind the formation of clustering sites for astrophysical processes, such as galaxy groups and clusters.

In galaxy formation and evolution, environment plays a role of major importance. Most galaxies are found in groups and clusters (e.g., Eke et al. 2004), and observations indicate that the main part of galaxy evolution takes place in the group environment, with significant post-processing occurring in clusters (Tran et al. 2008, 2009, hereafter T08 and T09). A detailed understanding of the total mass (dark and visible) and the structure of the mass density distribution is therefore necessary to understand both the processes of group and cluster formation and fundamental scaling relations (e.g., Leauthaud et al. 2010, Hoekstra et al. 2012, Von der Linden et al. 2014) as well as to distinguish the latter from intrinsic variances in astrophysical processes.

Most overdensities are detected using visible, that is, baryonic means. Common methods use galaxies (red sequence and spectroscopic association, e.g., Eke et al. 2004, Gladders & Yee 2005) or gas (X-ray emission or the SZ effect, e.g., Sunyaev & Zeldovich 1970, 1972, Finoguenov et al. 2007). While these methods are efficient, they might not always be as effective: they rely on the presence of baryonic matter, while the matter distribution is driven by dark matter. Furthermore, the subsequent classification relies on observing the results of complex (astrophysical) processes, which introduces a significant intrinsic scatter in properties such as X-ray temperatures, star formation rate, and galaxy morphologies in different structures of comparable mass.

Gravitational lensing is the only direct probe of the total mass distribution, in the sense that it does not rely on astrophysical assumptions. A lower signal-to-noise ratio (S/N) makes it a less efficient detection method except for massive structures, but in combination with complementary methods, it is a powerful independent tool. From a statistical perspective, as an independent, direct measurement, it can serve as a calibrator for mass-observable scaling relations (e.g., Leauthaud et al. 2010, Hoekstra et al. 2012, Von der Linden et al. 2014). In individual systems it provides an independent estimation of the (projected) density field and can shed light on aspects such as interaction, dynamical state, and offsets between the overall matter distribution and the baryonic matter.

Direct reconstructions of the density fields of individual systems do not have a high resolution and are predicted to show significant noise fluctuations (Van Waerbeke 2000). However, it gives an important qualitative indication of the dominant density distribution, independent of the presence of baryons. To determine the parameters of the matter distribution, either in a statistical survey or for individual systems, robust centroiding is an important task. Using other tracers of the center of mass, such as X-ray detections, brightest cluster/group galaxies (BCG/BGG), or the luminosity-weighted mean (LWM) position, gravitational lensing can significantly constrain halo masses and concentrations (e.g., George et al. 2012).

Galaxy clusters and, in the past decade, galaxy groups, are now identified in a robust manner, including examples of accretion of smaller structures onto existing clusters. However,

we have less observational evidence of the connection between structures on various scales, that is to say, the initial assembly of clusters from groups and galaxies. In this study, we perform a weak lensing analysis of SG1120−1202 (Gonzalez et al. 2005, hereafter G05), an assembling system of four galaxy groups at $z \sim 0.37$ discovered in the Las Campanas Distant Cluster Survey (LCDCS, Gonzalez et al. 2001). These groups are gravitationally bound and will merge into a galaxy cluster comparable in mass to Coma by $z = 0$. The supergroup, hereafter SG1120, is confirmed by X-ray imaging and optical spectroscopy and has already formed a red sequence (see, e.g., G05, T08, T09, and Kautsch et al. 2008, Just et al. 2011, Freeland et al. 2011).

The individual subgroups are in the low-mass regime of X-ray groups, $M_{200} \sim 10^{13}$ to $10^{14} M_\odot$ and $\sigma_v \sim 400$ km s^{-1}, and have not yet interacted. The aim of this study is to determine the total matter distribution in the system (dark and baryonic) and to constrain individual halo masses.

This paper is organized as follows. We summarize the general framework for weak lensing in Sect. 2. In Sect. 3 we briefly describe the data we use, while Sect. 4 covers the framework of measurement and analysis methods. In Sect. 5 we discuss the results and the scientific implications. Section 6 gives a summary of our conclusions.

Throughout this paper we assume a Planck (Planck Collaboration et al. 2014) cosmology with $\Omega_M = 0.3183$, $\Omega_\Lambda = 0.6817$ and $H_0 = 67.04$ km s^{-1} Mpc^{-1}.

2.2 Weak lensing framework

Gravitational lensing is the effect of curved space-time on the paths of light rays from distant sources to the observer as they pass through the potential of foreground structures. This geometrical effect leads to a displacement of point sources on the projected plane of the sky. The differential effect on extended sources leads to magnification and distortion effects. This is commonly described as a coordinate transformation

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \tag{2.1}$$

where the trace component κ is known as the convergence and the reduced symmetric part is determined by the gravitational shear (γ_1, γ_2).

Since we do not know the intrinsic source sizes or magnitudes, we can only measure the net distortion or reduced shear $(g_1, g_2) \equiv (\gamma_1, \gamma_2)/(1-\kappa)$:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = (1-\kappa) \begin{pmatrix} 1-g_1 & -g_2 \\ -g_2 & 1+g_1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \tag{2.2}$$

where the transformation is written as a multiplication of $(1-\kappa)$, which we do not identify, and a distortion matrix describing the alignment of lensed sources in the foreground potential.

The observed shape of a background source is not a pure tracer of the gravitational lensing effect, but the combined effect of an intrinsic galaxy shape and any distortion of that shape, including gravitational lensing. Systematic effects such as telescope aberrations and detector systematics likewise contribute, which need to be corrected for.

The uncertainties the intrinsic shapes introduce is referred to as shape noise, and the amount of background sources available determines the precision of the results. The shape noise averages out statistically if we assume that the background sources are randomly oriented intrinsically. If faint members of the foreground structure are not identified and removed from the sample of background sources, these shapes will not be affected by gravitational lensing, but might be aligned with the potential of the structure under investigation. These effects are known as intrinsic alignment (see, e.g., Mandelbaum et al. 2006). However, recent results suggest that intrinsic alignments should have negligible influence for current cluster weak lensing studies (Sifón et al. 2015b).

The average measured distortion, corrected for systematic effects, can then be related to the projected density distribution in the lensing structure through

$$\kappa(\boldsymbol{\theta}) = \frac{4\pi G}{c^2}\Sigma(\boldsymbol{\theta})\frac{D_{\rm ol}D_{\rm ls}}{D_{\rm os}}\,, \tag{2.3}$$

where $\boldsymbol{\theta}$ represents the angular coordinates on the plane of the sky, $\Sigma(\boldsymbol{\theta})$ is the projected density distribution, and $D_{\rm i}$ are the angular diameter distances between the observer, lens, and background sources (luminosity distances, sometimes written as $D_{\rm l}$, are not used throughout this paper).

Normalized by $4\pi G c^{-2}$, the convergence κ is therefore also known as the dimensionless surface mass density, directly related to the lensing density distribution and the lensing geometry. For axisymmetric lenses, $|\gamma|(\theta) = \bar{\kappa}(<\theta) - \kappa(\theta)$ with $\bar{\kappa}(<\theta)$ the mean surface mass density within a radial separation $\theta = |\boldsymbol{\theta}|$ to the lens centroid.

2.3 Data

For this project we made use of results of an extensive multiwavelength data set (see, e.g., G05, T08, and T09).

Key to our lensing analysis are optical data, consisting of high-resolution *HST*/ACS[1] imaging used for shape measurements, as well as *VLT*/VIMOS (Le Fèvre et al. 2003), *VLT*/FORS2 (Appenzeller et al. 1998), and *Magellan*/LDSS3 spectroscopy.

We also used the X-ray temperatures based on *Chandra*/ACIS imaging and stellar masses inferred from *VLT*/VIMOS *BVR* photometry (T08) and complement our optical color information with *KPNO*/FLAMINGOS near-infrared (NIR) K_s imaging.

We use $\alpha = 11^{\rm h}19^{\rm m}58\overset{\rm s}{.}0$, $\delta = 12^\circ03'33\overset{''}{.}0$ as center of coordinates, which roughly is the center of the VIMOS imaging data. To convert angular to physical separations, we use a reference redshift of $z = 0.37$; this is the median of the redshifts of the four BGGs.

2.3.1 HST imaging

The *HST*/ACS imaging data were taken in July 2005 and January 2006 and consist of ten pointings, forming a contiguous $11' \times 18'$ mosaic. Each tile was observed in F814W ($0\overset{''}{.}05$/pixel) for 2ks over four dithered exposures. Figure 2.1 shows the layout of the different pointings.

[1] Based on observations made with the NASA/ESA Hubble Space Telescope, obtained at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc.

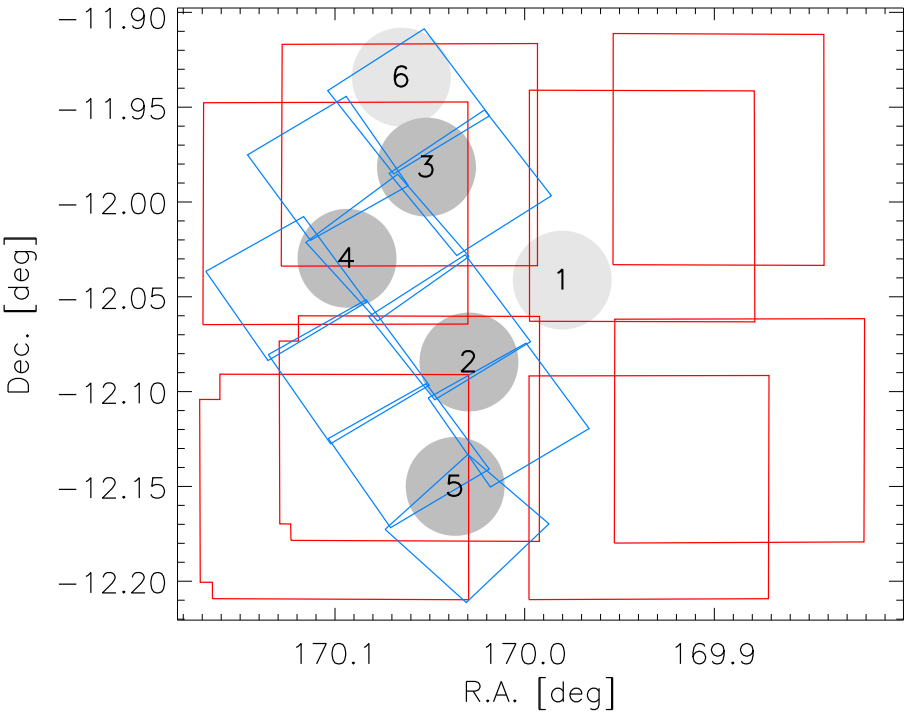


Figure 2.1: Layout of the *VLT*/VIMOS pointings (red) and *HST*/ACS pointings (blue). The detected X-ray peaks are shown as well (gray), with the radius of the circles 0.5 h^{-1} Mpc. The X-ray peaks 1 and 6 (light gray) are associated with structures at higher redshift, beyond SG1120 (G05).

We reduced the data with the same pipeline as employed in Schrabback et al. (2010), which uses `MultiDrizzle` (Koekemoer et al. 2003) to stack exposures and remove cosmic rays. It also includes careful refinement of shifts and rotations between exposures as well as optimized weighting.

Schrabback et al. (2010) found that using `MultiDrizzle` with the default cosmic-ray rejection parameters can cause central stellar pixels to be flagged as cosmic rays, especially when there are significant PSF variations between exposures. Galaxies are not affected, due to their shallower light profiles. To avoid differences in the effective stacked PSFs, we created separate stacks for stars and galaxies, with less aggressive cosmic-ray rejection for the former.

For a more detailed description of the reduction process, we refer to Schrabback et al. (2010).

2.3.2 Spectroscopy

We employed optical spectroscopy consisting of three subsets of data. The first subset of targets was selected from a magnitude-limited catalog ($R \leq 22.5$), with preference given to objects in visually overdense regions (G05), and observed using VIMOS. Follow-up spectroscopy was selected from K_s catalogs ($K_s \leq 20$) and carried out on LDSS3 and FORS2. Figure 2.2 shows the redshift distribution of the final target selection, using a redshift quality

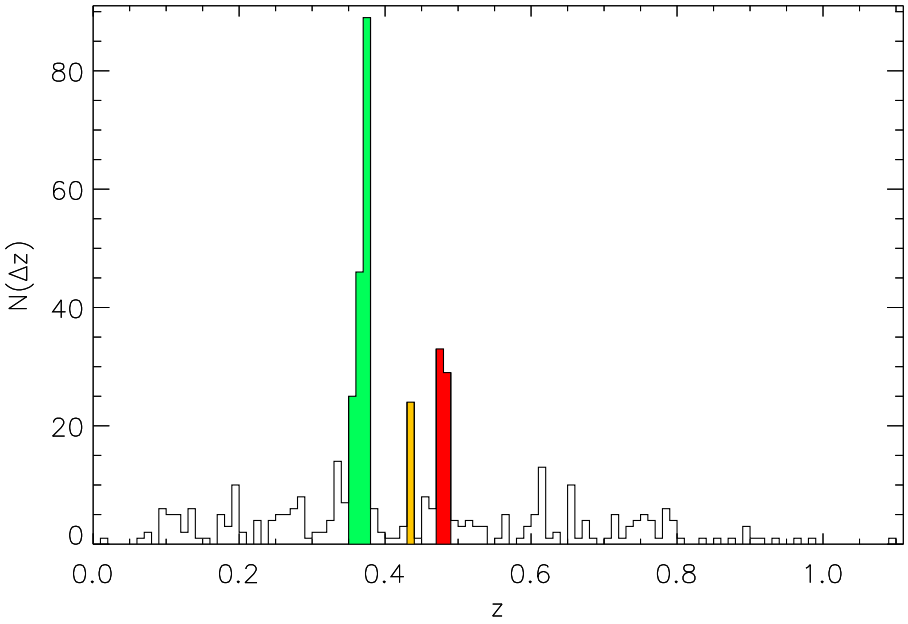


Figure 2.2: Redshift distribution of the spectroscopic targets around SG1120. Three significant peaks are identified between $0.35 < z < 0.37$, associated with SG1120; between $0.43 < z < 0.44$, unassociated but concentrated slightly north of peak 1; and between $0.47 < z < 0.49$, associated with X-ray peaks 1 and 6.

cut as defined in T08.

Members of the subgroups of SG1120 were initially selected as galaxies within 500 kpc of their respective X-ray peaks within the redshift range $0.32 \leq z \leq 0.39$. We narrowed the redshift range to $0.34 \leq z \leq 0.38$ without the loss of any members. Figure 2.3 shows a layout of the targets.

2.3.3 Subgroups

In Table 2.1, we give an overview of the properties of each subgroup, also given in G05, T08 and T09. We use the same numbering convention for the X-ray peaks and BGGs as these papers. Subtle differences in the number of group members are due to using the same selection criteria as T09, but a slightly different cosmology.

Figure 2.4 shows the radial (z, y) distribution of galaxies within 500 kpc of each subgroup, with $0.315 \leq z \leq 0.415$. The subgroup associated with X-ray peak 2 and BGG 1 is at a slightly lower redshift, with an estimated $30 \sim 40$ Mpc in angular diameter distance to the median redshift of the supergroup.

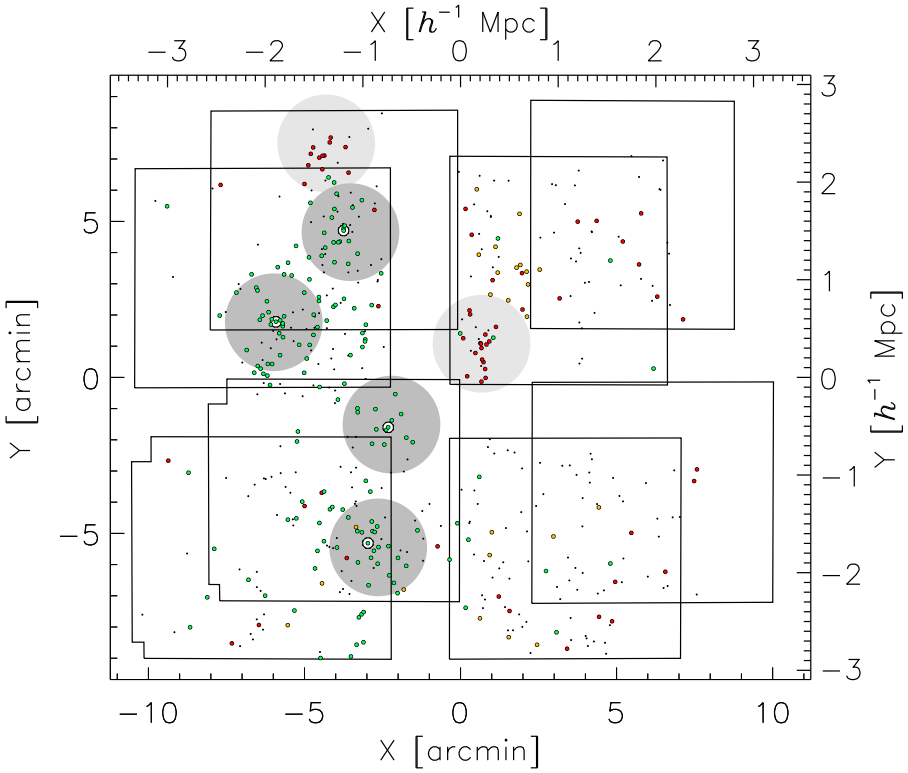


Figure 2.3: Layout of spectroscopic targets, overlaid with the *VLT*/VIMOS pointings. $(x, y) = (0, 0)$ corresponds to $\alpha = 11^{\rm h}19^{\rm m}58\overset{\rm s}{.}0$, $\delta = 12°03'33\overset{''}{.}0$. Colors correspond to the peaks in Fig. 2.2. The BGGs are indicated by larger circles.

2.4 Analysis

In this section we briefly describe our method of shape measurement. We discuss the redshift distribution and selection of background sources. After establishing a reliable background catalog with robust shapes, we describe how we obtain a qualitative reconstruction of the projected mass density and complement this with density profiles to the subgroups based on the BGG and X-ray peak positions.

2.4.1 KSB+ shape measurements

The art of measuring accurate galaxy shapes is an ongoing field of investigation, as witnessed, for instance, by the Shear TEsting Programmes and the GRavitational lEnsing Accuracy Testing (Heymans et al. 2006, Massey et al. 2007, Bridle et al. 2010, Kitching et al. 2012, Mandelbaum et al. 2015, hereafter STEP, STEP2, and GREAT08, GREAT10, and GREAT3). We make use of the KSB method (Kaiser et al. 1995), the most commonly used and tested technique in the past decade, and discuss its application to ACS data.

For this study we used the same approach as Schrabback et al. (2007, 2010, the TS

Table 2.1: Properties of galaxy groups in SG1120 and the two structures identified at higher redshift.

BGG ID	X-ray peak	z	T (keV)	σ_z (km s^{-1})	N
1	2	0.3522	$2.2^{+0.7}_{-0.4}$	303 ± 60	13
2	3	0.3707	$1.7^{+0.5}_{-0.3}$	406 ± 83	19
3	4	0.3712	$1.8^{+1.2}_{-0.5}$	580 ± 100	29
4	5	0.3694	$3.0^{+1.2}_{-1.0}$	567 ± 119	21
	1	0.4794	$2.3^{+0.4}_{-0.3}$	820 ± 101	19
	6	0.4801	...	...	12

pipeline in STEP and STEP2) based on the implementation by Erben et al. (2001). KSB uses the first-order effects of distortions induced by gravitational shear and PSF on the weighted second moments of the light distribution of a source to estimate the reduced shear. We detect objects in the same way using `SExtractor` (Bertin & Arnouts 1996).

We defined individual weights for this method based on the variance of the shear estimators from this pipeline as

$$w \propto \left(\sigma^2_{sn} + \sigma^2_{e_1} + \sigma^2_{e_2}\right)^{-1} , \tag{2.4}$$

where we assumed a minimum variance of σ^2_{sn}, the intrinsic shape noise. Based on the findings in the STEP analyses, we expect an underestimation of the shears by KSB+ of about a few percent. We applied the same empirical correction factor as Schrabback et al. (2010) to account for this expected bias.

The systematic distortion effects due to telescope and optic system give rise to shapes convoluted by a point spread function (PSF). The main source of variations of the ACS PSF is caused by changes in the telescope focus, causing spatial and temporal fluctuations (see, e.g., Schrabback et al. 2007, Rhodes et al. 2007).

A common strategy is to map the distortions caused by the PSF using the shapes of foreground stars, but the average number of stars in our ACS images is $\sim 20 - 40$. This leads to a poorly sampled PSF and an imperfect correction, causing significant residual distortions, especially detectable at the edges of the images where the tiles overlap slightly. We therefore adopted the same strategy as Schrabback et al. (2010) based on a principal component analysis of the PSF variation in dense stellar fields.

Furthermore, deterioration of the ACS CCDs over time due to constant exposure to cosmic rays in space leads to an effect called charge-transfer inefficiency (CTI), causing charge trails in the CCD readout direction (e.g., Rhodes et al. 2007, Massey et al. 2010, Schrabback et al. 2010). These effects will affect the measured shear pattern and the reconstruction of the projected density distribution, and it is therefore important to correct for them.

Here we applied the same parametric CTI model as described in Schrabback et al. (2010) for the correction of the KSB+ polarizations.

After constructing the shape measurement catalogs, we applied several common selection

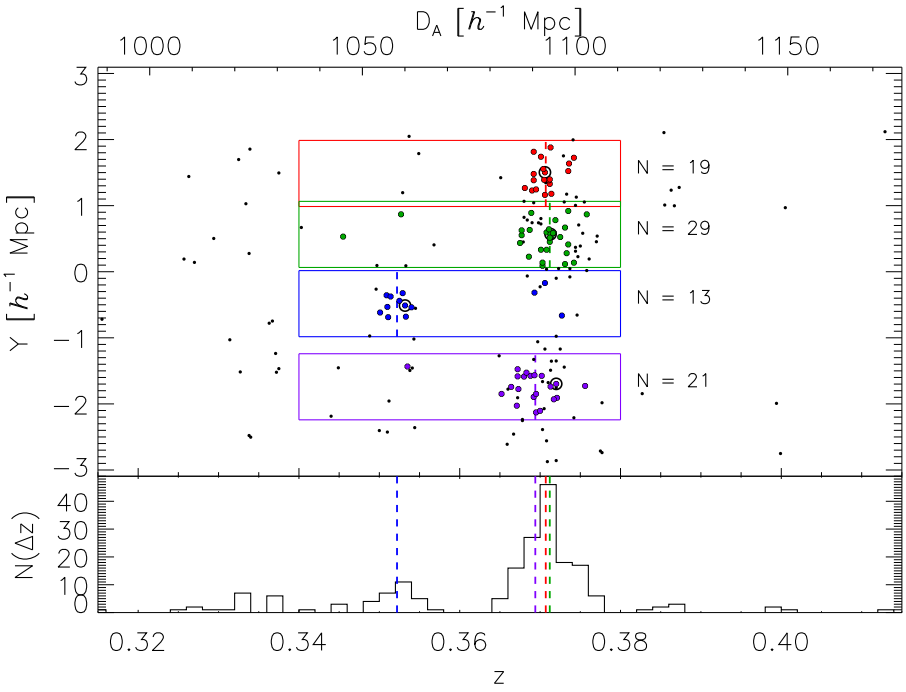


Figure 2.4: Radial (z, y) distribution of all objects (black dots, with a redshift quality of 3 as defined in T08). Objects within 500 kpc of an X-ray peak and the corresponding selection criteria are shown in different colors for easy distinction. The BGGs are indicated by larger circles, group redshifts by dashed lines.

criteria and cuts. These criteria are based on simulations and quality flags of the detection and shape measurement pipelines, and they depend on the noise properties, on the variance and convergence of the model fits, and on the object and PSF size.

A list of the various selection criteria can be found in the appendix. Sources that pass the criteria of all pipelines number 7012, for a source density of ~ 64 galaxies/arcmin^2 with `MAG_AUTO` magnitudes $i_{814} < 27.1$.

2.4.2 Redshift distribution

We acquired spectroscopic redshifts for 497 objects in our catalogs. The spectroscopic targets were selected based on magnitude, and preference was given to visually overdense regions, which means that these spectroscopically confirmed members do not give a complete picture of the galaxy distribution in SG1120. The brightest confirmed supergroup member has $i_{814} = 17.5$, while the spectroscopic survey remains $> 50\%$ complete to $i_{814} = 20.5$ (T09). We find that confirmed supergroup members have numbers peaking between magnitudes $19.5 < i_{814} < 20.0$.

To separate background and foreground sources, we considered that group members are expected to dominate number counts in the magnitude range of confirmed members, while

background number counts dominate at fainter magnitudes. We initially selected background sources as objects with $i_{814} \geq 22$ and assessed possible contamination by faint foreground objects. Figure 2.5 shows the number density of sources with $i_{814} < 22$ and $i_{814} > 22$, where we have used a Gaussian smoothing with a 20" kernel width.

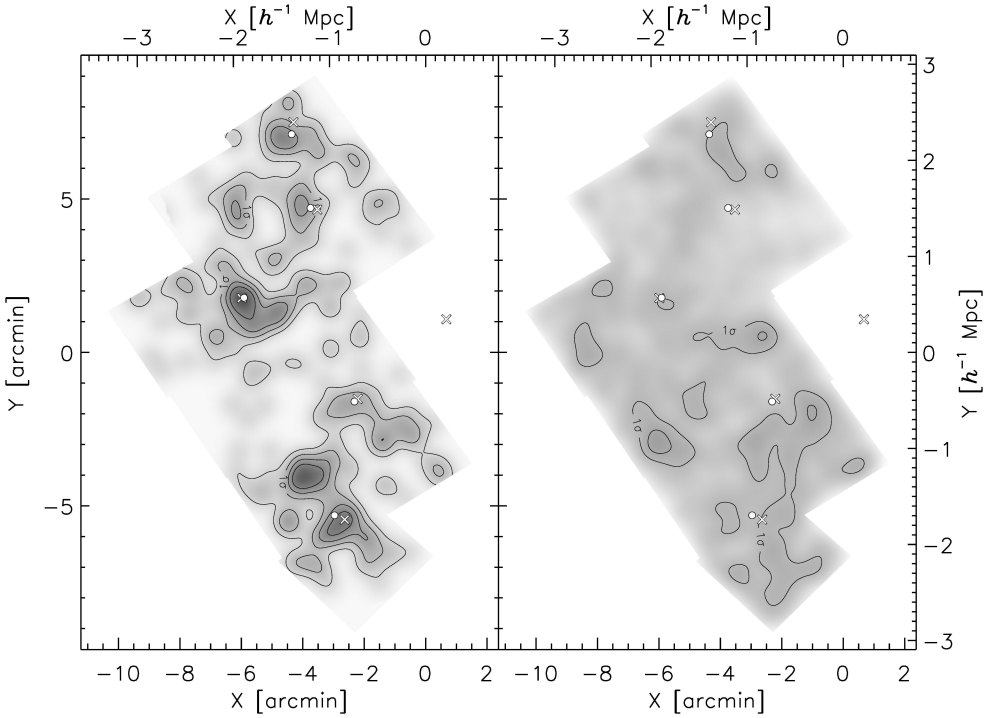


Figure 2.5: Number density of sources with $i_{814} < 22$ *(left)* and $i_{814} > 22$ *(right)*, smoothed using a Gaussian smoothing kernel with a width of 20". Dense regions are shown as dark in a normalized grayscale. Contours correspond to fluctuations in integer standard deviations in number density.

Because gravitational lensing is a geometric effect that has a non-linear dependence on redshift, we took the expected redshift distributions into account, following the same parametrization as in Schrabback et al. (2010). We show the total redshift distribution of sources with $i_{814} > 22$ in Fig. 2.6. For a given lens redshift, such as in this particular system, the lensing signal has a linear dependence on the lensing efficiency $\beta = \max\{0, D_{\rm ls}/D_{\rm os}\}$. We can therefore determine a mean lensing efficiency $\langle\beta\rangle$ for the sources with respect to each subgroup redshift.

As mentioned earlier, both X-ray peak 1 and 6 (G05) are associated with structures at higher redshift (both $0.46 \lesssim z \lesssim 0.48$). We must take the gravitational distortions caused by these background structures into account when trying to isolate the signal from SG1120. We therefore also determined a mean lensing efficiency for these structures.

We found average lensing efficiencies of $\langle\beta\rangle \approx 0.52$ for SG1120, corresponding to an effective background redshift of $z_{\rm eff} \approx 0.88$, and $\langle\beta\rangle \approx 0.42$ for the two background structures, corresponding to an effective background redshift of $z_{\rm eff} \approx 0.95$.

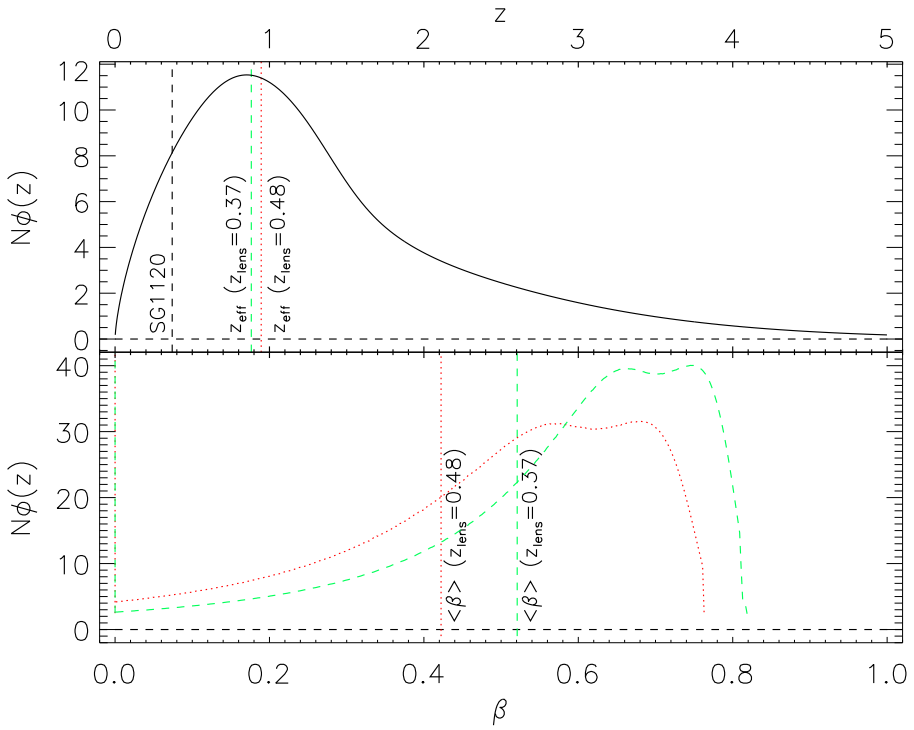


Figure 2.6: Parametric redshift distribution of sources with $i_{814} > 22$ *(upper panel)* and the corresponding distribution in lensing efficiency β *(lower panel)*. In the upper panel, the dashed green line corresponds to $z_{\rm eff} \approx 0.88$ with respect to SG1120 and the dotted red line corresponds to $z_{\rm eff} \approx 0.95$ with respect to the two structures at higher redshift. In the lower panel, the dashed green curve shows the distribution in β with respect to SG1120, with $\langle\beta\rangle \approx 0.52$ (dashed green vertical line), and the dotted red curve the distribution in β for the two structures at $z = 0.48$, with $\langle\beta\rangle \approx 0.42$ (dotted red vertical line).

Foreground contamination

An intrinsic redshift distribution of sources with $i_{814} > 22$ implies that some of these objects are faint foreground sources or members of the SG1120 structure. Based on our parametric redshift distribution, we estimate that $\sim$ 9% of our background sources to lie in front of SG1120.

Foreground sources are not lensed by the groups. We account for this dilution effect by assigning $\beta = 0$ to this part of the redshift distribution in our definition of the lensing efficiency above.

This assumes a random field of view, which is not the case for our observations, with known overdensities at $z \sim 0.37$ and $z \sim 0.48$. However, Fig. 2.5 suggests no significant correlation between the distribution of these sources and the galaxy distribution of SG1120. To estimate possible variations in the number density n of sources with $i_{814} > 22$, we derived an average number density profile around the group centers, as shown in Fig. 2.7.

We used radial bins between 10" $< \theta <$ 95" to avoid the BGGs and the edges of the

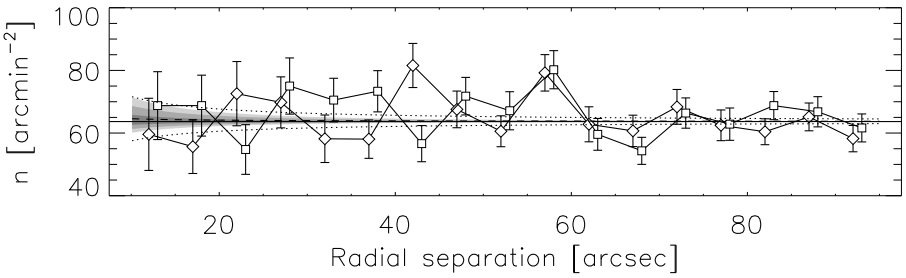


Figure 2.7: Variations in galaxy number density n as a function of radial distance from the lens positions, using the X-ray peaks (diamonds) and BGG positions (squares). Data points are slightly offset for clarity. Overplotted is the average number density of ~ 64 galaxies/arcmin^2 of the whole ACS mosaic (black solid line) and the best-fit radial profile (dashed) with 1σ errors (dotted). The estimated effect of the lensing magnification, $\mu^{\alpha-1}$, is shown in grayscale, varying the slope of the luminosity function between $0 < \alpha < 3$. Different shades in grayscale correspond to steps of 0.5×10^{14} in group mass M_{200}.

ACS coverage. We also considered only the group centers of groups 2 through 5, as group 6 is not entirely covered by the ACS pointings. Finally, we averaged the signal over all four subgroups to increase the signal-to-noise ratio.

We then quantified the radial dependence of the galaxy number density by fitting a parameterized profile given by $n(\theta) = (1 + a/\theta)n_{\rm bg}$, with θ in arcseconds and with $n_{\rm bg} = 64/\mathrm{arcmin}^2$ fixed. (In fact, if we allow $n_{\rm bg}$ to vary, we recover $n_{\rm bg} = (64 \pm 3)/\mathrm{arcmin}^2$.) We found $a = 0.14 \pm 1.11$, consistent with no trend in galaxy number density with radial separation from the lens centers.

To interpret this radial number density profile, we have to consider both the presence of unidentified faint group members and the effect of the lensing magnification μ (see, e.g., Bartelmann & Schneider 2001, and references therein). The presence of unidentified group members would increase the number density. Magnification increases both the observed flux of background sources, leading to an increase in n, and the solid angle behind the lenses, causing a dilution of n (not to be confused with the dilution of the shear signal caused by unlensed foreground objects in the background source sample). It then depends on the slope α of the luminosity function whether the lensing magnification causes a net increase or decrease in number density by $\mu^{\alpha-1}$, where μ and α depend of the source redshift. Both effects were shown by Hildebrandt et al. (2009). A decrease could cancel the effect of unidentified group members.

We wish to obtain a rough estimate of the expected influence of magnification to check whether it is smaller than the statistical uncertainty. For this we ignored the redshift dependencies of μ and α and considered a wide range $0 < \alpha < 3$, which was simply chosen to assess all possible variations in the magnification without making assumptions about the luminosity function. We used a group mass of $M_{200} = 1.0 \times 10^{14}$, where M_{200} is defined as the total mass within a radius of r_{200} of a halo, where the mean density of the halo is $\bar{\rho}(< r_{200}) = 200\rho_{\rm crit}(z)$. For a robust estimate, we also considered variations in M_{200} of $\pm 0.5 \times 10^{14}$. The results are plotted in Fig. 2.7 in grayscale.

We estimate possible magnification effects to be smaller than the statistical uncertainties from potential residual group member contamination. We expect that any residual excess contamination by member galaxies of SG1120 in the source sample must be small and com-

parable to regular line-of-sight variations. In Sect. 5.2, we confirm that we do not need to apply a dilution correction for excess contamination from the supergroup itself.

2.4.3 Lensing analysis

Our approach to determining the matter distribution in SG1120 is twofold. First, we show that the distribution of light (galaxy number densities, BGGs, and X-ray peaks) is closely correlated with the underlying mass distribution. Second, we determine the density profile parameters for each subgroup, taking into account the effect of each subgroup and background structure simultaneously.

Reconstruction of the mass distribution

We used a Kaiser-Squires (KS, Kaiser & Squires 1993) inversion technique to reconstruct the surface mass density. We smoothed the data onto a rectangular grid, using a Gaussian smoothing kernel with a width of 20", equal to the smoothing used for the galaxy number densities in Fig. 2.5.

We investigated possible systematic errors in our data by changing the phase of the shear by $\frac{1}{2}\pi$, which corresponds to rotating the background galaxies by $\frac{1}{4}\pi$. The distortion caused by weak lensing does not introduce a curl in the shear field, and the resulting reconstructed map should display only noise in the absence of systematic errors.

Density profile parameters

Earlier studies indicate that the groups are infalling for the first time and have not yet interacted, although X-ray measurements show a possible onset of interaction (G05). We considered the groups as individual overdensities with spherically symmetric density distributions and derived halo parameters for each group, including the background structure around X-ray peak 6.

We considered two types of density profiles and two possible choices of group centroids. We considered the Navarro-Frenk-White (NFW, Navarro et al. 1996) density profile and compared this to the singular isothermal sphere (SIS) model.

The SIS profile is determined by a single free parameter, the halo velocity dispersion σ_γ, where the subscript γ is used to distinguish this parameter, derived from a two-dimensional model of the projected mass density, from other derivations of velocity dispersion, such as the one-dimensional σ_z derived from the redshift distribution. The advantages of this profile are its simplicity and the linear dependence of the lensing signal on the squared velocity dispersion. The tangential component of the shear with respect to the group center is given by

$$\gamma_\mathrm{t}(\theta) = \frac{2\pi}{c^2}\sigma_\gamma^2\frac{\beta}{\theta}\,, \tag{2.5}$$

where θ indicates the separation from the center in radians. This allows for a straightforward interpretation of any possible correlation between the fitted parameters of different subgroups.

The NFW profile is usually expressed in terms of its mass and concentration and depends on redshift. The halo mass M_{200} is given by

$$M_{200} \equiv 200\rho_\mathrm{crit}(z)\frac{4}{3}\pi r_{200}^3 = 100\frac{H(z)}{G}r_{200}^3. \tag{2.6}$$

The concentration c_{200} is defined as the relation between the characteristic shape of the density profile and r_{200}. The analytical formulas for the shear signal of an NFW profile can be found in Wright & Brainerd (2000) and Bartelmann (1996).

Because of the lower S/N, the centers of dark matter haloes should not be estimated directly from the lensing data when determining density profile parameters. Instead, one has to rely upon visible tracers such as peaks in the X-ray emission of hot gas or the brightest or heaviest galaxy (e.g., in terms of a stellar mass as derived in T08) in the group or cluster. If the fitted halo model is offset from the true underlying halo, the fit is inferior and the introduced systematic uncertainties can be significant (George et al. 2012). In particular, the halo mass will on average be underestimated, while the uncertainties, most often determined from confidence levels, will be increased. This leads to both a biased and a less effective study.

As described in George et al. (2012), there are several choices possible as tracer of the halo center. These can be based upon a central galaxy, several or all of the associated galaxies, or the X-ray flux. In this study, the haloes under consideration are part of a coalescing system, and an offset from the true halo center of some or all of these tracers is not unlikely. However, the BGGs of the subgroup are also the most massive group galaxies (MMGG, George et al. 2012) in terms of stellar mass and magnitudes in most observed bands and coincide well with the X-ray peaks (T08). We derived the parameter values using both options and determined whether these are consistent.

Given the close angular separation of the X-ray peaks, we did not compute azimuthally averaged profiles. Instead, we computed the total lensing distortion $\boldsymbol{g} = \sum \boldsymbol{g}_i$ for each background source induced by each of the six foreground structures. This is valid if we assume $g \ll 1$, which is certainly the case for the sources where the distortion is not dominated by one of the lensing structures.

We then determined profile parameters for each subgroup using a χ^2 minimization. For X-ray peak 1, we assumed $\sigma = 820$ km s^{-1} from G05 and an order of magnitude $M_{200} = 3.7{\times}10^{14} h^{-1} M_\odot$ and assessed the effect of omitting the influence of this background structure.

2.5 Results

In this section we discuss the reconstructed density distribution and best-fit profile parameters, and we show that SG1120 is consistent with expectations from hierarchical structure formation, even though the system is not relaxed.

2.5.1 Matter distribution

In Fig. 2.8 we show the reconstruction of the projected surface mass density. We detect significant peaks near three of the foreground structures. We do not detect a significant peak in the density distribution near X-ray peak 2.

We considered the results of our mass reconstruction in a qualitative manner. The peaks in our surface mass density reconstruction coincide very well with the peaks in galaxy number density (Fig. 2.5) and X-ray emission (G05), within smoothing scales. We found no significant 'dark' overdensities either, and small offsets between peaks using various tracers

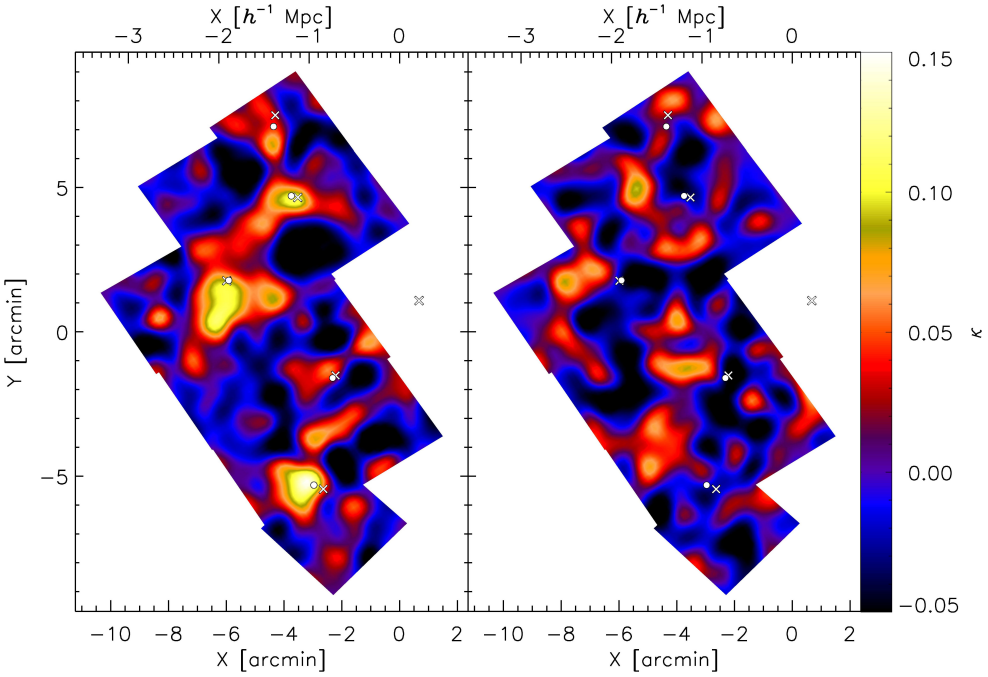


Figure 2.8: Smoothed map of the reconstructed projected density distribution (left) and the imaginary control signal (right) where the shear signal is rotated out of phase. X-ray peak positions are indicated by white crosses and BGG positions by white circles.

are expected in a coalescing system. Finally, the map shows significantly stronger peaks than the control map.

2.5.2 Individual groups

SIS velocity dispersions

We present the results of the joint χ^2 minimization fit of SIS profile parameters around the X-ray peaks in Fig. 2.9. The reduced χ^2 value is $\chi^2_\nu = 1.4$.

The combined contours of Fig. 2.9 show no features that indicate significant degeneracies between the individual group σ_γ values. While it is to be expected that nearby mass concentrations influence the shear pattern around an individual lens, we conclude that noise is a dominant factor in these results. More massive haloes or smaller halo separations can be expected to increase correlations.

The resulting σ_γ values are given in Table 2.2. Consistent with the reconstructed mass map in Fig. 2.8, we do not detect a very significant lensing signal around X-ray peak 2, barely exceeding the 68% confidence limit.

The velocity dispersion associated with X-ray peak 1 is necessarily kept constant, as the peak lies outside the ACS mosaic. Upon inspection, it turns out that varying this parameter between $0 \leq \sigma_1 \leq 820$ km s^{-1} does not alter the results by more than 10% of the 68%

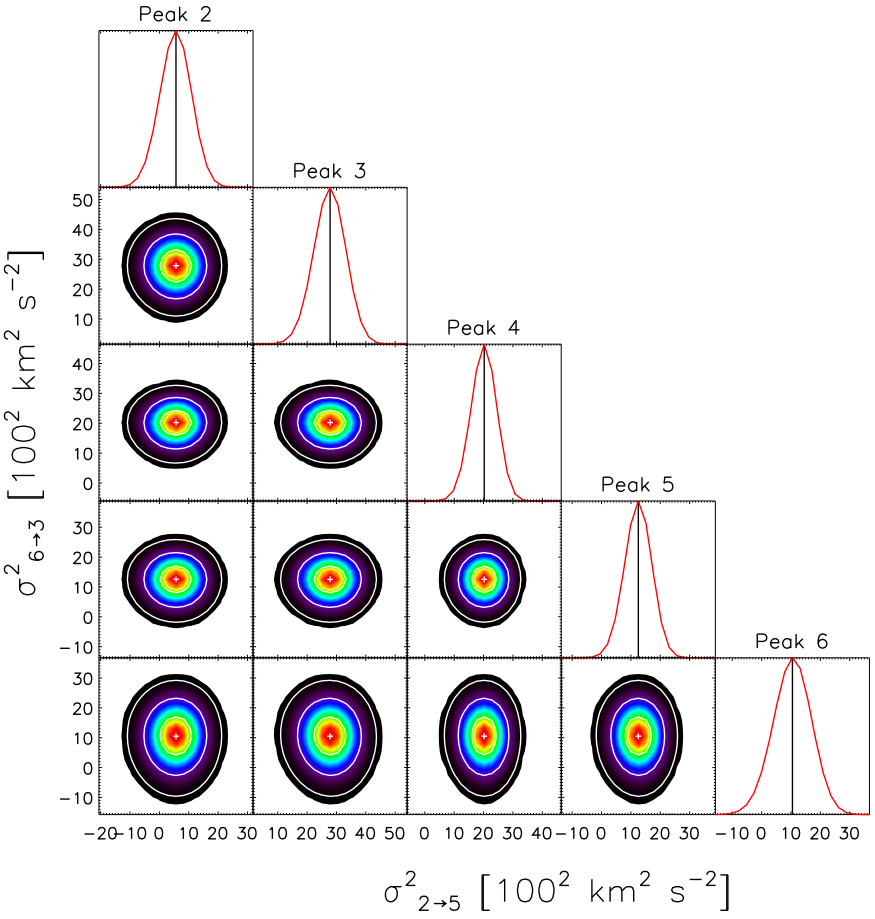


Figure 2.9: Marginalized 2D χ^2 distributions of the simultaneous fit to the individual subgroup velocity dispersions, together with the marginalized 1D likelihoods for each subgroup. Overplotted are the 68.3%,95.4%, and 99.7% confidence levels.

confidence interval for X-ray peak 2, which lies closest to peak 1. The effect is even smaller for the other groups.

Similar to our assessment of systematics for the mass map reconstruction, we repeated the fit to a control signal by changing the phase of the shear by $\frac{1}{2}\pi$. The results are consistent with a control signal of $g_c \approx 0$. Because of its less favorable lensing geometry ($\langle\beta\rangle = 0.42$), the constraints for group 6 are weaker, although it is still detected at a significance of $\sigma \approx 1.6$.

Finally, we determined how much our results would be affected if the signal were boosted by a dilution factor of $1 + (a + \sigma_a)/\theta = 1 + 1.25/\theta$ for group member contamination, as discussed in Sect. 4.2, using a conservative 1σ upper limit. We find that this does not alter the results by more than 37% of the 68% confidence intervals, justifying our earlier approach.

We repeated the fit around the BGGs as tracers of the halo centers. The results are very similar, with the fitted values also given in Table 2.2. There is some difference with up to 2σ

Table 2.2: Profile parameter fit results

Subgroup (X-ray ID)	σ_z (km s^{-1})	σ_γ (X-ray) (km s^{-1})	σ_γ (BGG) (km s^{-1})	M_{200} (X-ray) ($10^{14}h^{-1}M_\odot$)	M_{200} (BGG) ($10^{14}h^{-1}M_\odot$)
2	303 ± 60	240^{+90}_{-190}	$230^{+95}_{...}$	$0.3^{+0.1}_{-0.2}$	$0.2^{+0.2}_{-0.2}$
3	406 ± 83	530^{+45}_{-55}	425^{+60}_{-70}	$1.1^{+0.4}_{-0.4}$	$0.9^{+0.4}_{-0.3}$
4	580 ± 100	450^{+45}_{-50}	445^{+45}_{-50}	$1.6^{+0.5}_{-0.4}$	$1.6^{+0.5}_{-0.4}$
5	567 ± 119	355^{+55}_{-70}	480^{+45}_{-50}	$0.8^{+0.4}_{-0.3}$	$1.6^{+0.5}_{-0.4}$
6	...	325^{+85}_{-120}	305^{+90}_{-135}	$0.7^{+0.6}_{-0.5}$	$0.7^{+0.2}_{-0.1}$

deviations between the results for peaks 3 and 5, where the separation between X-ray peak and BGG is also the largest. The quality of the fit, in terms of a reduced χ^2 value, is the same.

M_{200}

In the same manner, we determined NFW profile parameters from the distortion pattern in the ACS field around the subgroups.

Weak lensing data of individual groups or low-mass clusters do not have sufficient signal-to-noise to provide useful constraints on M_{200} and c_{200} simultaneously. Therefore, we employed the mass-concentration relation given in Mandelbaum et al. (2008), restricting the fit to one free parameter, M_{200}. The results of these fits are summarized in Table 2.2, both for the X-ray centroids and BGGs as tracers of the halo centers.

Scaling relations

G05 showed that the subgroups were consistent with the local $T_X - \sigma_z$ relation (Xue & Wu 2000), a fact which did not change with more spectroscopic data in T09. Here we did not determine 1D velocity dispersions from the redshift distribution of group members, but assumed the projections of 3D halo models. Hence, we are not limited by group member identification. As mentioned before, group centroiding can be a problem.

Although the parameters of individual groups have shifted in this analysis, on average the groups still lie on the local $T_X - \sigma_z$ relation, showing a scatter of similar magnitude as the data in Xue & Wu (2000, Fig. 2.10).

Leauthaud et al. (2010) constrained the $L_X - M_{200}$ scaling relation using weak lensing data of groups in the COSMOS field. The supergroup as a whole is consistent with this scaling relation as well, within the scatter (Fig. 2.10).

These interpretations would be reinforced if we did not take X-ray peak 2 into account and considered the conclusion in George et al. (2012) that BGGs/MMGGs are better tracers of group halo centers than X-ray centroids.

Even though individual groups do not always lie precisely on the determined scaling relations, differences in environment and their effect on the astrophysical processes behind

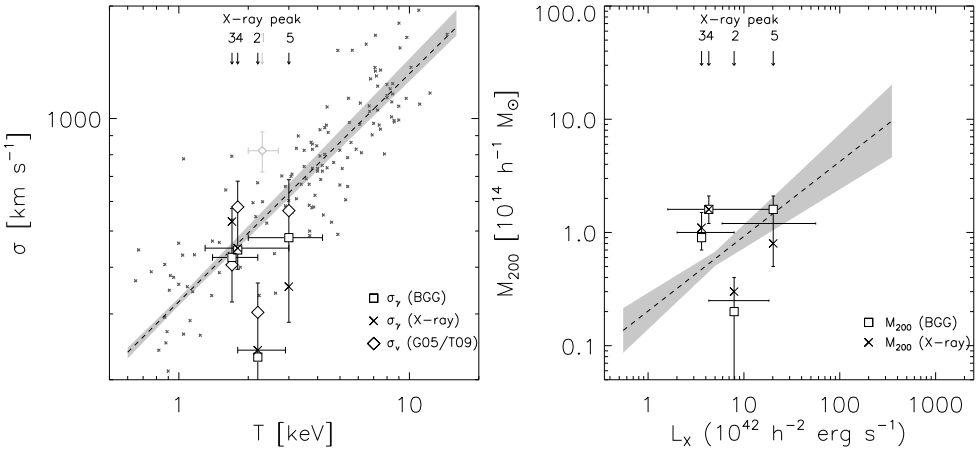


Figure 2.10: Comparison of the properties of SG1120 with observed $\sigma - T$ (Xue & Wu 2000, left) and $L_X - M_{200}$ relations (Leauthaud et al. 2010, right). Results based upon X-ray peaks and BGGs as center of mass are indicated by crosses and open squares, respectively, while velocity dispersions from G05 and T09 are shown as open diamonds. Horizontal error bars are plotted at the vertical median.

the observables used in these analyses create intrinsic scatter around these relations, which is averaged out in a stacking analysis such as employed in Leauthaud et al. (2010).

2.6 Summary

We have performed a weak lensing analysis of the coalescing supergroup SG1120 and showed that the underlying density distribution of matter is well traced by both visual galaxy light and X-ray emission. The subgroups of SG1120 have not yet interacted, but are expected to do so within short timescales, as projected separations are of about $1 - 4$ Mpc (G05). As such, the system is a unique demonstration of hierarchical structure formation.

Slight offsets between the peaks in the galaxy distributions, X-ray gas, and the total matter distributions are well within smoothing scales used and are consistent with an unrelaxed system on the verge of merging. We found that using either X-ray peaks or BGGs as tracers for the halo centers (George et al. 2012) has a minor impact on the derived halo parameters, with results consistent within 2σ error bars. We consider these conclusions to be an indication of the robustness of our results.

Furthermore, while the groups are close enough to be gravitationally bound (G05), the individual group halo masses are low enough compared to their separations to treat them as individual lenses, within parameter error bars.

The fitted profile parameters are consistent with well-demonstrated scaling relations, within the intrinsic scatter created by astrophysical variations (Leauthaud et al. 2010). This is further confirmation that the observed structure of SG1120 is consistent with the paradigm of hierarchical structure formation, providing a unique example of this theoretical framework.

Structures such as SG1120 are rare. In fact, SG1120 should be seen as a single piece of a much larger puzzle, where confirmation from studies of similar structures is a necessity.

The structure of SG1120 is uniquely oriented in the plane of the sky, and the subgroups show no signs of interaction yet, making it well suited to distinguish the various components and overdensities. An example of a well-studied heavier structure is the Cl 1604 supercluster (Gal et al. 2008), where the complex structure presents difficulties in determining accurate masses, either using spectroscopic velocity dispersions (e.g., Lemaux et al. 2012) or weak lensing analyses of a few selected subclusters (Margoniner et al. 2005, Lagattuta 2011).

Especially the extension of studies like these to individual systems of lower mass like SG1120 will present a significant challenge, both in detecting such rare coalescing systems and in obtaining robust and accurate lensing measurements, given the lower S/N. An interesting approach is the combination of large existing spectroscopic group catalogs (e.g., Eke et al. 2004, Berlind et al. 2006, Tempel et al. 2012, Robotham et al. 2011) and recent or currently ongoing large sky imaging surveys of various width and depth, designed for lensing (e.g., Heymans et al. 2012b, Gilbank et al. 2011, de Jong et al. 2013) that are supported by extensive spectroscopic and color information.

Acknowledgements

We thank the anonymous referee for constructive and efficient comments that helped to improve this paper and the robustness of our conclusions. MS acknowledges support from the Netherlands Organization for Scientific Research (NWO). TS acknowledges support from the Netherlands Organization for Scientific Research (NWO), NSF through grant AST-0444059-001, and the Smithsonian Astrophysics Observatory through grant GO0-11147A. Observations taken by NASA HST G0-10499, JPL/Caltech SST GO-20683, and Chandra GO2-3183X3.

Facilities. VLT (VIMOS), VLT (FORS2), Magellan (LDSS3), HST (ACS), SST (MIPS), CXO (ACIS).

2.A Quality and selection criteria for background sources

We assigned several quality flags to the source catalogs during detection and shape measurement.

We used the same rms noise model and deblending parameters as Schrabback et al. (2010) for object detection with `SExtrator`. In addition to detection flags, we required at least eight adjacent pixels with values more than 1.4σ above the background. We defined an initial S/N cut by flagging objects with `FLUX_AUTO/FLUXERR_AUTO`< 10.

We furthermore selected sources with a minimum size compared to the smearing induced by the PSF. We excluded sources for which the half-light radius r_h (as defined in Erben et al. 2001) compared to that of the average star is not smaller than $r_h > 1.2 r_h^*$.

Finally, we selected sources with a KSB shape measurement S/N (defined in Erben et al. 2001) larger than 4, to be consistent with KSB+ studies using a similar definition of the source S/N. In this pipeline, the effect of smearing and shearing by the PSF is for an important part described by the P^g tensor. To avoid being dominated by noise, we excluded sources for which $\mathrm{Tr}(P^g)/2 < 0.1$ (see Erben et al. 2001, for technical details and terminology).

In the final source selection, the catalog of 8273 galaxies is reduced to 7012, ~ 64 galaxies/arcmin^2, that pass all quality criteria from detection and shape measurement.

3

Weak lensing by very low redshift groups: analysis of systematics and robust shape measurements

Weak gravitational lensing is one of the most direct ways to study mass distributions on a wide range of scales. Here we attempt to complement dynamical and luminosity based mass estimations of nearby light galaxy groups with measurements based on weak gravitational shear.

Shape measurements are derived using two pipelines: the Shapelets technique, describing the intensity distribution of faint background sources in Gauss-Hermite expansions, and the extensively used KSB technique. This allows adequate flexibility in modelling the various distortions that affect the images, such as gravitational shear and systematic effects like the PSF.

We present shear estimates obtained from wide-field imaging data of 79 light galaxy groups from the *Zürich Environmental Survey*. We discuss the level of control of systematic errors and compare results between pipelines. The robustness of the methods is promising. We then derive estimates for the velocity dispersion by fitting the lens profile for a singular isothermal sphere, and discuss estimates for the mass and concentration of the groups, by fitting a Navarro, Frenk & White profile. This provides results to be compared to those of groups at intermediate redshifts, exploring ranges in mass and redshift that have not been studied extensively.

M. Smit, K. Kuijken, T. Schrabback, A. Cibinel, C.M. Carollo, and A. Amara

3.1 Introduction

Weak gravitational lensing, since its first detection by Tyson et al. (1990), has become an important and well tested tool in studying the distribution of mass in the universe. The systematic distortion of background images due to the gravitational bending of light rays by a foreground density field can be described by a straightforward geometric model (see e.g. Hoekstra 2005, for a review), instead of relying on complicated physical models of luminous matter or sometimes sparse dynamical tracers.

In the past decade a standard cosmological model has crystallised out of a wide range of measurements of increasing precision. In this model, the matter content of the universe is dominated by a dark, gravitationally interacting component which is thought to form the background density structure for the formation of galaxies and galaxy concentrations as fundamental cosmological building blocks. One of the key aspects in understanding galaxy and structure formation is therefore testing the relation between the visible over-densities and the associated dark matter halos.

On larger scales, up to those of clusters and superclusters, and on small scales, the lensing by individual galaxies first detected by Brainerd et al. (1996), successful estimates of the dark matter content have been made during the last decade (see e.g. Hoekstra et al. 2004, Mandelbaum et al. 2006, Heymans et al. 2006). On intermediate scales however, when trying to detect the common halos of (light) galaxy groups, progress has been slower (see e.g. Mandelbaum et al. (2006) and Leauthaud et al. (2010) for results in the high mass group regime). An important reason for this is the lack of systematic surveys of galaxy groups in combination with robust, extensive catalogues of galaxy groups.

Hoekstra et al. (2001) reported the first measurements of the average mass and mass-to-light ratio (M/L) of light galaxy groups (~ 4 members) around a median redshift of $z = 0.33$, using groups in the CNOC2. This sample was extended in Parker et al. (2005), who found an average velocity dispersion of $\sigma \simeq 245 \pm 18$ km s^{-1} and $M/L_B \simeq 185 \pm 28\ h\ M_\odot/L_{B\odot}$.

As most galaxies are probably found in groups (see e.g. Eke et al. 2004), it is a necessity to understand the group environment and its effect on galaxy evolution. There are indications that a significant part of galaxy evolution already takes place in these environments, before being assembled in clusters (see e.g. Tran et al. 2009). Lensing might prove to be an important independent tool, when studying light, unrelaxed structures with low numbers members.

In recent years, well-defined group catalogues have become available from large surveys like the Sloan Digital Sky Survey (SDSS, Abazajian et al. 2003) and the 2-degree Field Galaxy Redshift Survey (2dFGRS, Colless et al. 2001), that allow systematic studies of galaxy concentrations. Weak lensing may provide a crucial, independent complement to dynamical and luminosity-based mass estimations for these systems that are sometimes sparsely populated and in often unknown phases of evolution. This research, using groups from the *Zürich ENvironmental Survey* (ZENS, Carollo et al. 2013), may be a suitable pathfinder for the Kilo-Degree Survey (KiDS), which is scheduled for 2009 and has an extensive overlap with the 2dFGRS and this sample.

Weak lensing itself has its own limitations, due to intrinsic variations in the shapes of distant sources (a statistical limitation called shape noise) and the ever present systematic errors (most notably the distortions induced by the PSF). In the last decade, several methods for shape measurements have been proposed and used (see e.g. Kaiser et al. (1995) [*KSB*], Hoekstra et al. (1998) and Erben et al. (2001) [*KSB+*], Bernstein & Jarvis (2002) [*BJ02*], Re-

fregier (2003) and Kuijken (2006, hereafter KK06) [*Shapelets*], Miller et al. (2007) [*Lens*fit], among others). In order to make a sensible analysis of the expected weak signal of these light systems, a thorough understanding of these systematics and an adequate method for correcting for these is required.

This paper is ordered as follows. In the next section, we briefly describe the data set. The third section outlines the methods of shape measurement while the fourth presents the actual measurements, the systematic distortions present in the data, and the robustness of our shear estimates. In section 5 we present the signal obtained from these measurements and we discuss several fits to the shear profiles. The final section contains the conclusions.

Unless specifically mentioned otherwise, we assume a ΛCDM cosmology with $\Omega_m = 0.3$, $\Lambda = 0.7$ and $H_0 = 70h$ km s^{-1} Mpc^{-1}. Throughout the paper we discuss various selection criteria for stars and background sources to be used in the final quantitative analysis. Unless mentioned otherwise to highlight or discuss the effect of a certain criterion, figures of source distributions or source comparisons will show the same *final* selection, to make comparing figures throughout the paper intuitively clear, even where the final selection is not yet completely discussed. Finally, we discuss the determination of the group centres. In figures and discussions where group (centre) positions are mentioned, we assume the luminosity weighted mean (LWM) position of the group members as explained in Section 5, except where indicated otherwise.

3.2 Data

For this analysis we use Wide Field Imager (WFI) observations of 79 nearby galaxy groups from ZENS, ESO Large Programme 177.A-0680 (Carollo et al. 2013). ZENS is based on a sample of 1630 galaxies, members of 141 galaxy groups extracted from the 2dFGRS (Colless et al. 2001), and specifically from the Percolation-Inferred Galaxy Group (2PIGG) catalogue (Eke et al. 2004). The 141 ZENS groups are a random selection of the complete 2PIGG sample of 185 which are found in the very thin redshift slice $0.05 < z < 0.0585$, and have at least 5 confirmed members in the 2PIGG catalogue, down to a magnitude $b_J = 19.45$. The ZENS sample is thus statistically complete, and free from possible biases deriving from distance effects. It covers two orders of magnitude in dynamical mass scales, ranging from poor ($\sim 10^{12} M_\odot$), to rich galaxy groups.

This paper is based on the optical B and I images for the first set of 79 groups out of the total sample of 141, that were acquired during several observing runs in January-February 2005 and April 2006-March 2007 with the WFI camera mounted at the Cassegrain focus of MPG/ESO 2.2m Telescope at La Silla. Each group was observed in 5 dithered exposures in the B- and I-band, with single exposure times between 131 and 144 seconds. In each band, the final science image was obtained by taking the median of the five dithered exposures, so as to remove cosmic ray and CCD defects (see Carollo et al. 2013, for a detailed discussion of the data reduction).

The 79 groups have a median velocity dispersion of $\sigma \sim 210$ km sec^{-1} ($38 \leq \sigma \leq 691$ km sec^{-1}) and a median total luminosity of $L \sim 5.0 \times 10^{10} L_\odot$ ($1.1 \times 10^{10} \leq L \leq 5.9 \times 10^{11} L_\odot$). The median number of catalogued group members is 8, with the number of members ranging from 5 to 71.

The observations cover a total area of the sky of ∼22 deg^2 to a depth of $I_{AB} \sim 22$ at S/N

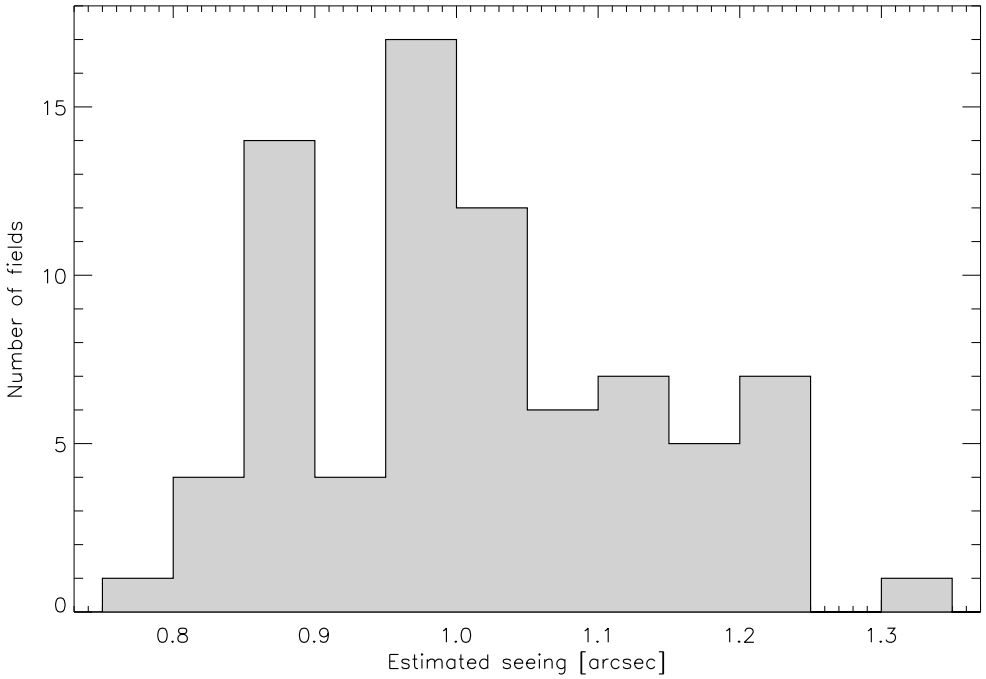


Figure 3.1: Histogram of the seeing of individual fields, as determined from the location of the stellar locus.

~ 10 for extended background sources which we use for our analysis, for a total of more than 10^5 sources ($\sim$ 2000 per field, ~ 2 arcmin^{-2}). The selected fields have a median seeing in the I-band of $\sim 1.0''$ arcsec (Figure 3.1), ranging between $0.75''$ and $1.35''$.

Figure 3.2 shows the magnitude distribution of detected sources that pass the selection criteria that will be discuss in the following sections, i.e. the sources shown are the sources used in the final scientific analysis, but including sources of magnitudes brighter than the $I_{AB} \geq 19$ magnitude criterion used. The magnitude distribution is shown as

$$n_{\rm eff}(I) = \frac{N_{\rm eff}(\Delta I)}{N_{\rm eff,tot}} \tag{3.1}$$

where $N_{\rm eff,tot}$ is the *effective* total source count, including individual weights discussed in Section 3, so the sum of the histogram equals 1.

3.3 Methods

Weak lensing by a foreground density distribution introduces a systematic alignment of background sources. Hence, the (reduced) gravitational shear can be estimated from the shapes of these sources. Any systematic effect that distorts these shapes and can mimic the gravitational lensing signal or affect our shape measurements, such as atmosphere, camera distortion, and pixelisation, needs to be identified and corrected for.

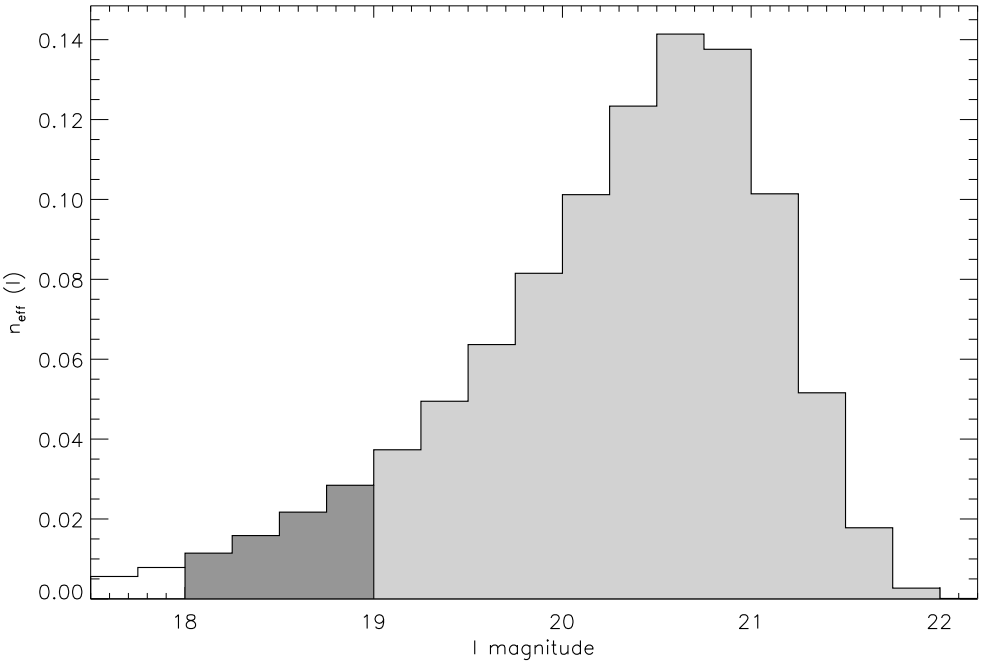


Figure 3.2: Magnitude distribution of the background source sample, determined by the selection criteria discussed in this paper, but including sources of all magnitudes, represented as part of the effective total source selection, so the sum of the histogram equals 1. The light grey area represents the sources used for the final quantitative analysis. The dark grey area represents sources with $18 \leq I \leq 19$, which are not used in the final analysis.

In this section we briefly discuss gravitational lensing and the Shapelets method introduced in KK06. The KSB method and various modifications have been extensively used and were discussed in the aforementioned papers. Therefore, we only briefly touch upon the implementation we use. Both methods have been tested in the Shear TEsting Programme (Heymans et al. 2006, Massey et al. 2007, hereafter STEP) and for a thorough discussion of the results, we refer the reader to these papers.

3.3.1 Weak gravitational lensing terminology

The light rays emanating from different parts of an extended source in the background of a gravitational lens will experience different parts of the (projected) lens potential and will therefore be deflected slightly differently. The resulting change of the source shape, essentially a coordinate transformation $I(x, y) \rightarrow I(x', y')$ is described in terms of the convergence κ and the shear: (γ_1, γ_2)

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{3.2}$$

Since we don't know the intrinsic source sizes or magnitudes, we can only measure the net *distortion* or *reduced shear* $(g_1, g_2) \equiv (\gamma_1, \gamma_2)/(1-\kappa)$:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = (1-\kappa) \begin{pmatrix} 1-g_1 & -g_2 \\ -g_2 & 1+g_1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{3.3}$$

where the transformation is written as a multiplication $(1-\kappa)$, which we don't identify, and a distortion matrix describing the alignment of lensed sources.

The *tangential* shear g_+, often termed the E-mode, is defined as the reduced shear g_1 in the coordinate system with the lens at the origin and the x-axis tangential to the line connecting the observed angular positions of the lens and the source. The g_2 component in the same reference frame is defined as the cross shear $g_\times$ or B-mode.

Following KK06, we define the ellipticity (e_1, e_2) of an object such that a distortion $(-e_1, -e_2)$ will circularise the best-fitting model with constant-ellipticity isophotes. This way, a gravitationally lensed, intrinsically elliptical source can be described as a circular source being sheared twice: first by (e_1, e_2) and then by (g_1, g_2) (see also BJ02).

Background sources display an intrinsic variation in ellipticity and orientation, which is a form of random (shape) noise in estimating the reduced shear. For a source sample with shape noise that is indeed random, we expect $\langle e_i \rangle = 0$, allowing us to recover an estimate for $\langle g_i \rangle$ using a sufficiently large source sample.

3.3.2 Shapelets

In the Shapelets formalism introduced by Refregier (2003), the light distribution of a source is expanded in the orthonormal basis set of Gauss-Hermite functions. This allows for a flexible model and has the advantage that the behaviour of these basis functions under simple transformations (such as an applied shear or smearing by a PSF) is well understood. A PSF model can then be constructed from the shapelet expansions of bright stars in the image, for which the expansions of sources can be corrected in a consistent manner.

In this analysis we use the KK06 implementation of Shapelets (the KK pipeline in STEP), which describes sources as sheared circular objects. As mentioned above, the measured ellipticities are then geometric: a superposition of the intrinsic shape, expressed as a shear, and the gravitational shear. To correct for PSF effects, this sheared circular galaxy model is convolved with the constructed PSF model and fitted to the observed sources.

$$M = P \cdot (1 + g_1 S_1 + g_2 S_2 + \delta_1 T_1 + \delta_2 T_2)(c_0 C^0 + c_2 C^2 + \ldots) \tag{3.4}$$

Here M is the fitted model, the C^n are the basis functions describing a circular source, c_n are the coefficients to be fitted, S_i and T_i are the first order complex shear and translation operators respectively, γ_i and δ_i are the fitted shear and translation (to ensure an optimal centroid fitting) and P is the PSF convolution matrix. Not only is this numerically more stable than a deconvolution of the sources, it also allows for error propagation of the pixel noise.

The sources and stars are all expanded to a fixed order of $n = 12$ in basis functions, to avoid truncation biases in the shape measurements. The images contain enough stars to ensure a well-sampled PSF, $N_* \approx 900$ ($n_* \approx 1$ arcmin^{-2}), and we model the PSF variation of the field by a polynomial of order 5.

Based on the pixel noise, we define individual weights for the sources as

$$w = \left(\sigma_{sn}^2 + \sigma_{g_1}^2 + \sigma_{g_2}^2\right)^{-1} \tag{3.5}$$

where σ_{sn} is the intrinsic shape noise, to avoid spuriously large weights, and σ_{g_i} are the estimated uncertainties in the ellipticity components.

In addition to other selection criteria described in section 4, we select sources with a minimum size compared to the smearing induced by the PSF. We exclude sources for which $\beta_i > 1.1\beta_*$, where β is the scale radius of the Gaussian basis of the shapelet expansion, with subscripts indicating the original source before convolution with the PSF and the stars, respectively.

3.3.3 KSB+

KSB uses the first order effects of distortions induced by gravitational shear and PSF on the weighted second moments of the light distribution of a source to estimate the reduced shear. We have obtained an independent shape measurement catalogue using KSB+ based on the implementation by Erben et al. 2001, with modifications by T.S. (the TS pipeline in STEP).

We define individual weights for this method analogous to those for the KK pipeline, based on the variance of the shear estimators from this pipeline. Based on the findings in the STEP analyses, we expect an underestimation of the shears by KSB of $\sim 8\%$. We have applied a 'fudge-factor' to account for this expected bias, based *only* on the results from simulations.

In addition to other selection criteria described in the following section, we filter for KSB shape measured S/N ≥ 4, to be consistent with KSB+ studies using a similar definition of the source S/N. In this pipeline, the effect of smearing and shearing by the PSF is for an important part described by the P^g tensor. To avoid being dominated by noise, we exclude sources for which $\mathrm{Tr}(P^g)/2 < 0.1$ and the half light radius of the source compared to that of the average star is not at least $r_h > 1.2r_h^*$ (see Erben et al. 2001, for technical details and terminology).

3.4 Shape Measurements

We explain and justify an a priori selection and exclusion scheme, completely 'blind' to the final scientific result.

For our analysis we use as mentioned earlier only sources which are detected at a minimum S/N of ~ 10 in flux. Any sources which suffer from very close neighbours, truncation effects due to chips edges or other uncertainties in their photometry are flagged and excluded. Any remaining neighbour that is extended enough to affect the intensity distribution of the source will induce power in high order coefficients. Since these should be close to zero in our shapelet expansions, we flag and exclude these sources as well.

We estimate a shape noise of $\sigma_{sn} \simeq 0.25$ in the KK measurements, and $\sigma_{sn} \simeq 0.28$ in the TS measurements. The presence of noise in any real data incurs a small fraction of unrealistically large ellipticities, expressed as a shear greater than unity. We exclude any sources with estimated errors $\sigma_g > 0.4$. In the case where we calculate statistical averages, sensitive to outliers, we also exclude total ellipticities $|g| > 2$, which is not too strict in order to avoid inducing a bias in the observed distribution in ellipticities.

At the edges of the field of view, systematic distortions are often most severe, but close to edges of the detector there are not always enough stars to optimally model this. In our selection, we exclude sources from the corners of our images.

For bright sources, the stellar locus is well separated from the galaxy distribution. For faint sources, contamination of one by another is almost a certainty. We only use unsaturated stars with $I_{AB} \leq 22$ for PSF modelling. As mentioned in the previous section, we exclude galaxies smaller or comparable to the widest stellar images from our selection, on a field by field basis to account for variations in the seeing.

In the rest of this section, we discuss the robustness of our shape measurements and the sensitivity to residual systematic effects. We estimate the redshift distribution of the sources and possible contamination by faint foreground sources. This leads to the final selection criteria for background sources to be used in the gravitational lensing analysis of section 5.

3.4.1 Robustness

To investigate the robustness of our shape measurement methods, we have made several consistency checks and compared the results.

Comparing methods

In the left panel of Figure 3.3, we compare the shear estimates derived by the KK and the TS pipelines. Taking into account the uncertainties in both sets of measurements, we fit a linear relation

$$g_{i,TS} = (1 + m)g_{i,KK} + c \tag{3.6}$$

where m represents a multiplicative bias and c an additive bias. Comparing $\sim 1.5 \times 10^5$ sources, we find an excellent agreement between the pipelines on this data set, with very tight constraints on the biases, giving $m \simeq 0.1 \pm 0.1$ % and $c = 0.0 \pm 0.1$ %. Only for the noisiest outermost contour, a slight bias at very high ellipticities can be argued for.

Although no bias means that the measurements are *on average* the same over the whole selection, the variance perpendicular to the best fit line gives an indication of the tightness of the correlation, or the variation between individual measurements. We measure a standard deviation around the best fit line of $\Delta g \simeq 5.6$ %, which can be seen as a small, negligible contribution of random noise compared to the estimated shape noise of our source selection.

This result is a confirmation of the robustness of our shape measurement techniques and for this data set, we can conclude that the two pipelines essentially give the same measurements. The estimated uncertainties and the associated weights do vary between the methods.

Stacked images and single exposures

Using the same initial detection catalogues derived from the stacked science images, we have run the Shapelets pipeline on individual exposures for a single field. These single frames are noisier by a factor of $\sqrt{5}$ and the contribution of the PSF anisotropy pattern to the reduced shear can change significantly from exposure to exposure (see Figure 3.4). These single exposures may contain more defects that are removed in the final dithered stacking, which is why the stacked images were used for detection.

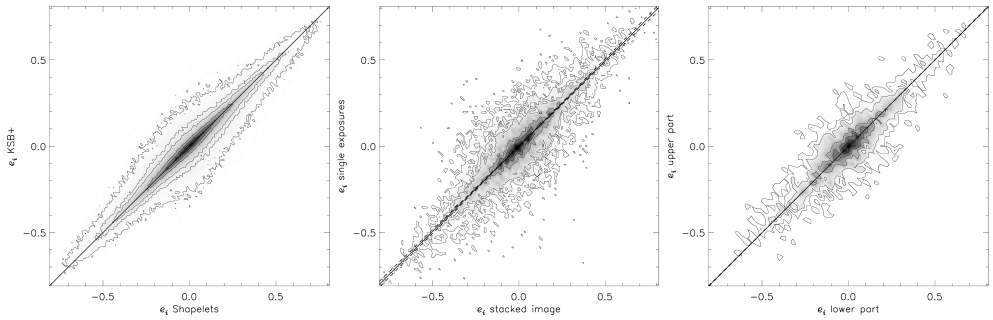


Figure 3.3: *Left:* Shear estimates obtained using the KK Shapelets pipeline compared to shear estimates using the TS KSB+ pipeline. Only at the outermost contour, high ellipticities might start to show a very slight bias. *Middle:* Shear estimates obtained from single exposures of a single field compared to shear estimates obtained from the final stacked image. *Right:* Ellipticities as measured in the "lower" area of the field of view compared to ellipticities of the same sources as measured in the "upper" area of an overlapping field of view. From left to right, the contours exhibit more noise due to statistical power. The spread around the best fit line is comparable. In each figure, the greyscale represents the number count of sources as a 2D histogram. The solid contours, where plotted, correspond to 90%, 70%, 50%, 30%, 10%, 7%, and 4% in counted sources. The solid lines represent the best fit line $e_y = (1 + m)e_x + c$ and the dashed lines represent a 1σ deviation in m. For the left panel $m = 0.1 \pm 0.1$ % and $c = 0.0 \pm 0.1$ %. For the middle panel, $m = -1.1 \pm 2.2$ % and $c = 0.0 \pm 2.2$ % . For the right panel, $m = 0.2 \pm 0.3$ % and $c = 0.1 \pm 0.3$ % .

One of the key questions in the development of accurate shape measurement techniques is the effect of stacking exposures with different PSF on the shapes of the measured sources. Since we have argued earlier that the Shapelets formalism is in principle a method with good flexibility to fit a variety of shapes, we investigate here the performance on the individual PSF patterns for a field where there is significant variation, and the resulting stacked PSF.

In the second panel of Figure 3.3, we compare our shape measurements to ellipticities derived from single exposures, logically only selecting source coordinates actually present in all these exposures. This means we select only stars and background sources from areas on the stacked images that are covered by each of the five exposures, resulting in chip-like regions, as shown in Figure 3.5.

The resulting ellipticities seem slightly noisier than those of the stacked frame, but taking into account the slightly larger measurement error estimates, we find no significant bias between the measurements. The best linear fit gives $m \simeq -1.1 \pm 2.2$ % and $c = 0.0 \pm 2.2$ %.

As we have only $\sim$ 6000 sources in this particular field, we lack the statistics from the comparison between the KK and TS pipelines, which results in *visually* noisier contours. The deviation from the best fitted line is comparable: we measure a standard deviation of $\Delta g \simeq 12$ %, roughly the expected factor $\sim \sqrt{6}$ higher compared to the spread around the best fitted line in the left panel of Figure 3.3.

Another important observation is the lack of bias in the fit between ellipticities. This means that, although the uncertainty estimates are sensitive to the level to which sources extend above the noise, when taking the proper weight factors into account both approaches give the same result, adding to the robustness of the shape measurements.

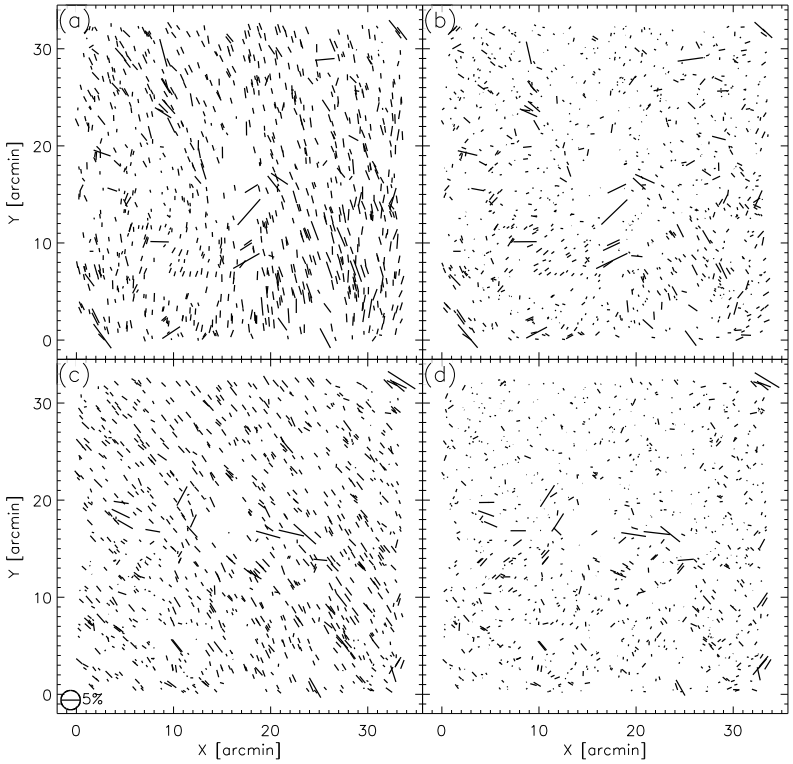


Figure 3.4: Variation of the PSF and PSF correction over exposures of one field. Panels (a) and (b) shows the PSF modelling of an exposure where the pattern displays an almost pure negative g_1 quantitatively; visually, the stars seem to be aligned vertically. In the exposure shown in panels (c) and (d), the PSF introduces an almost pure negative g_2: the pattern of the PSF seems predominantly diagonal.

Overlapping fields

A few of the groups lie close enough to another group from this sample, that observation overlap slightly. This gives us the opportunity to compare measurements of the same sources at different locations in the field of view. In the third panel of Figure 3.3, we compare ~ 600 sources that are detected in two fields.

The choice in which measurement to use as 'reference' value and which as 'comparison' value is completely free. In these data, the x and y image coordinates are almost perfectly aligned with right ascension and declination respectively. In Figure 3.3 we compare the measurements of the northernmost of the two fields, in which the sources lie predominantly in the lower part of the field (i.e. have a lower y image coordinates), to the measurements in the southernmost of the two fields, in which the same sources are detected higher in the field.

We have also made the same comparison selecting on the image x coordinate and we have even made the comparison using random selections. The results in all cases are identical within significance. It is less instructive to fit a linear relationship in this case, but for completeness we note here that we find $m = 0.2 \pm 0.3$ % and $c = 0.1 \pm 0.3$ %. A more

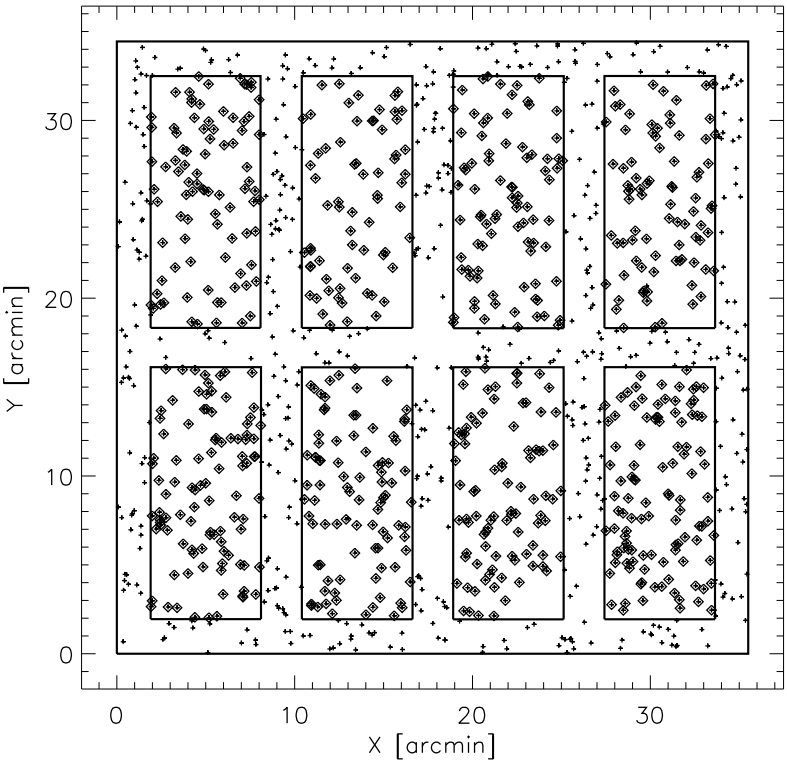


Figure 3.5: The areas on the stacked image that are covered by all five individual exposures. As an example, we show the positions of the stars over the field. Diamonds represent stars in areas with a full depth of exposures.

important result is that the variation around the best fit line is, within significance, the same as for the comparison between the KK and the TS pipelines. We find a standard deviation of $\Delta g \simeq 5.0$ %.

3.4.2 Possible systematics

Given the consistency of our shape measurements between independent methods and observations, we expect any remaining systematic signal to be carried in the actual shapes of background sources and not to be introduced by the measurement thereof.

Since this survey specifically targets groups of galaxies, a concern with this dataset is that we are not dealing with random fields. Although there are a few exceptions, in most of the fields we expect an overdensity concentrated near the center of the field of view. Several diagnostics that are devised for random fields, such as the estimation of a remaining correlation between the PSF and the corrected source shapes, become difficult and less meaningful to interpret.

Near the edges and corners of any field of view, fitted models such as those for the PSF variation or astrometric corrections are expected to become less constrained. For a sample of

lenses predominantly located near the center of the field of view, possible systematics induced by these effects occur at roughly the same distance from each lens and will most likely not disappear by averaging over the lenses.

Astrometric effects

Roughly 30% of the fields have sufficient overlap with SDSS to estimate the possible presence of astrometric distortions. Using the SDSS positions as a reference, we determine the displacement of sources in our catalogues and investigate whether the displacements are random or display a pattern dependent on position on the field. The median displacement is ~ 1.8 pixels or $\sim 0.4''$.

Intuitively, one can understand the effect of astrometry as a distortion as follows. If the field is slightly stretched in the x direction (or compressed in the y direction), this would induce a small, positive e_1: on average, all sources would display a slight horizontal alignment. Similarly if the field is stretched in the y direction (compressed in the x direction), the effect would be a small, negative e_1.

Following the same reasoning for the diagonal directions for e_2, it is clear that we can detect a possible effect by measuring the variation of the displament $(\Delta x, \Delta y)$ of our detection with respect to SDSS as a function of position:

$$\frac{\partial(\Delta x)}{\partial x} - \frac{\partial(\Delta y)}{\partial y} \qquad \text{and} \qquad \frac{\partial(\Delta y)}{\partial x} + \frac{\partial(\Delta x)}{\partial y} \tag{3.7}$$

for e_1 and e_2 respectively.

Translating this into $(e_+, e_\times)$ with respect to the groups, we do not detect a significant deviation from zero at any distance from the expected lens positions, with maximum error bars of the order of 10^{-4}.

PSF discontinuities

As already mentioned, a dominant source of systematic distortion of source shapes is the PSF. Using the shapelet expansions of the stars, we investigate in Figure 3.6 how much spurious signal can be expected to be contaminating the true lensing signal. The differences between the shapes of the stars and the PSF model fitted to those stars should be consistent with zero for any observation. This provides a diagnostic for residual PSF distortions present in the data that should not be affected by image geometry or lens positions in the fields of view, as long as the distribution of foreground stars is sufficiently random.

Plotted are both the tangential and cross distortion deduced from the stellar shapes, mimicing a $(g_+, g_\times)$ induced by the original PSF, and the residual effect carried in the stars after correction. The average PSF pattern exhibits only a small cross signal which largely disappears after correction. The tangential pattern shows more variation, however. While the *average* residual distortion after correction for the PSF vanishes by construction, there remains a subpercent but significant postive signal close to the group centers and a similar but negative signal on large scales.

To investigate the cause of this residual signal, we have looked at the spatial variation of the stellar residuals over all fields. In Figure 3.7 we show one of the results of these tests, the variation of residual ellipticity of foreground stars over the field. We have used all $\sim 6 \times 10^4$ stars of all fields and their positions with respect to the central pixel of each image.

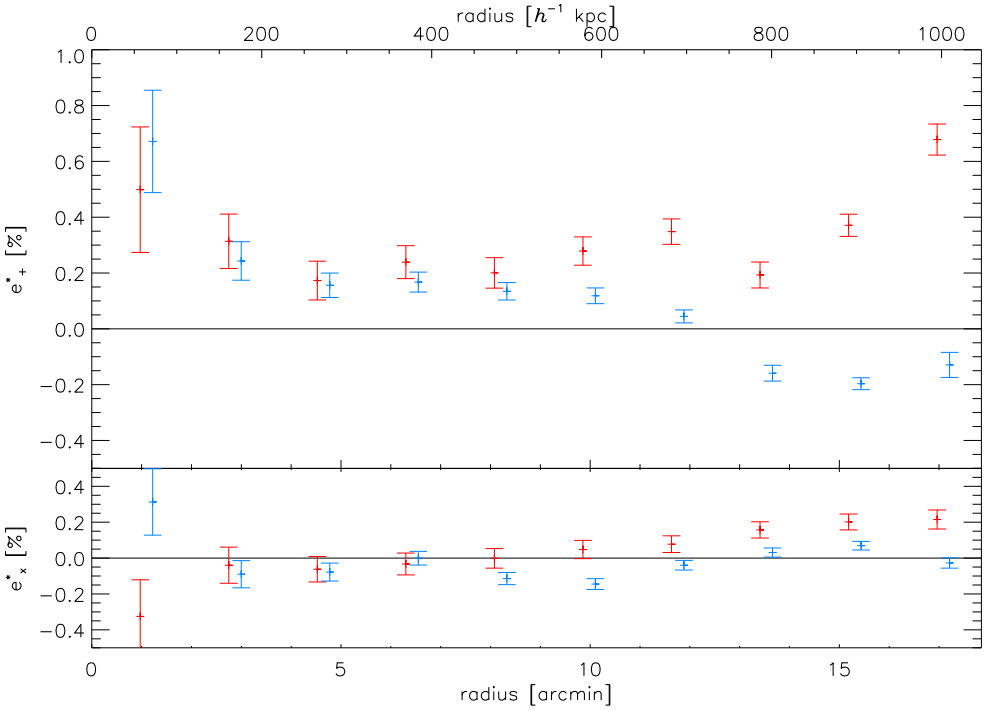


Figure 3.6: Tangential ellipticity (e_+) and cross signal ($e_\times$) with respect to the ZENS group centres, carried in the stars before (red) and after (blue, slightly offset to the right) correction for the PSF. The total number of stars used is $\sim 6 \times 10^4$.

The layout of the WFI chips stands out clearly, showing higher residuals in regions that lie between chips in one or more exposures. This is to be expected, as the PSF model is dominated by the chip-like regions covered by all individual exposures. This model is less representative in the regions where one of more exposures don't contribute to the PSF, resulting in a less accurate PSF correction in these regions.

To test the effect of these PSF residuals in the stars, we plot in Figure 3.8 the same tangential and cross distortions as in Figure 3.6, but using only stars that are covered by all individual exposures. The result is a significant improvement in the residual tangential PSF distortion. Except for two of the outer bins, the signal carried by the stars disappears after PSF correction, showing that the PSF model is much more accurate in these regions. The residual cross signal is also dimished further, being consistent with zero on all scales.

For our PSF model and correction and for the analysis and interpretation in Section 5, we will only select stars and background sources that are covered by all individual exposures. This selection reduces the number of stars and background sources by approximately 35%, to $N_* \approx 3.5 \times 10^4$ and $N_s \approx 10^5$. A more efficient approach would be to account for the PSF of each individual exposure and combine the information, similar to the *lens*fit pipeline (Miller et al. 2007), but this is beyond the scope of this project.

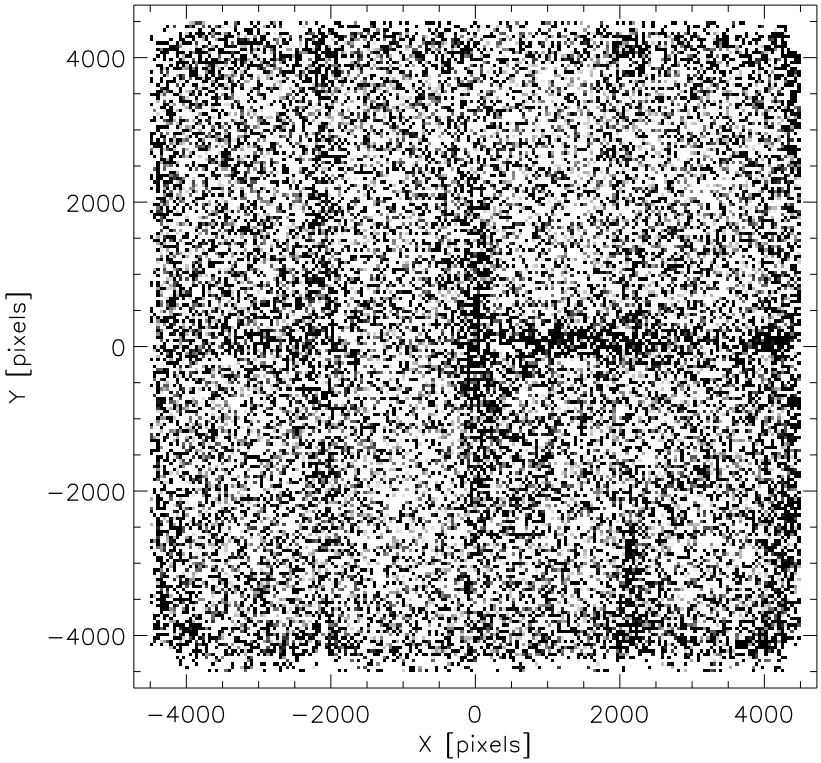


Figure 3.7: Residual absolute ellipticities of the stars after correction for the PSF, represented as a grey scale, dark where the residual signal is higher and lighter where the PSF correction is more efficient. The residuals are most prominent in regions of the image not covered by all individual exposures, where the PSF model is probably less accurate, resulting in the grid pattern around the eight chip-like regions as shown in Figure 3.5.

Other effects

As we show later, we attempt to separate the foreground from the background using a magnitude selection. Any faint foreground members contaminating our source sample will not carry a lensing signal, diluting the final detection. Another possible source of error may come from faint group members, not identified as such spectroscopically, that contaminate our sample of background sources. This is a possibility, as the survey specifically targets areas on the sky with an overdensity at the lensing redshift. If this affects our signal, we expect it to be diluted as well (Mandelbaum et al. 2006).

In the next section, where we determine the foreground-background separation and calculate lens signal and efficiency, we will show that this contamination is likely to be minimal.

3.4.3 Background redshift distribution

Since the groups are relatively close by, the lensing geometry is far from optimal. However, an advantage is that we can use many background sources, selecting all sources with $I_{AB} \geq 19$,

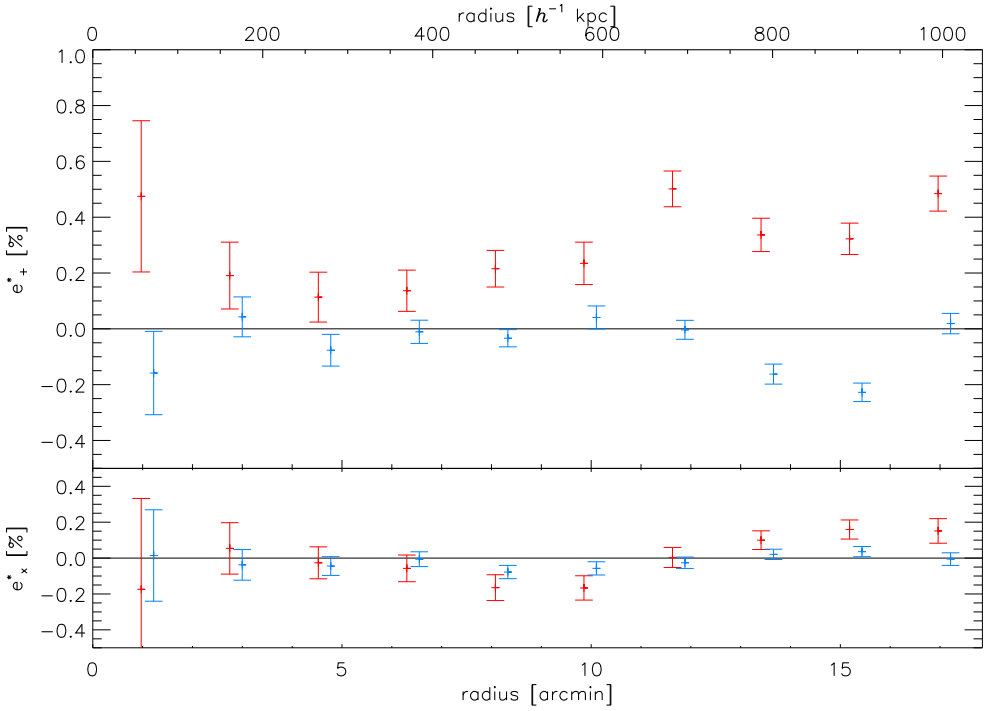


Figure 3.8: The tangential ellipticity (e_+) and cross signal ($e_\times$) carried in stars that are present in all individual exposures, before (red) and after (blue, slightly offset to the right) correction for the PSF. The total number of stars used is $\sim 3.5 \times 10^4$.

resulting in a selection of $19 \leq I_{AB} \leq 23$. Using the first epoch results of the VIMOS VLT Deep Survey (Le Fèvre et al. 2005, hereafter VVDS), we estimate a median redshift of the selected background sources of $z \sim 0.43$.

In Figure 3.9 we show the expected redshift distribution of the selected source sample, taking into account the full VVDS redshift distribution for each magnitude and the individual weights described in Section 3. The dashed line represents the median redshift of the ZENS group sample.

The strength of the gravitational shear signal for a given lens depends of the distances of the observer to the lens (D_d), the background sources (D_s) and the distance from the lens to the source (D_{ds}). With lens redshifts between $0.05 < z < 0.0585$, we calculate an angular diameter distance range of the lenses to be $202 \leq D_d \leq 233\ h^{-1}$ Mpc. Taking into account the variation in lens redshifts and the full estimated background redshift distribution, we calculate the variation in lensing efficiency D_{ds}/D_s, shown in Figure 3.10.

The weighted mean lensing efficiency is found to be $D_{ds}/D_s = 0.86$, shown as the dashed line in Figure 3.10, offset from the main peak due to the skewdness of the distribution. Since the gravitational shear signal is linear in D_{ds}/D_s, we can use this single number to relate reduced shear to physical quantities such as velocity dispersion and lensing mass.

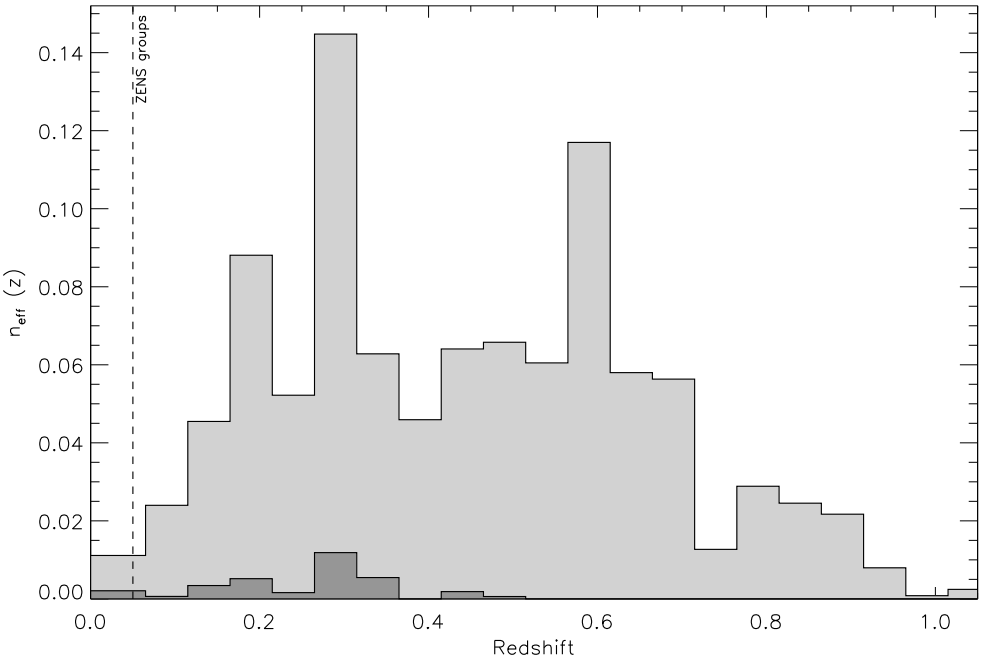


Figure 3.9: Expected redshift distributions for selected background sources with $I \geq 19$, represented as part of the effective total source selection. The high redshift tail of the distribution extends beyond the plot. The dashed line shows the median ZENS group redshift. For completeness, the dark grey histogram shows the expected redshift distribution of sources with $18 \leq I \leq 19$, corresponding to Figure 3.2.

Foreground-background separation

Comparing to the VVDS redshift distribution, we expect about $\sim 0.6\%$ of our background sources to be at the same or lower redshift than the groups. More strictly, sources at or just beyond the lens redshift can be seen as contamination as well as they will not carry much signal and mostly contribute noise. As a more conservative estimate, we expect $\sim 3.5\%$ of our background selection to lie at a redshift below $z = 0.1$. For foreground objects unrelated to the galaxy groups, this should be a fair indication. However, as said before, the ZENS targets areas on the sky with an overdensity at the lensing redshift, so these estimates should be read as lower limits due to possible faint group members that haven't been spectroscopically identified in the 2PIGG catalogue.

If there is a significant contamination of unidentified, faint group members, we expect an increased source density near the center of the groups, instead of a uniform background distribution. We attempt a more accurate estimate of the possible contamination by randomizing positions of background sources before only selecting sources that fall on the chip-like areas as described in Section 4. We use 1000 realizations for each of the fields and mask the areas where one or more of the individual exposures has a gap between the chips. We then count the sources in bins around the group centers and compare this to the actual source count in Figure 3.11.

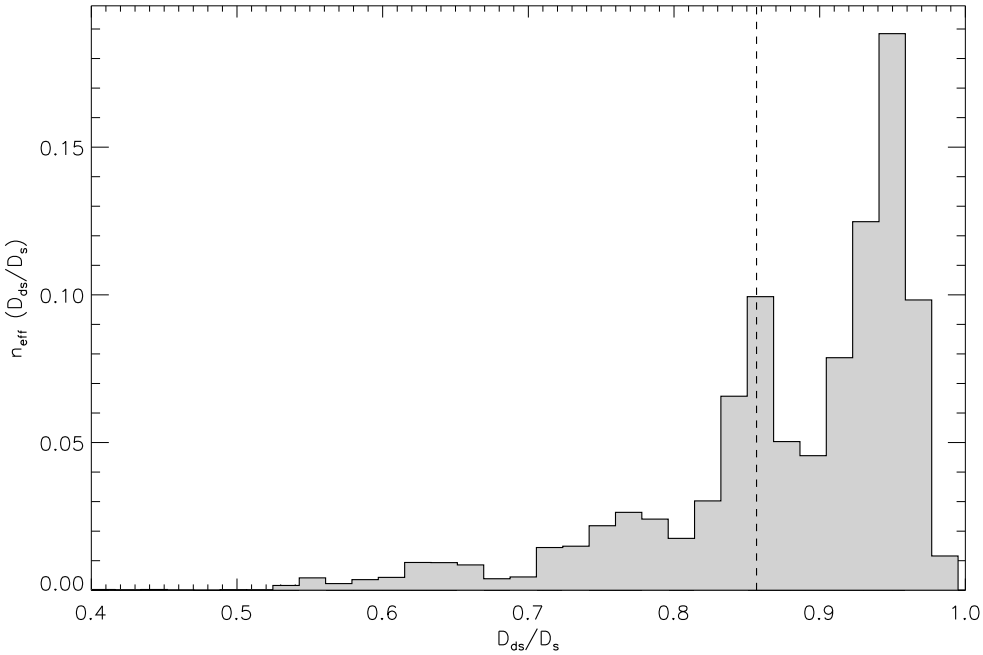


Figure 3.10: Expected lensing efficiency D_{ds}/D_s distribution for selected background sources with $I \geq 19$, represented as part of the effective total source selection. The dashed line shows the weighted mean lensing efficiency.

We normalize the average source count of the 1000 realizations to 100% in each bin and normalize the source count of the actual detected sources accordingly. The shaded grey area represent the variation in the randomizations. The red line representing the normalized, actual source count shows slightly more deviation from the simulated source counts than is expected from statistical variation, but we do not see a clear increase in source density in the central bins.

This method allows us to take into account the variation in areas masked between fields, due to different dither patterns, as well as source density variations between fields and the expected locations of the group centers with respect to the center of the images. This does not account for more complex effects, such as spacial gaps in detection due to very bright stars (an example of which can be seen in the center of Figure 3.4) or extraction flags due to coinciding bad pixels from different exposures. It should therefore be kept in mind that Figure 3.11 serves as an indication.

3.5 Gravitational shear signal and mass estimates

Weighting each field according to the density of background sources and their associated weights, we derive from the 2PIGG catalogue an average velocity dispersion of $\sigma = 229 \pm 21$ km s^{-1} and derived dynamical mass (Eke et al. 2004) of this group sample of $M_\sigma \simeq 4.1 \pm 1.1 \times 10^{13} h^{-1} M_\odot$.

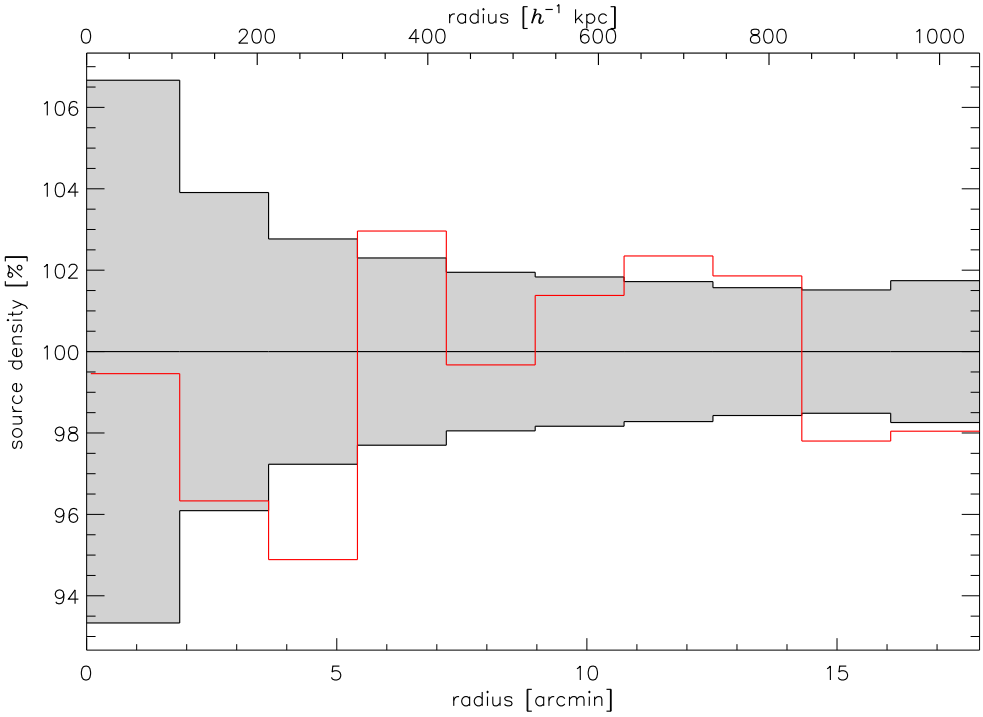


Figure 3.11: Source counts in bins around the group centres *(red line)*, compared to 1000 randomisations. The mean simulated source counts are normalised to 100% and the actual data counts are normalised accordingly. The shaded grey area represents the statistical variation in the randomisations.

This section will cover the estimation of the reduced shear signal, using a weighted statistical mean and convex hull peeling (CHP), to assess the influence of outliers. We estimate the signal around the luminosity-weighted mean (LWM) positions of the member galaxies and around the brightest group galaxies (BGGs). To relate our shear signal to physical quantities such as velocity dispersion an mass distribution, we fit a singular isothermal sphere (SIS) and a Navarro, Frenk & White (Navarro et al. 1996, NFW) profile to the ensemble average, and briefly touch upon how well the dynamical and luminous properties of the group correlate with the lensing signal.

3.5.1 Convex Hull Peeling

Mathematically, the convex hull of a set of points S in $\mathbb{R}^n$ is the intersection of all convex sets in $\mathbb{R}^n$ containing S. For our purposes, the convex hull of a two dimensional set of (g_1, g_2) points is the minimum subset of points that, if connected, forms a polygon that encloses the rest of the set.

The CHP or "onion peeling" method for determining the data center of a convex set, equivalent to the median for one dimension, consists of removing (peeling away) the convex hull and repeating the process for the remaining subset of points, until one point or no points are left. If one point is left, it determines the two dimensional median. If no points are left, the

mean of the last convex hull determines the two dimensional median, similar to the median in one dimension.

The main advantage of CHP is a lower sensitivity to outliers. Disadvantages include a less accurate determination of the underlying (g_1, g_2) (depending on the point cloud distribution) and the inability to apply weights to individual shear estimates. Here, we use it as an alternative to a direct weighted average, to asses the sensitivity to outliers of our shear determination.

3.5.2 Group centre of mass

One of the key problems in dealing with light galaxy groups is the determination of the center of mass. We discuss a priori the motivation for the two most likely tracers for the group center of mass, the LWM position and the BGGs.

Without external information, such as x-ray observations for heavy groups and clusters (e.g. Leauthaud et al. 2010), the only way to estimate the position of the center of mass of each group is using the luminous components, i.e. the member galaxies. One can either use the average position of the member galaxies, possibly weighted by their luminosities, or assume that the brightest member galaxy is also the heaviest and is on average located close the the center of the group halo. One can't use the lensing signal itself to estimate the position of the center of mass, not only because in this case the groups are too light to be detected in a projected mass reconstruction, but also because the signal would, by construction, be biased towards positive noise peaks.

For this group sample, the weighted and unweighted average positions correspond well with each other, and we decide to use the luminosity weighted mean positions. For groups with only a few members and a dominantly luminous galaxy, the LWM position and the position of the BGG lie close to each other, whereas the richer groups provide more statistical power.

In Figures 3.12 and 3.13 we plot in dark and light grey the predicted tangential shear signal around the LWM group positions and the BGGs, using for simplicity a SIS model based on the 2PIGG velocity dispersion and the background source positions and weights. To account for the complex masking we applied, we used the locations of the sources used for shear measurements in order to make an accurate prediction for this lens sample and data set. First we predict the signal around the LWM positions, both for the case where these would be a perfect tracer for the center of mass and if the BGGs would actually represent the center of mass. We then do the same for the expected signal around the BGGs in both cases. It is clear from Figures 3.12 and 3.13 that an accurate estimate for the center of mass is paramount.

Overplotted are the weighted mean tangential shear measured with the KK and TS pipelines in each bin. As in Figure 3.3, the two pipelines agree very well within error bars. In both plots, the cross signal is consistent with zero, as is required by theory. In the left plot, a clear positive signal can be seen around the LWM group positions, whereas in the right plot, the signal is consistent with zero on almost all scales around the BGGs. We also note the diminished signal in the outer bins and propose that this is caused by residual PSF distortion, similar to Figure 3.8.

This *suggests* that, at least for this group sample, the LWM positions represent the positions of the center of mass better than the BGGs. However, we should be cautious against

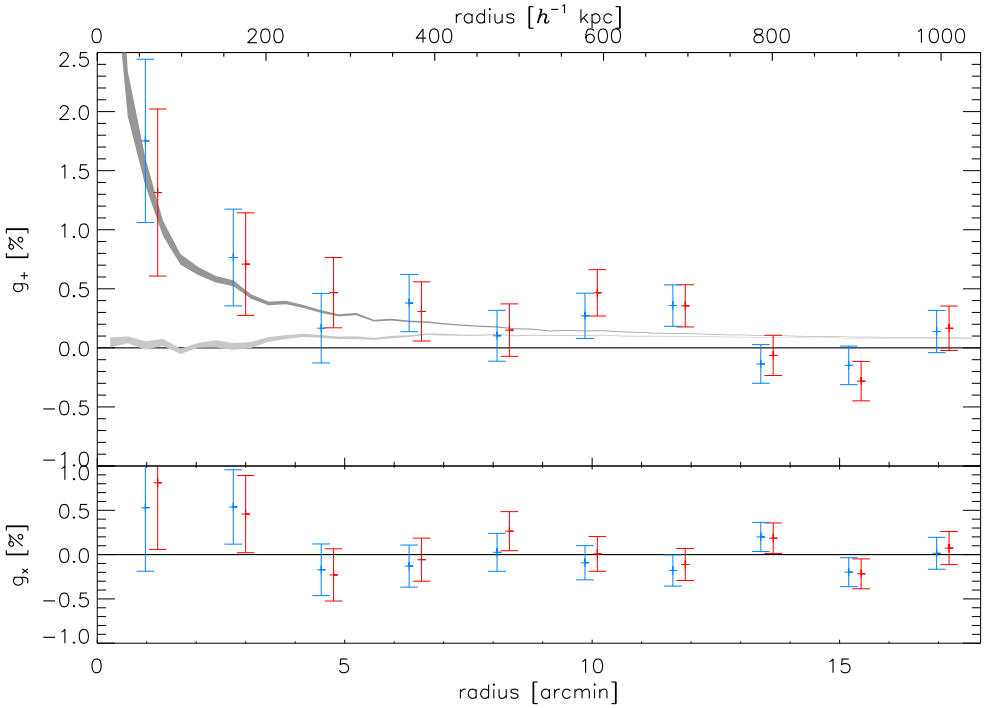


Figure 3.12: Binned tangential shear measured around the LWM group positions. The KK *(blue)* and TS *(red)* measurements agree well within error bars. The dark grey shaded area is the predicted tangential shear for the group velocity dispersions from the 2PIGG catalogue, assuming the centre of mass to coincide with the LWM group positions and assuming a SIS density profile for each group, where the thickness represents the spread due to the variation in σ within the group sample only. The light grey shaded area is the same around the BGGs.

deciding on the measurement that corresponds best to expectations and note that the measurement is very noisy. We have motivated our choice for two possible tracers for the underlying center of mass and we will continue our analysis using the signal around the LWM group positions, but we do not presume this assumption to be a general conclusion for light galaxy groups.

One should keep in mind that for a less accurate estimate of the group center of mass, the signal is more diluted. This means that a possible improvement in group center determination could increase the signal and our result might be slightly biased low. However, this effect could be well within error bars and the conclusion that our result is a lower limit is likely too strong.

We further note that all bins are completely independent in the sense that each background source is used only once, neglecting the $\sim$ 600 sources ($\sim$ 0.5%) detected in two fields. At larger radii, the effect of background large-scale structure (LSS) can give a correlation (Hoekstra 2003). Since the groups are at such a low redshift, this effect might be present, depending on the actual redshift of the lensed background sources.

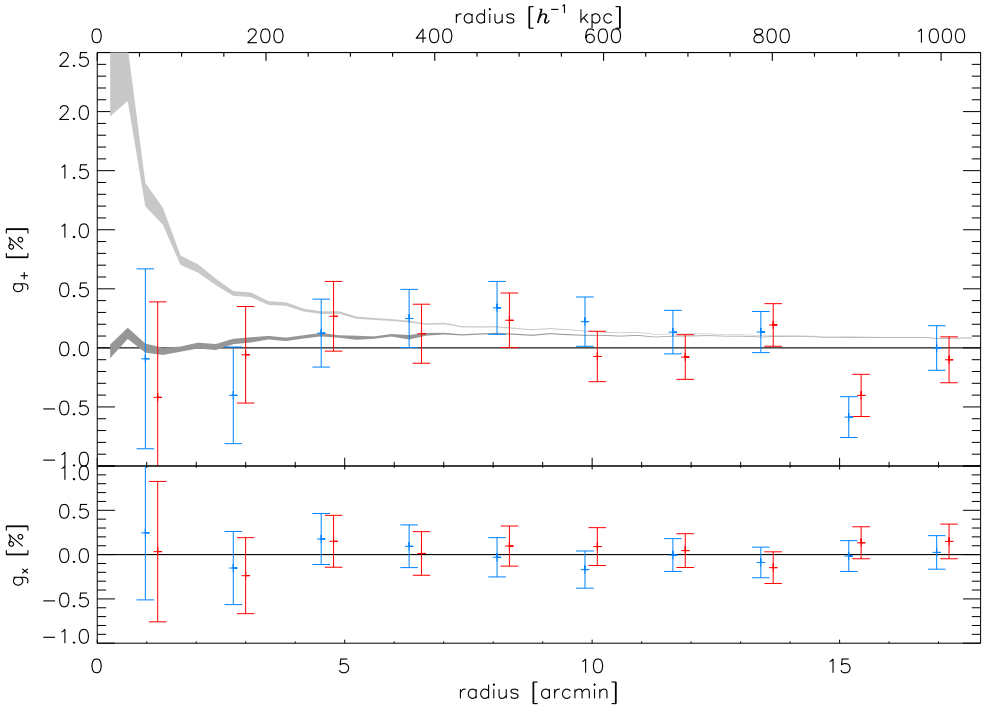


Figure 3.13: Binned tangential shear measured around the BGGs. The KK *(blue)* and TS *(red)* measurements agree well within error bars. The dark grey shaded area is the predicted tangential shear for the group velocity dispersions from the 2PIGG catalogue, assuming the centre of mass to coincide with the LWM group positions and assuming a SIS density profile for each group, where the thickness represents the spread due to the variation in σ within the group sample only. The light grey shaded area is the same around the BGGs.

3.5.3 Velocity dispersion

To compare our lensing signal directly to the dynamical velocity dispersion σ_{dyn}, we fit a SIS model. In Figure 3.14, we show the resulting fit to the KK and TS measured tangential shear, using a weighted mean and a CHP radial binning of the data. Based on Figure 3.8, we exclude the outer three bins from our fit, although we note that both in Figure 3.8 and Figure 3.14 the outermost bin agrees very well with prediction and fit.

We derive $\sigma_\gamma \simeq 283^{+94}_{-150}$ km s^{-1} for Shapelets and $\sigma_\gamma \simeq 286^{+99}_{-164}$ km s^{-1} for KSB+. If we use CHP to bin the KK data, we find $\sigma_\gamma \simeq 251^{+90}_{-151}$ km s^{-1}. In Table 3.1, we summarize the results of these fits.

The results agree very well with each other and with the dynamical estimate of $\sigma_{dyn} = 229 \pm 21$ km s^{-1}, although the error estimates show that this is a noisy, albeit significant, result.

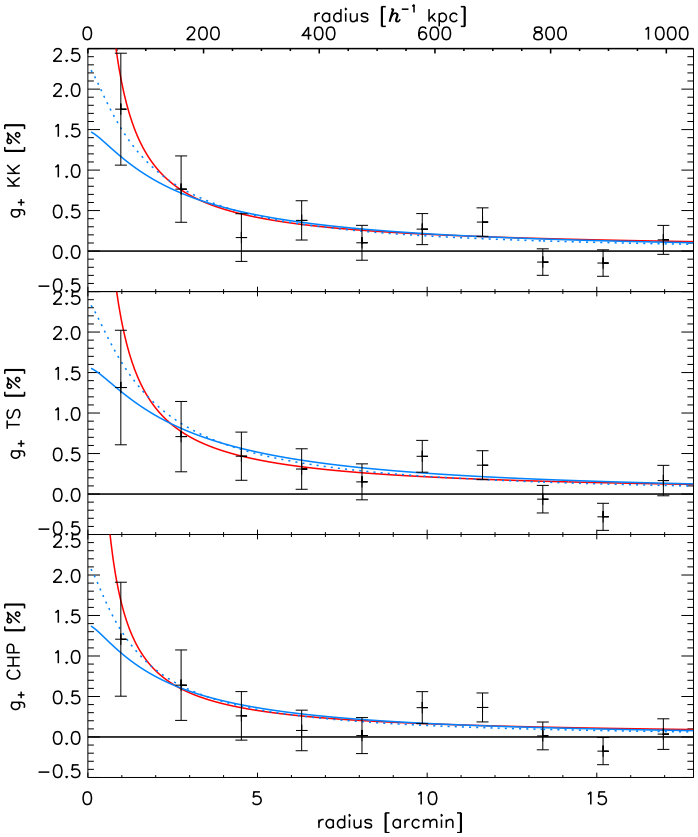


Figure 3.14: The best fit tangential shear profiles for the weighted averaged KK *(upper panel)*, the weighted averaged TS *(middle panel)*, and CHP KK *(lower panel)* bins. The red line is the best fit SIS profile and the blue lines are the best fit NFW profiles, assuming mass-concentration relations with $c_0 = 7$ *(solid)* and $c_0 = 10$ *(dotted)*.

3.5.4 Mass and concentration

We attempt to derive an estimate for the mass and concentration of the average group halo by fitting a NFW profile, where we keep M_{200} and c as free parameters. In Figure 3.15, we show the χ^2 distributions over the parameter space, we we have again excluded the three outermost bins from the fit. We also summarize the results in Table 3.1.

While M_{200} is relatively well determined, it is difficult with the low signal-to-noise of our result to get constraints on the concentration parameter c. Only for the weighted KK bins a lower limit on c can be determined.

Various studies have indicated that the halo mass and concentration are actually correlated, both simulations (see e.g. Neto et al. 2007, Duffy et al. 2008) and observations (Mandelbaum et al. 2008, M08). Following M08, we assume a mass-concentration relation of the

Method	Full group sample				
	SIS	NFW			
		Full fit		$c_0 = 10$	$c_0 = 7$
	σ_γ	M_{200}	c	M_{200}	M_{200}
KK	283^{+94}_{-150}	$2.0^{+3.3}_{-1.4}$	$6.8^{+5.8}_{-4.0}$	$2.3^{+2.5}_{-1.6}$	$2.8^{+3.6}_{-2.1}$
TS	286^{+99}_{-164}	$3.2^{+4.3}_{-2.4}$	$3.1^{+4.5}_{-3.1}$ (1)	$2.7^{+3.0}_{-2.0}$	$3.5^{+3.9}_{-2.5}$
CHP	251^{+90}_{-151}	$1.9^{+2.4}_{-1.2}$	$2.0^{+6.5}_{-2.0}$ (1)	$1.6^{+1.9}_{-1.2}$	$2.0^{+2.6}_{-1.6}$

Table 3.1: The best fit profile values for the weighted mean KK and TS measured shear and CPH bins for comparison, for the full group sample. Velocity dispersion σ_γ in [km s^{-1}] and masses M_{200} in $[10^{13}h^{-1}M_\odot]$. From left to right: the best fit SIS velocity dispersion; the best NFW fit for two independent parameters; M_{200} for $c_0 = 10$ and $c_0 = 7$.
(1) Lower limits not constrained.

form

$$c = \frac{c_0}{1+z}\left(\frac{M}{M_0}\right)^{-\beta} \tag{3.8}$$

where we assume $\beta = 0.1$, $M_0 = 10^{14}h^{-1}M_\odot$.

Figure 3.15 shows this relation for $c_0 = 7$, 10, and 13. Our sample of light galaxy groups falls right in the gap in mass range in M08, between L_* lenses and richer galaxy groups. Given the low lensing signal of our data, we only fit the resulting one parameter profile for $c_0 = 7$ and 10 and present the results in Table 3.1. Figure 3.14 shows the corresponding best fit NFW profile for a fixed mass-concentration relation with $c_0 = 7$ and $c_0 = 10$.

3.5.5 The heaviest groups

Since most of our groups have only a few members, with more than 75% having $N \leq 10$, an important question is which observed group property (taken from the 2PIGG catalogue directly) is a good indicator for the total group mass. Since gravitational lensing is sensitive to all gravitational matter, luminous and dark, it is instructive to compare to group properties based on the luminous content only.

To investigate how the lensing signal as a tracer of the total (projected) mass distribution correlate with various possible tracers for group mass from the 2PIGG catalogue, we split our group sample by dynamical velocity dispersion σ_{dyn}, total group luminosity L and group richness N. For simplicity we again assume a mass-concentration relation with $c_0 = 7$ for the NFW profile in all cases.

Table 3.2 presents the results for each pipeline and binning method for the various selections. Although there is an overall increase in significance for the estimate of σ_γ, there is no observable trend over selection method, pipeline or binning method. For the M_{200} estimate, the selection seems to have no consistent effect.

It is clear that for this study, we are limited by the signal-to-noise of the lensing signal. These considerations are important, however, when comparing studies of galaxy groups based

Method	Heaviest groups					
	Selection					
	σ_{dyn}		L		N	
	σ_γ	M_{200}	σ_γ	M_{200}	σ_γ	M_{200}
KK	356^{+102}_{-145}	$3.6^{+5.7}_{-3.0}$	333^{+105}_{-161}	$3.5^{+5.5}_{-2.9}$	294^{+101}_{-164}	$2.9^{+4.4}_{-2.4}$
TS	329^{+102}_{-152}	$3.0^{+4.7}_{-2.5}$	274^{+100}_{-172}	$2.1^{+2.9}_{-1.7}$	263^{+94}_{-157}	$2.4^{+3.2}_{-2.0}$
CHP	251^{+98}_{-187}	$2.5^{+3.6}_{-2.1}$	276^{+101}_{-178}	$2.4^{+3.7}_{-2.0}$	297^{+99}_{-157}	$3.2^{+4.5}_{-2.6}$

Table 3.2: The best fit profile values for the weighted mean KK and TS measured shear and CPH bins for comparison, for the 'heaviest' groups. Velocity dispersions σ_γ in [km s^{-1}] and masses M_{200} in $[10^{13}h^{-1}M_\odot]$, with $c_0 = 7$ assumed in all cases. From left to right: σ_γ and M_{200} for groups with the highest dynamical σ_{dyn}; total luminosity L; richness N .

on and using different methods, such as lensing, dynamics and, for heavier groups, X-ray observations.

3.6 Conclusions

We have been able to determine the gravitational shear signal around a sample of nearby light galaxy groups, making use of extended sky coverage, providing us with more than 10^5 background sources, and careful indentification of possible sources of systematic errors.

3.6.1 Shape measurements

A good understanding of possible systematic effects is paramount for an accurate determination of a subtle signal.

We have used two independent shape measurement methods, the Shapelets pipeline of Kuijken (2006) and the TS implementation of KSB+ described in Erben et al. (2001), both tested in STEP. Both methods yield very consistent results and using the shapelet pipeline we also find consistency between different single exposures and in overlapping fields. This leads us to conclude that the shape measurements are robust and reliable.

Our observations are well covered by foreground stars, with on average $N_* \approx 900$ per field for a well sampled PSF. We have used the residual differences between the actual stellar shapes and the PSF model as a diagnostic that should not be affected by the specific properties of this data set. We have shown that, except for the regions in the image stacks that are not covered by one or more individual exposures, the PSF model seems to be very accurate, leading to residual stellar distortions that are consistent with zero in both ellipticity components on almost all scales. We therefore conclude that a single, overall PSF model is not accurate in these regions with partial coverage and exclude for our analyses both stars and sources from these areas. We do caution for a residual tangential PSF distortion on large scales at the $\sim 0.1\%$ level in the stars and based on these tests do not include those scales in the final lensing analyses.

Using SDSS coverage of $\sim 30\%$ of our observations to test for any astrometric distortion effect, we found excellent agreement, with uncertainties at the $\sim 0.01\%$ level in order of magnitude. Using the VVDS magnitude-redshift observations, we estimate our foreground-background separation to be accurate, expecting less than 1% contamination of faint foreground sources. To account for the image geometry and complex masking pattern, we have used 1000 randomisations of background positions to look for any trend that could indicate contamination of our source sample by unidentified faint group members and found none. We are confident that any of these effects are negligible compared to the intrinsic variance for this data set.

3.6.2 Measured signal

Using both the LWM group positions and the BGGs as tracers for the underlying centre of mass, we find a significant tangential signal around the LWM positions. The absence of a signal around the BGGs underlines the sensitivity of the signal to the accuracy in the determination of the center of mass.

Analysing the signal around the LWM group centres, we use both a statistical mean and a CHP binning scheme, as the former is an unbiased estimator while the latter is less sensitive to outliers. Well within error bars, we find good agreement between shape measurement pipelines and between binning methods, showing the robustness of the detected signal.

3.6.3 Profile shapes

We have fitted both a SIS and a NFW profile to our measured signal and derived estimates for the velocity dispersion σ_γ, halo mass M_{200} and concentration c, summarised in Tables 3.1 and 3.2, that agree well with dynamical estimates (Eke et al. 2004) and studies of light galaxy groups at higher redshift (Parker et al. 2005). Although the unfavourable lensing geometry, range in group size and low average group mass leads to large error bars, the estimations for σ_γ and M_{200} are significant. We conclude that it is not possible to accurately constrain c as a free parameter with this data set.

We also assumed a mass-concentration relation and looked at plausible values for the normalisation constant c_0 based on M08. A value of $c_0 = 7$ seems to agree best with results from both pipelines and binning methods. Using this value, we have investigated the correlation of our lensing signal with group properties that possibly trace the halo mass, velocity dispersion σ_{dyn}, luminosity L, and richness N. Although there seems to be an increase in lensing signal overall, the increase in uncertainties from using half of our data set limits the significance of this trend.

3.6.4 General conclusion

Using independent shape measurement methods, we find overall a good agreement between gravitational lensing and dynamical methods on the different estimations of the average group mass content. Our lensing results, although having a low signal-to-noise ratio, agree well with those obtained from groups at higher redshift. Furthermore, we have shown that weak gravitational lensing works for these low redshift lenses, where the weak signal barely extends

above the intrinsic shape noise of our sample of background galaxies, filling in gaps in mass and redshift ranges.

3.6.5 Future work

Our weak lensing study can be an important, independent result complementing other studies of the group environment (see Carollo et al. 2009, *in prep.*), which is the most common galaxy environment and perhaps dominant in galaxy evolution (see e.g. Tran et al. 2009). Given the current sample, the lightness of the groups and the lensing geometry, we expect to have reached the limits of what can be obtained with our current data. Our results show that lensing can be an useful, practical technique to study matter distributions down to low redshift and low density.

We look forward to robust catalogues of galaxy groups becoming available for large scale surveys better or specifically suited for lensing, such as the CFHTLS (see e.g. Parker et al. 2007) or the soon to commence KIDS, which will cover an important part of this group sample. More detailed studies will be possible with these big surveys, allowing one to investigate trends e.g. with group richness, redshift and proximity to LSS. Given the promising results so far in this field, the study of galaxy groups using gravitational lensing has much potential left to be explored.

Acknowledgements

We like to thank Malin Velander for making her NFW fitting routines available to us. The CHP method was first brought to our attention by a poster presentation of Hyunsook Lee at Penn State University. MS acknowledges support from the Netherlands Organization for Scientific Research (NWO). TS acknowledges support from the Netherlands Organization for Scientific Research (NWO), NSF through grant AST-0444059-001, and the Smithsonian Astrophysics Observatory through grant GO0-11147A.

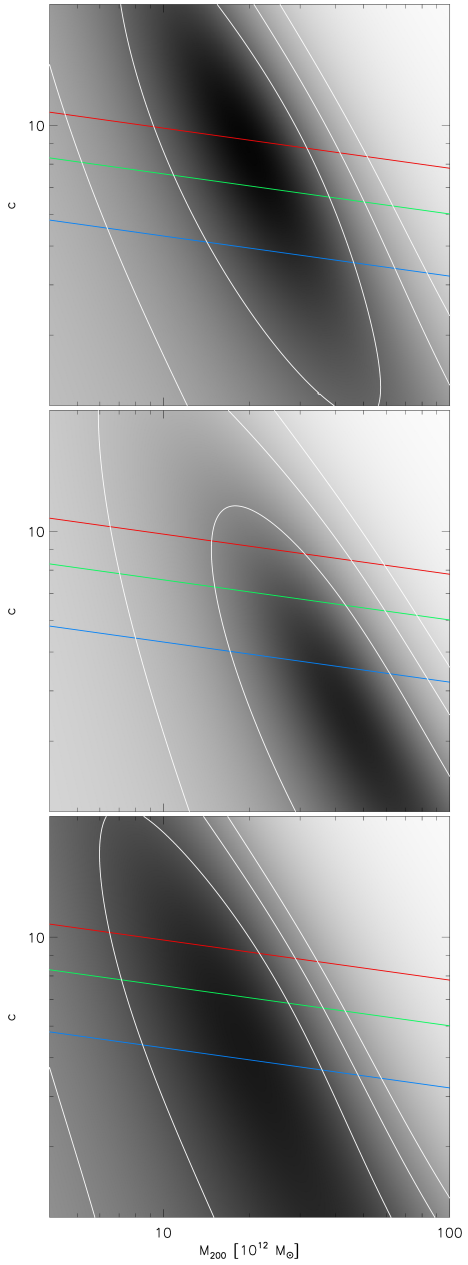


Figure 3.15: The χ^2 distributions *(greyscale)* for the M_{200}, c fits to the weighted averaged KK *(upper panel)*, the weighted averaged TS *(middle panel)*, and CHP KK *(lower panel)* bins. The white contours show the 68%, 95%, and 99% confidence limits. The red, green and blue lines represent mass-concentration relations assuming c_0 = 13, 10, and 7 respectively.

4

Chasing the peak: optimal statistics for weak shear analyses

Weak gravitational lensing analyses are fundamentally limited by the intrinsic distribution of galaxy shapes. It is well known that this distribution of galaxy ellipticity is non-Gaussian, and the traditional estimation methods, explicitly or implicitly assuming Gaussianity, are not necessarily optimal.

We aim to explore alternative statistics for samples of ellipticity measurements. An optimal estimator needs to be asymptotically unbiased, efficient, and robust in retaining these properties for various possible sample distributions. We take the non-linear mapping of gravitational shear and the effect of noise into account. We then discuss how the distribution of individual galaxy shapes in the observed field of view can be modeled by fitting Fourier modes to the shear pattern directly. This allows scientific analyses using statistical information of the whole field of view, instead of locally sparse and poorly constrained estimates.

We simulated samples of galaxy ellipticities, using both theoretical distributions and data for ellipticities and noise. We determined the possible bias Δe, the efficiency η and the robustness of the least absolute deviations, the biweight, and the convex hull peeling estimators, compared to the canonical weighted mean. Using these statistics for regression, we have shown the applicability of direct Fourier mode fitting.

We find an improved performance of all estimators, when iteratively reducing the residuals after de-shearing the ellipticity samples by the estimated shear, which removes the asymmetry in the ellipticity distributions. We show that these estimators are then unbiased in the absence of noise, and decrease noise bias by more than $\sim 30\%$. Our results show that the convex hull peeling estimator distribution is skewed, but still centered around the underlying shear, and its bias least affected by noise. We find the least absolute deviations estimator to be the most efficient estimator in almost all cases, except in the Gaussian case, where it's still competitive ($0.83 < \eta < 5.1$) and therefore robust. These results hold when fitting Fourier modes, where amplitudes of variation in ellipticity are determined to the order of 10^{-3}.

The peak of the ellipticity distribution is a direct tracer of the underlying shear and unaffected by noise, and we have shown that estimators that are sensitive to a central cusp perform more efficiently, potentially reducing uncertainties by more than 50% and significantly decreasing noise bias. These results become increasingly important, as survey sizes increase and systematic issues in shape measurements decrease.

M. Smit, and K. Kuijken
Astronomy & Astrophysics, **Volume 609**, A103 (2018)

4.1 Introduction

Since the first gravitational shear detections (Tyson et al. 1990), the statistical analysis of weak gravitational lensing effects has become recognized as a competitive cosmological tool. With the advent of precision cosmology, meaningful interpretations of statistical agreement or tension between various models and datasets become increasingly important.

Weak gravitational lensing produces slight magnification and distortion effects by bending the paths of light rays. Although analyses of the former have produced important scientific results (e.g., Hildebrandt et al. 2009, Van Waerbeke et al. 2010) and it has in fact been demonstrated that combined analyses can give better constraints (Hildebrandt et al. 2011, Ford et al. 2012), most scientific information has come from the analysis of weak shear distortions. To access that information, one has to be able to (1) measure the shapes of lensed background sources accurately, (2) understand the intrinsic distribution of these shapes and the effects of shear and noise on statistical inference, and (3) obtain the statistical power to probe the subtle perturbations of this distribution by weak shear.

For the first part, a multitude of shape measurement methods have been explored, among which are foremost methods based on surface brightness moments (e.g., Kaiser et al. 1995, Rhodes et al. 2000) and model fitting methods (e.g., Kuijken 1999, Bernstein & Jarvis 2002, Hirata & Seljak 2003, Refregier & Bacon 2003, Kuijken 2006, Miller et al. 2007, Kitching et al. 2008), with various alternative or combined approaches (Bernstein & Armstrong 2014, Herbonnet et al. 2017, Zhang et al. 2015).

Community-driven projects for optimal and robust shape estimates (Heymans et al. 2006, Massey et al. 2007, Bridle et al. 2010, Kitching et al. 2012, Mandelbaum et al. 2015) have led to a further decrease in measurement variances and a better understanding of remaining systematic effects and biases (e.g., Voigt & Bridle 2010, Bernstein 2010, Kacprzak et al. 2012, Melchior & Viola 2012, Refregier et al. 2012).

For the last part, the last two and a half decades have also known dramatic improvements in statistical power. Surveys that are finished, ongoing, and planned such as COSMOS[1] (Leauthaud et al. 2007), CFHTLenS[2] (Heymans et al. 2012b), RCSLenS[3] (Hildebrandt et al. 2016), KiDS[4] (de Jong et al. 2013), DES[5] (Dark Energy Survey Collaboration et al. 2016), LSST[6] (Ivezić et al. 2019), Euclid[7] (Laureijs et al. 2011) steadily increase in size (sky coverage and depth) and imaging quality, including a significant improvement in understanding and correcting for systematic effects (e.g., Heymans et al. 2012a,b, for CFHTLenS).

This increasing statistical power is necessary to overcome the inference limit set by the intrinsic galaxy shape distribution, known as shape noise. Unlike many forms of noise, such as measurement uncertainties that are often dominated by Poisson processes, there is no reason that the ellipticities of background galaxies follow a Gaussian distribution. In fact, studies of galaxy morphologies (Lambas et al. 1992, Rodríguez & Padilla 2013) suggest that late type galaxies may exhibit a roughly uniform axis ration distribution.

This departure from Gaussianity is clearly demonstrated in Section 4.2, when comparing

[1] http://cosmos.astro.caltech.edu/
[2] http://www.cfhtlens.org/
[3] http://www.rcslens.org/
[4] http://kids.strw.leidenuniv.nl/
[5] http://www.darkenergysurvey.org/
[6] https://www.lsst.org/
[7] http://www.euclid-ec.org/

the shape distribution of the CFHTLenS shape measurements catalog (Heymans et al. 2012b, Figure 4.2) to a simulated Gaussian distribution (Figure 4.1). This implies that commonly used Gaussian estimators, such as the (weighted) mean estimate of the central peak of the distribution or the variance for its width, are not necessarily optimal for the inference of the underlying gravitational shear.

For example, if the tails of the ellipticity distribution decline more slowly than the Gaussian $\exp\left(-x^2\right)$, then more elliptical galaxies contribute more shape noise. There have been many weighting and clipping schemes suggested to minimize biases and uncertainties in weak shear inference (Bonnet & Mellier 1995, Van Waerbeke et al. 2000, Bernstein & Jarvis 2002). Alternative approaches include distribution symmetrization (Zhang et al. 2017), or using ensembles of galaxies in Bayesian analyses or nulling techniques (Bernstein & Armstrong 2014, Herbonnet et al. 2017), so that the step of individual shape measurement before inference of the underlying shear is bypassed.

In this article, we explore an alternative approach by reviewing statistical estimators that are more suited to a distribution with a pronounced central cusp and slowly declining tails. Estimator optimality would include a low or vanishing estimator bias and a high accuracy by a low spread in estimates. These aspects should be robust for various possible distributions, as samples of background galaxies are comprised of different populations.

We then highlight the use of these estimators in fitting the shear pattern in the field of view with Fourier modes (Fourier Mode Fitting, FMF). This approach provides an alternative to smoothed gridding and locally sparse and therefore poorly constrained estimates. It provides statistical information constrained by the whole field of view, and incorporates fluctuations in background number densities and estimated measurement uncertainties automatically. For subsequent scientific analyses, the Fourier model allows for relatively straightforward, analytic approach to fundamental quantities, such as a power spectrum or mass density reconstruction.

We note that we focus on the statistical inference from samples of measured shapes, for various possible intrinsic shape distributions, that is, the propagation of shape noise. This is a single but fundamental step in improving the accuracy and fidelity of weak lensing analyses. We do not perform a subsequent cosmological analysis, which would require addressing other well-known sources of bias and systematic effects. These include for example selection and detection biases (e.g., Hirata et al. 2004, Miller et al. 2013, Jarvis et al. 2016) among others on the instrumental and computational side. Other sources include physical effects that affects the interpretation of the measured signal, such as the effects of baryons, or the redshift distribution and intrinsic alignments of lensed background galaxies background sources. The shear signal we recover in this paper would represent a combined signal, which would then need to be interpreted.

The remainder of this paper is organized as follows. We will briefly review the necessary definitions of galaxy shapes and the weak lensing formalism in Sect. 2, referring the reader to excellent reviews such as Bartelmann & Schneider (2001), Schneider (2006), Hoekstra & Jain (2008), for more in-depth approaches. We review the necessary statistical framework in Sect. 3, where we discuss galaxy shape distributions and statistical estimators, including definitions for efficiency and bias, before expanding on FMF. In Sect. 4 we describe the various possible simulations and data, and analysis methods. In Sect. 5 we discuss the results and the scientific implications. Section 6 gives a summary of our conclusions.

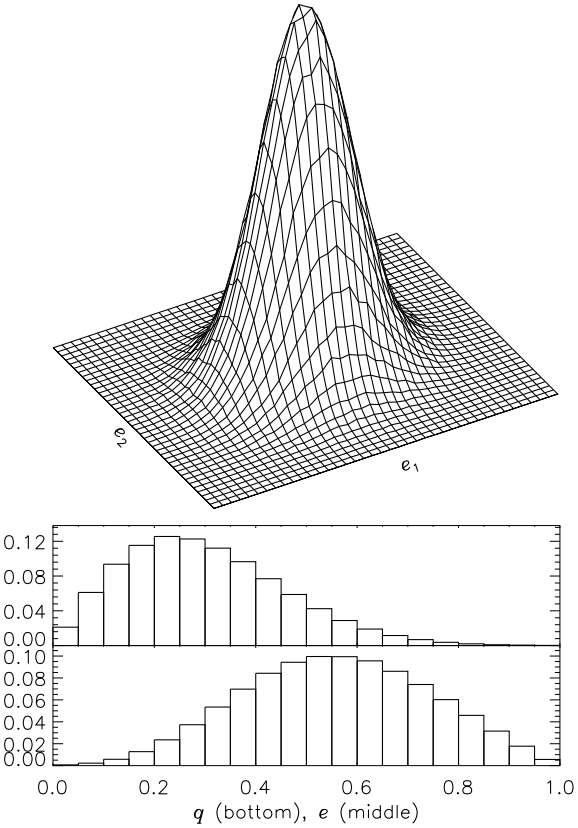


Figure 4.1: Gaussian ellipticity distribution and corresponding axis ratio distribution. Top: a 2D histogram of ellipticities. Middle: histogram of the absolute ellipticity $|e|$. Bottom: histogram of the ellipse axis ratio q.

4.2 Weak lensing

Gravitational lensing is the effect of curved space-time on the paths of light rays from distant sources to the observer as they pass through the gravitational potential of foreground structures. This geometrical effect leads to a displacement of point sources on the projected plane of the sky. The differential effect on images $I(x, y)$ of extended sources leads to magnification and distortion effects, know as the convergence κ and the shear $\gamma = \gamma_1 + i\gamma_2$, directly related to the surface mass density. This is commonly described as a coordinate transformation

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \tag{4.1}$$

resulting in the lensed image $I(x', y')$.

Weak lensing magnification analyses (e.g., Hildebrandt et al. 2009, Van Waerbeke et al. 2010) require the intrinsic (distribution of) source sizes or magnitudes. In weak shear anal-

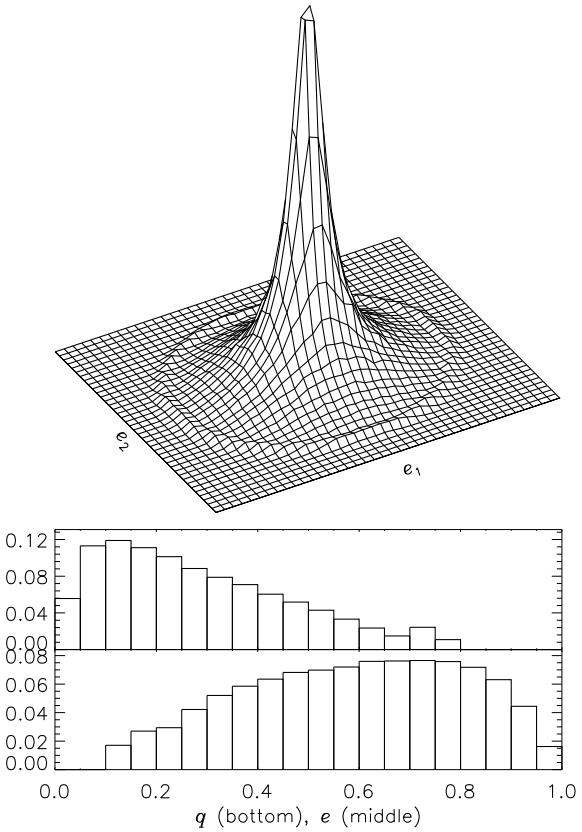


Figure 4.2: Ellipticity and axis ratio distributions of the CFHTLenS catalog. Top: a 2D histogram of ellipticities. We note that the ring-like feature at $e \approx 0.8$ is due to noisy outliers forced to a maximum e by the shape measurement pipeline, but see also Figure 4.8. Middle: histogram of the absolute ellipticity $|e|$. Bottom: histogram of the ellipse axis ratio q.

yses, the focus lies on the net distortion or reduced shear $g = g_1 + ig_2 \equiv (\gamma_1 + i\gamma_2)/(1 - \kappa)$:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = (1 - \kappa) \begin{pmatrix} 1 - g_1 & -g_2 \\ -g_2 & 1 + g_1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \tag{4.2}$$

where the transformation is written as a multiplication of $(1 - \kappa)$ (which leads to the magnification) and a traceless distortion matrix describing the alignment of lensed sources in the foreground potential.

The distortion effect of weak lensing shear on images of background galaxies depends on their intrinsic shape distribution. While galaxies often have complex morphologies, it is adequate to describe images by their quadrupole brightness moments or their ellipticities, and the respective response to weak shear distortions.

A common definition of the shape of an image with elliptical isophotes is the ellipticity $e = e_1 + ie_2$, defined as the reduced shear needed to create this image from an image with

circular isophotes (Bernstein & Jarvis 2002, Kuijken 2006). This gives an axis ratio $q = \frac{b}{a}$ as

$$q = \frac{1-|e|}{1+|e|} \quad \Leftrightarrow \quad |e| = \frac{1-q}{1+q} = \frac{a-b}{a+b}\,, \tag{4.3}$$

and position angle θ via

$$e = |e|\,(\cos 2\theta + i \sin 2\theta)\,. \tag{4.4}$$

As an example, we compare a Gaussian (e_1, e_2) distribution to the distribution observed in the CFHTLenS shape measurement catalog in Figures 4.1 and 4.2.

This complex notation gives a most straightforward formulation of the resulting ellipticity $\tilde{e}$, after transforming an image with ellipticity e by a distortion g, by Seitz & Schneider (1997)

$$\tilde{e} = \frac{e+g}{1+g^* e} \quad \text{for} \quad |g| \leq 1\,, \tag{4.5}$$

with g^* the complex conjugate of g.

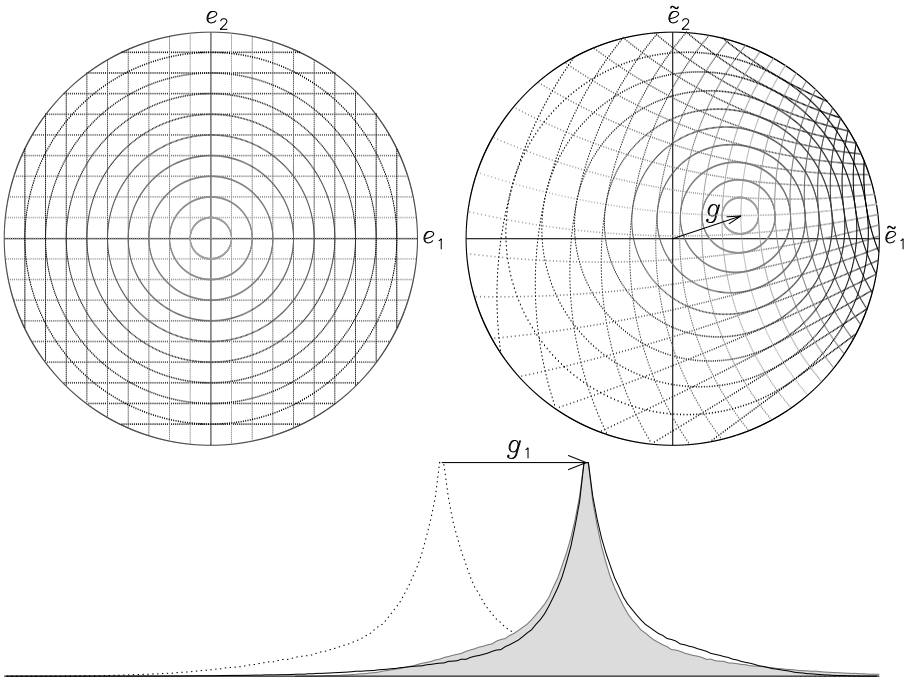


Figure 4.3: Top: the non-linear mapping of ellipticities (with $|e| \leq 1$) by an exaggerated gravitational shear of $g = 0.33 + 0.11i$. Bottom: the asymmetry introduced in the ellipticity distribution, highlighted for the e_1-component.

The non-linear effect of gravitational shear on the ellipticity parameters is shown in the top panel of Figure 4.3. Through statistical estimation, we can attempt to infer from an ensemble of galaxy shapes the underlying shear, if we assume the intrinsic ellipticity distribution $P(e)$ to be centered around zero ellipticity. In other words, one assumes no preferred direction on the sky.

This non-linear response to weak shear distortions gives rise to the asymmetry in the observed ellipticity distribution, as shown in the bottom panel of Figure 4.3. The shifted central peak of the distribution is unaffected by this non-linearity and therefore a direct tracer of g.

The canonical approach is a weighted mean μ, where weighting schemes attempt to minimize systematic effects from noise, size and brightness. As observed by Seitz & Schneider (1997), the expectation value $\langle\tilde{e}\rangle$ does not depend on $P(e)$ in the absence of noise. The mean of an ensemble of measured ellipticities is then an asymptotically unbiased estimator for the underlying shear g.

In the presence of noise, however, these estimations suffer from unavoidable biases in the estimated shear (Melchior & Viola 2012, Kacprzak et al. 2012). Furthermore, the variance of an estimator such as the mean, or more generally, the scale of the estimator distribution, does depend on the intrinsic ellipticity distribution $P(e)$. Informally put, the smaller the estimator variance, the more 'trustworthy' the estimates and the more efficient the estimator. A more efficient estimator reduces the uncertainties in and therefore the error bars or confidence intervals of parameter estimates.

The smearing of the sheared distribution by noise affects central value estimations, but the peak location itself is still an unbiased tracer of the shear.

4.3 Statistical framework

In this section, we discuss various estimators, after reflecting upon estimator properties, such as bias, efficiency, and robustness, and their interpretation. We then propose ways to apply this to fitting individual Fourier modes to a shear field.

4.3.1 Bias, efficiency and robustness

We will use the term bias, or Δe, when referring to the difference between the central value of an estimator, such as the expected value or mean $\langle\hat{e}\rangle$, and the population parameter e. We will use the term residuals, or $r_i = e_i - \hat{e}$, when talking about the differences between one sample estimate and the elements of that sample, that is, the individual measurements $e_i = e_{i,1} + ie_{i,2}$.

We note that we write $r_i = e_i - \hat{e}$ for simplicity throughout this paper, but we employ Equation 4.5 to calculate the residuals, unless specifically noted otherwise. The absolute residual ellipticity of a single measurement with respect to the sample estimate is then the norm $|r_i|$.

The difference $\Delta e = \langle\hat{e}\rangle - e$, commonly referred to as simply the bias of the estimator, is formally called the mean-bias $\mu_{\Delta e}$. An estimator is then called asymptotically mean-unbiased, if for an increasing number of estimations $\hat{e}$, the mean estimate $\mu_{\hat{e}}$ converges toward the parameter value of the underlying population. This is commonly simply referred to as unbiased. Here we have changed notation from $\langle\hat{e}\rangle$ to $\mu_{\hat{e}}$, to emphasize the method of determining the central value of a set of estimates.

We do this, because there are other possible definitions of unbiasedness, such as median-unbiasedness, in which case the median estimate $M(\hat{e})$ converges toward the true parameter value. By the central limit theorem, it is often appropriate to assume an asymptotically normal distribution of the estimator $\hat{e}$ (not to be confused with the distribution $P(e)$ of the population

parameter e), when the number of estimations increases. This validates the general use of mean-unbiasedness. In practice, sample sizes needed for convergence toward a normal estimator distribution can be very large and one should take care when assuming asymptotic normality when making statistical inferences from a few measurements.

The efficiency of an estimator can be defined in terms of its variance. For unbiased estimators, this variance is bounded from below by the Cramér-Rao lower bound (Rao 1945, Cramér 1946), which in short means that there is an absolute maximum efficiency that can be obtained. For some distributions, such as the Gaussian distribution, this limit can be calculated analytically[8]. In other cases, it is useful to define a relative efficiency

$$\eta_{\hat{e}} = \frac{\sigma_0^2}{\sigma_{\hat{e}}^2}, \tag{4.6}$$

where σ_0^2 is the variance of a comparison estimator, such as the mean. Then, if for example $\eta_{\hat{e}} > 1$, the estimator has a lower variance than the mean and is therefore more efficient in finding the central value of the population parameter distribution. An estimator that achieves the Cramér-Rao lower bound for all possible parameter values is for this reason also known as a minimum variance estimator.

Again, if the assumption of asymptotic normality is not appropriate, another definition of the scale of distribution of the estimator can be used instead of the variance, such as the median absolute deviation (MAD). In such cases, care should be taken with the coverage of that scale, which is simply the percentage of estimates with lower residuals than the scale. In case of a Gaussian distribution, the standard deviation has a coverage of 68.3%. The MAD has, by definition, a coverage of 50%.

To avoid comparing apples with oranges, we will use chosen percentiles as scale, so the coverage is defined. For instance, we define the 68.3% scale $s_{68.3}$ as the residual value for which 68.3% of the estimates has an equal or lower residual. In case of asymptotic normality, $s_{68.3}^2$ will converge to the same value as the estimator variance.

We note that we can do this, since in our simulations the true population parameter value e is known[9]. In general, the coverage of a definition of scale is not known, confusing the interpretation of any relative efficiency.

In conclusion, we define the efficiency of an estimator $\hat{e}$, relative to the mean μ_e, at a certain percentile coverage p, as

$$\eta_{\hat{e};p} = \frac{s_{\mu_e;p}^2}{s_{\hat{e};p}^2}. \tag{4.7}$$

Finally, we label an estimator $\hat{e}$ as robust (in a qualitative manner), when $\hat{e}$ retains low or zero bias and high efficiency in a wide range of possible distributions. A robust estimator is desirable, since it makes the choice of estimator for a parameter with unknown distribution more objective. As an example, the mean is optimally efficient in case of a Gaussian parameter distribution, but since the mean has low resistance against departures from Gaussianity (such as outliers), it is not the most robust.

[8] We will omit a more detailed discussion, since it's applicable only to certain distributions and not (directly) relevant to this discussion.

[9] More accurately, the underlying shear g is known.

Since we work with relative efficiencies, a conclusive statement about robustness is not straightforward. We will therefore use robustness to indicate that an estimator is equally or more efficient than the Gaussian estimator in most or all cases.

4.3.2 Estimators

We have explored various alternatives for well known estimators, which are optimal under Gaussian assumptions, like the mean and variance. By definition, the mean $\hat{e}_\mu$, or μ_e, minimizes the variance of the residuals, which makes it a least squares estimator.

In general, optimization estimators are solutions $\hat{e}$ that minimize a loss function

$$S_{\hat{e}} = \sum_i \rho\,(e_i; \hat{e})\ , \tag{4.8}$$

such as $\rho = r_i^2 = (e_i - \hat{e})^2$ for the mean.

For this paper, we considered two other optimization estimators, the least absolute deviations estimator (LAD) and the biweight (BI) estimator, and an ordering estimator, namely convex hull peeling (CHP). In section 4.3.3, we describe Fourier mode fitting (FM), using a LAD regression approach.

Least absolute deviations

LAD is an optimization approach where the loss function to be minimized is the sum of the absolute deviations, instead of the commonly used least squares minimization:

$$S_{\mathrm{LAD}} = \sum_i |r_i|. \tag{4.9}$$

In the one dimensional case, this is the median. In more than one dimension, we talk about the marginal median, when in each dimension the median is taken independently, or the spatial median, when minimizing the sum of the distances of measurements to a point. In many practical cases[10], the spatial median is unique, contrary to the marginal median, which can have multiple solutions. This is one of the reasons we used the spatial median throughout the rest of the paper.

Another reason is that e_1 and e_2 should not be seen as independent parameters of the shape. An ellipticity is defined by an absolute elongation $|e|$ and a position angle θ. The latter is defined within the context of a chosen frame of reference and therefore so are e_1 and e_2. In other words, using the marginal median would introduce an artificial anisotropy, as can be seen in Figure 4.4.

In concreto, for a set of $(e_{i,1}, e_{i,2})$ measurements, the mean $\hat{e}_\mu$ as an estimator for the net reduced shear $g_1 + ig_2$ minimizes the squared residuals

$$S_\mu = \sum_i (e_{i,1} - g_1)^2 + (e_{i,2} - g_2)^2. \tag{4.10}$$

A LAD estimate minimizes the absolute residuals,

$$S_{\mathrm{LAD}} = \sum_i \sqrt{(e_{i,1} - g_1)^2 + (e_{i,2} - g_2)^2}\,, \tag{4.11}$$

[10] Formally speaking: when the norm is strictly convex.

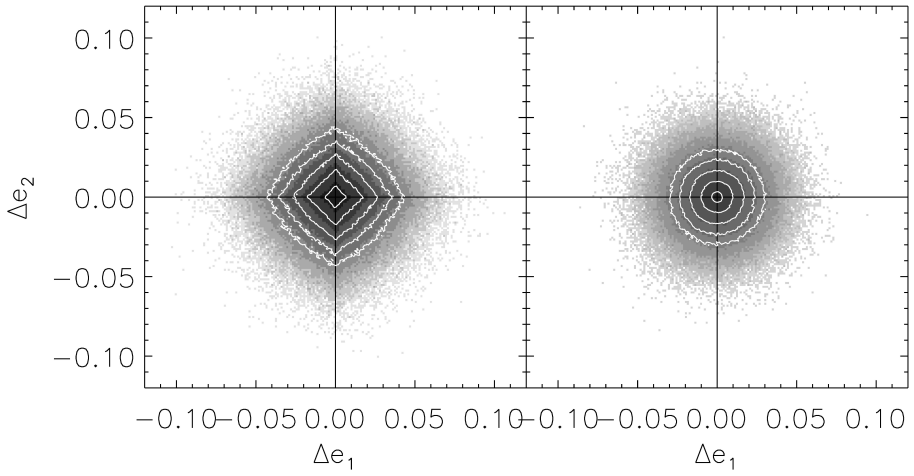


Figure 4.4: Comparison of the marginal median (left) to the spatial median, or LAD estimation (right). Plotted are the estimation biases Δe for 10^6 simulation runs, shown as a density in grayscale. Overplotted are arbitrary contours of increasing density (equal in both plots), to highlight the anisotropy in $P(\Delta e)$. Since e_1 and e_2 depend on the choice of reference frame, the marginal median introduces an artificial anisotropy. For the LAD estimations, the residual distances $|r_i|$ do not depend on the choice of reference frame.

which reduces the effect of outliers on the estimate. In one dimension, the LAD estimate arises as the central value maximum likelihood estimator of the Laplace distribution, which has a central cusp and more slowly declining tails.

There is no general analytic solution for LAD optimization. LAD can however be formulated as a linear optimization problem for which several iterative methods exist (e.g., simplex-based methods, Barrodale & Roberts 1973). In practical weak shear analyses, convergence is generally rapid.

The biweight

An alternative optimization approach is a bi-square weighted loss function (Beaton & Tukey 1974), called the biweight for short, given by

$$\nabla S_{\mathrm{BI}} = \sum_i r_i \left(1 - \left(\frac{r_i}{k}\right)^2\right)^2 = 0 \quad \text{for} \quad |r_i| < k\,, \tag{4.12}$$

where $r_i = (e_i - \hat{e})$ are again the residuals and k is a tuning parameter, usually determined by (an estimate of) the scale of the measured distribution.

A robust choice for k is the median absolute deviation (MAD), setting $k = c \cdot \mathrm{MAD}$, where $c = 6.0$ is optimal for estimation of location for a broad range of distributions (Mosteller & Tukey 1977). A common approach is iteratively correcting an initial estimate M_0 by the

normalized sum in Equation 4.12:

$$M_{n+1} = M_n + \frac{\sum_i r_{i,n}\left(1 - \left(\frac{r_{i,n}}{k}\right)^2\right)^2}{\sum_i \left(1 - \left(\frac{r_{i,n}}{k}\right)^2\right)^2}, \tag{4.13}$$

which can be interpreted as a normalized weighting of the residuals. In this case, the weight of a certain measurement increases toward the (current) central estimate, which makes this estimator a useful complement to the mean and LAD estimators.

In turn, a robust choice for M_0 is the (spatial) median. Note that measurements with residuals $|r_{i,n}| \geq k$ have effectively zero weight, although these points are not 'clipped' from the sample, since the residuals can change with each iteration. Convergence usually requires few iterations.

Convex hull peeling

The convex hull of a set of points X in $\mathbb{R}^n$ can be defined as the intersection of all convex sets in $\mathbb{R}^n$ that contain X. Informally put, the convex hull is the smallest subset of points that 'surrounds' the rest of the set (see Figure 4.5).

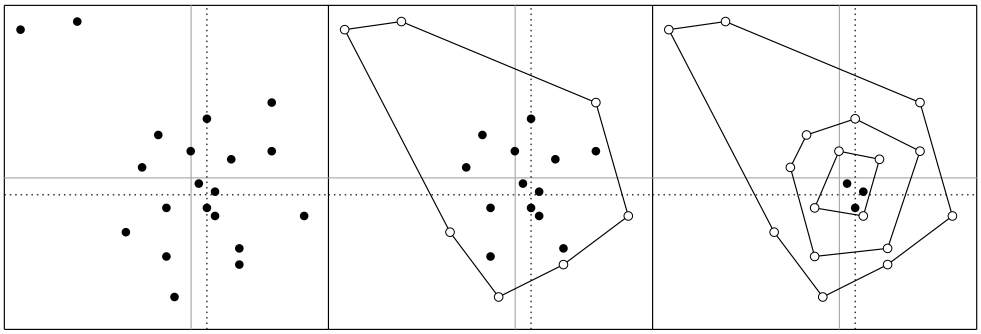


Figure 4.5: The method of CHP. The left panel shows a scatter plot with two outliers. The arithmetic mean is shown as a gray, solid line and the dotted line represents the mean without the two outliers. The middle panel shows the convex hull of the set of points, which is then removed from the set. The right panel shows the final result after repeating the process, until the final set of points is equal to its own convex hull.

There exist various algorithms for determining the convex hull (e.g., Preparata & Shamos 1985). For this paper, we used Delaunay triangulation based on the divide-and-conquer method (Lee & Schachter 1980).

In the process of CHP, the convex hull of a set of data points is determined and subsequently 'peeled' from the set, after which the process is repeated (Figure 4.5). When the remaining set of points is equal to its own convex hull, the final estimate is determined from these points, for example using the mean or LAD. This makes CHP an ordering approach, much like obtaining the familiar median for the one-dimensional case by sorting the data,

instead of optimization[11]. Among other aspects, it shares the resistance of the median against outliers.

In this paper, we are interested in the use of CHP in the two-dimensional case of (e_1, e_2) measurements in the complex plane, but CHP can be used in higher dimensions as well (see e.g., Lee 2007, for applications to SDSS quasar data).

Weighting and collinearity

When using real data, a weighting of ellipticity measurements is necessary to avoid or mitigate effects, such as noise or intrinsic size and ellipticity, that would confuse or bias the estimation of the underlying shear. For LAD and biweight optimization, weighting schemes are readily introduced, analogous to the weighted mean. For CHP, we suggest a possible weighting scheme, analogous to the one-dimensional weighted median, as follows.

The convex hull comprises a set of points in the $(e_{i,1}, e_{i,2})$-plane, with w_i the associated weights, given by the measurement pipeline. The minimum weight on the convex hull is then subtracted from these weights, after which all points with updated weight $w_i = 0$ are peeled from the sample. Note that this removes at least one point per iteration, but can lead to point-by-point peeling and large computation times. A solution with lesser precision but increased speed would be given by binning the weights in discrete steps.

We also note a possible collinearity problem of multiple ellipticity measurements with finite precision coinciding. In that case, triangulation has no solution. By combining these points into one measurement by combining the weights, this problem is resolved.

4.3.3 Fourier mode fitting

One can model a signal, in our case a varying ellipticity, over a one-dimensional range $-L < x < L$, writing that signal as a linear superposition of waves, or (Fourier) modes, $A_n \cdot \cos(k_n x \pm \phi_n)$, where A_n and ϕ_n are the amplitude and phase of the signal mode respectively, and $k_n \equiv \frac{n\pi}{L}$ are the wave numbers of the modes, showing the periodicity over the range $2L$.

It is useful to rewrite this model linearly in its coefficients $a_n \cdot \cos(k_n x) + b_n \cdot \sin(k_n x)$, where amplitude and the phase are now given by $A_n^2 = a_n^2 + b_n^2$ and via $\frac{b_n}{a_n} = \tan(\phi_n)$.

This one-dimensional model is readily extended to two dimensions, by considering that each coefficient depends similarly on y. This gives us $\alpha_{mn;\pm} = \cos(k_m x \pm l_n y)$ and $\beta_{mn;\pm} = \cos(k_m x \pm l_n y)$, or

$$\begin{aligned} e(x,y) = \sum_{m,n} \quad & a_{mn} \cos(k_m x) \cos(l_n y) \\ + \quad & b_{mn} \cos(k_m x) \sin(l_n y) \\ + \quad & c_{mn} \sin(k_m x) \cos(l_n y) \\ + \quad & d_{mn} \sin(k_m x) \sin(l_n y) \ , \end{aligned} \tag{4.14}$$

where the wave numbers k_m and l_n represent the spatial frequencies in the x and y directions, respectively. In two dimensions, we make a terminological distinction between a full Fourier mode, as given by Equation 4.14, and the individual waves comprising it. The amplitude of the fluctuations in ellipticity are now given for each mode in m, n by $a_{mn}^2 + b_{mn}^2 + c_{mn}^2 + d_{mn}^2$.

[11] Indeed, in one dimension, both approaches to the median are the same.

This linear model is fitted in a relatively straightforward manner to a sample of measured or simulated ellipticities. In the absence of noise and for a well-behaved field of view, each wave component of a Fourier mode is independent and can be fitted separately. We will discuss the effect of noise in Section 4.5.2.

Applying statistics

To apply these statistics to a shear field consisting of discrete Fourier modes, which by construction is centered around $e = 0$, the ellipticity measurements should be properly weighted by the model of the Fourier mode under consideration. We considered the information carried by an ellipticity measurement, which is proportional to the value of the fitted model M, where $M(x, y)$ can for instance be a single wave like $\cos(k_m x)\cos(l_n y)$, or a full mode.

Measurements close to the nodes of a wave carry the least information, whereas measurement close to extrema, or antinodes, carry the most amplitude information. We considered that each ellipticity measurement e_i theoretically infers an estimate of the amplitude A, where $A \in \{a_{mn}, b_{mn}, c_{mn}, d_{mn}\}$ by $\hat{A}_i = e_i \cdot M^{-1}$. In the case of Gaussian variations around the model, that is, measurement error distribution, the information scales as the inverse variance of that distribution, and therefore as the square of the model:

$$\hat{A} = \frac{\sum M^2 \cdot \frac{e}{M}}{\sum M^2} = \frac{\sum M \cdot e}{\sum M^2}, \tag{4.15}$$

where we recover the well known analytic LSQ form. This can be seen as an inverse variance weighting based on the model-to-noise ratio. For different error distributions, one can allow a general scaling of the information with the model by M^n, and therefore

$$\hat{A} = \frac{\sum M^{n-1} \cdot e}{\sum M^n}. \tag{4.16}$$

For application with our proposed weighting scheme for CHP, it is instructive to view the multiplication by weights as shifting the data points, so the central data point(s) or CHP value matches the amplitude to be estimated. For this purpose, it is practical to write Equation 4.16 as

$$\hat{A} = \frac{\sum |M|^{n-1} \cdot \mathrm{sgn}(M) \cdot e}{\sum |M|^{n-1}} \cdot \frac{\sum |M|^{n-1}}{\sum M^n}, \tag{4.17}$$

where the (e_1, e_2) data points are first shifted by $\mathrm{sgn}(M)$, and then weighed by $|M|$, before the weighed estimate is normalized as usual. We show this in Figure 4.6, where we plot the (e_1, e_2) values, the same (e_1, e_2) points shifted by $\mathrm{sgn}(M)$, with in this case $M = \cos(k_m x)\cos(l_n y)$, and then the associated distribution of the weights $|M|$ over the complex ellipticity plane as normalized 2D histograms.

4.4 Simulations and data

For this paper, we tested various forms of $P(e)$, using samples of random ellipticities, assumed to be centered around zero, which we sheared by Eq. 4.5. We have used several approaches to obtaining these samples.

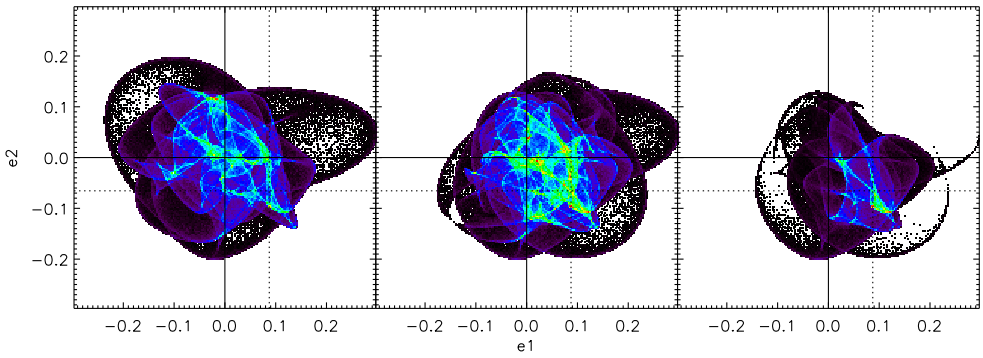


Figure 4.6: Ellipticity distribution of a superposition of 9 Fourier modes over the complex (e_1, e_2) plane, where we show how we recover a single (e_1, e_2) amplitude, indicated by the dotted lines. Left: the Fourier modes, centered around $(0, 0)$ (normalized number counts). Middle: the (e_1, e_2) points shifted by $\mathrm{sgn}(M)$, with M the model of the amplitude (normalized number counts). Right: the resulting distribution of weights over (e_1, e_2), showing a shift toward the amplitude under consideration.

Firstly, we simulated a uniform q distribution, which seems to fit real data adequately (e.g., Lambas et al. 1992, Rodríguez & Padilla 2013), without assuming any physical mechanism that would explain this distribution.

Secondly, we modeled background galaxies as randomly orientated triaxial ellipsoids, and derived the projected ellipticities following Stark (1977), using axis ratio distributions fitted to observed ellipticity distributions (Lambas et al. 1992).

In both cases, we compared our results to samples with added Gaussian noise, using real data shape measurement error distributions to simulate the effect of noise.

Thirdly, we sampled real data, using shape measurement catalogs from weak lensing observations.

Finally, we compared these various ellipticity distributions and the results from each estimator to results in case when $P(e)$ follows a Gaussian distribution. We examined the behavior of bias and efficiency of each estimator under the effect of noise, the input shear and the sample size.

4.4.1 Simulated ellipticity distributions

Uniform samples

We produced random samples with a uniform q-distribution, as an ideal version of the observed distribution of spiral galaxies in for example Lambas et al. (1992), Rodríguez & Padilla (2013), henceforth referred to as a uniform sample. We used an axis ratio cut-off of $q \approx 0.2$ to account for a finite galaxy thickness, following Lambas et al. (1992), which gives rise to standard deviations in each ellipticity component of $\sigma_e \approx 0.25$, comparable to the samples drawn from data.

The resulting axis ratio and ellipticity distributions are shown in Figure 4.7

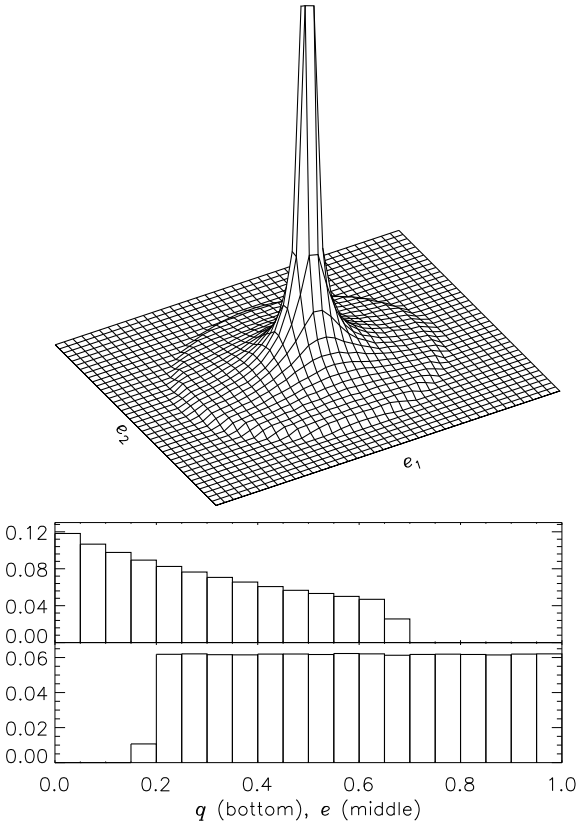


Figure 4.7: Ellipticity distributions for a uniform axis ratio distribution. Top: a 2D histogram of ellipticities. Middle: histogram of the absolute ellipticity $|e|$. Bottom: histogram of the ellipse axis ratio q. A cut-off near $q \approx 0.2$ is suggested by observations and produces standard deviations in each ellipticity component of $\sigma_e \approx 0.25$, comparable to most survey shape measurement catalogs.

Projected ellipsoids

A triaxial ellipsoid with axes $\tilde{a} \geq \tilde{b} \geq \tilde{c} \geq 0$ can be described by

$$(cx)^2 + \left(\frac{cy}{b}\right)^2 + (z)^2 = \text{constant}\,, \tag{4.18}$$

with $b = \tilde{b}/\tilde{a}$ and $c = \tilde{c}/\tilde{a}$. As given by Stark (1977), such an ellipsoid is seen as an ellipse in projection, given by

$$(j/f)x'^2 + 2(k/f)x'y' + (l/f)y'^2 = \text{constant}\,, \tag{4.19}$$

where (x', y') are the coordinates in the projection plane and

$$f \equiv c^2 \sin^2\theta \sin^2\varphi + (c/b)^2 \sin^2\theta \cos^2\varphi + \cos^2\theta\,, \tag{4.20a}$$

$$j \equiv c^2(c/b)^2 \sin^2\theta + c^2 \cos^2\varphi \cos^2\theta + (c/b)^2 \sin^2\varphi \cos^2\theta\,, \tag{4.20b}$$

$$k \equiv ((c/b)^2 - c^2) \sin\varphi \cos\varphi \cos\theta\,, \tag{4.20c}$$

$$l \equiv c^2 \sin^2\varphi + (c/b)^2 \cos^2\varphi\,, \tag{4.20d}$$

with φ and θ the first two orientation angles of the ellipsoid.

For simulations of projected ellipsoids, we assumed Gaussian distributions for b and c, following Lambas et al. (1992). For elliptical galaxies, we used $b = 0.95$ and $c = 0.55$ with standard deviations $\sigma_b = 0.35$ and $\sigma_c = 0.2$. For disk galaxies, we used $b = 1.00$ and $c = 0.25$ with standard deviations $\sigma_b = 0.13$ and $\sigma_c = 0.12$.

The axis ratio was then recovered via

$$q = \sqrt{\frac{j + l - \sqrt{(j-l)^2 + 4k^2}}{j + l + \sqrt{(j-l)^2 + 4k^2}}}\,, \tag{4.21}$$

and the ellipticity through Equation 4.3. The orientation angles of the ellipsoids were randomly distributed. The resulting axis ratio and ellipticity distributions are shown in Figure 4.8.

We will refer to these simulated samples as disk and elliptical samples. We also used combined samples with a disk to elliptical ratio derived from the CFHTLenS catalog. (See Table 4.1.)

4.4.2 Data: CFHTLenS

We used data from Canada-France-Hawaii Telescope Lensing Survey (CFHTLenS, Heymans et al. 2012b). The CFHTLenS survey analysis combined weak lensing data processing with THELI (Erben et al. 2005, 2009, 2013), shear measurement with *lens*fit (Miller et al. 2007, 2013, Kitching et al. 2008), and Bayesian photometric redshift measurement (BPZ, Benítez 2000, Coe et al. 2006) with PSF-matched photometry (Hildebrandt et al. 2012). A full systematic error analysis of the shear measurements in combination with the photometric redshifts is presented in Heymans et al. (2012b), with additional error analyses of the photometric redshift measurements presented in Benjamin et al. (2013).

For our analyses, we selected 4.2 million objects that are well determined and resolved (*lens*fit fitclass = 0, non-zero *lens*fit weight, star_flag = 0, `CLASS_STAR` ≤ 0.5). We excluded objects that lie within a mask, with the exception of large, conservative masks around relatively faint stars and stellar haloes (MASK ≤ 1, see Erben et al. 2013).

The CFHTLenS shape catalog is not an exact representation of the ellipticity distribution of the observed galaxy population, as it includes measurement noise present in any real data set. Selecting sources on *lens*fit weight w or signal-to-noise ratio $\nu_{\rm SN}$ could on the other hand introduce selection biases in the galaxy population we wanted to to study. We decided to use two sets of sources: the complete set, described above, to optimally sample the complete source population, and a conservative subset with $w \geq 15$ and $\nu_{\rm SN} \geq 20$, to reduce the uncertainty in observed ellipticity, at the possible cost of a bias in the selection.

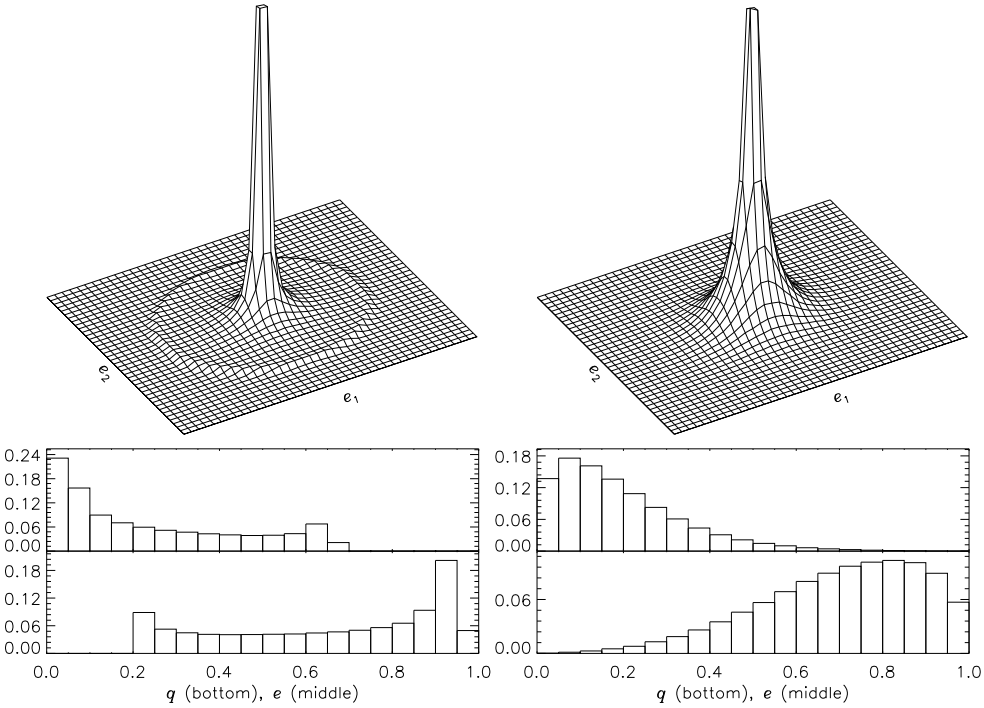


Figure 4.8: Ellipticity and axis ratio distributions for distribution of projected ellipsoids. Left: disk galaxies. Right: elliptical galaxies. Top: a 2D histogram of ellipticities. Note that the ring-like feature in the left panel is the result of a finite disk thickness. Middle: histogram of the absolute ellipticity $|e|$. Bottom: histogram of the ellipse axis ratio q.

For both sets, we split these sources by BPZ spectral type into red ($T_{\rm BPZ} < 1.5$) and blue ($1.5 < T_{\rm BPZ} < 3.95$) galaxies, with a further division between Sbc ($1.5 < T_{\rm BPZ} < 2.5$) and Scd ($2.5 < T_{\rm BPZ} < 3.95$). We found that our conservative selection reduced the number of galaxies to roughly 25%, almost independent of spectral bin for $T_{\rm BPZ} < 3.1$. For higher spectral types, the subset decreased linearly to roughly 10% for the highest spectral bin, which was an indication that our selection did indeed introduce a modest sample bias.

Table 4.1 gives an overview of the selected CFHTLenS data, while Figure 4.9 shows the respective distributions.

We drew random subsets from the selected CFHTLenS ellipticities, which we then sheared by Eq. 4.5. This introduced the implicit assumption that, after the bias corrections described in Heymans et al. (2012b) and Miller et al. (2013), the central ellipticity was zero, and these random subsets were approximately drawn from an unsheared, noise-free background galaxy population.

4.4.3 Simulated noise

In any realistic shape measurement catalog, ellipticities not only have shape noise due to a finite intrinsic distribution, but suffer from measurement uncertainties as well. For this

Table 4.1: Overview of the CFHTLenS data used. Column 1 gives the division between BPZ spectral type (red: $T_{\rm BPZ} < 1.5$, blue: $1.5 < T_{\rm BPZ} < 3.95$, Sbc: $1.5 < T_{\rm BPZ} < 2.5$, Scd: $2.5 < T_{\rm BPZ} < 3.95$). Column 2 gives the number N of objects selected. Column 3 gives the 1D Gaussian ellipticity standard deviation σ_e, using both ellipticity components after bias correction. In parentheses, we give N and σ_e for sources with $w \geq 15$ and $\nu_{\rm SN} \geq 20$.

Color	N		σ_e	
All	4216334	(912828)	0.286	(0.242)
Red	553633	(151939)	0.267	(0.242)
Blue	3662701	(760889)	0.289	(0.242)
Sbc	870295	(219929)	0.294	(0.262)
Scd	2792406	(540960)	0.288	(0.232)

reason, we wanted to study the effect of noise or our simulated, noiseless ellipticity samples.

Measurement uncertainties depend primarily on pixel noise and therefore vary with image size and brightness. This means that errors on the ellipticities are not drawn from a single distribution. To mimic the effect of a skewed composite error distribution for our simulated samples, we randomly sampled the CFHTLenS weight w.

Miller et al. (2013) calculated an approximately inverse-variance weight using the width of the ellipticity likelihood surface by

$$w = \left[\frac{\sigma_e^2 e_{\rm max}^2}{e_{\rm max}^2 + 2\sigma_e^2} + \sigma_{\rm pop}^2 \right]^{-1} , \tag{4.22}$$

where σ_e^2 is the variance in ellipticity of the likelihood surface, $\sigma_{\rm pop}^2$ is the ellipticity variance of the galaxy population, and $e_{\rm max}$ is a maximum ellipticity, to reflect a finite edge-on disk thickness.

Using $e_{\rm max} = 0.804$ from Miller et al. (2013) and refining $\sigma_{\rm pop}^2 \approx 0.242$ using the CFHTLenS catalog itself[12], we obtained a distribution in ellipticity variance σ_e^2 for each w. From this, we produced noise by assuming a Gaussian distribution with the ellipticity as mean and σ_e^2 as variance.

Estimation of errors

To assess errors on bias and efficiency from our simulations, we simply divided our simulations randomly in smaller subsets and determine the statistical variations, assuming t-distributions. While this approach may seem to lack finesse compared to a full bootstrap, the significance of our results is high enough for a proof of concept.

[12]Miller et al. (2013) cite $\sigma_{\rm pop}^2 = 0.255$ as prior, but this would lead to a negative σ_e^2 for the maximum weight in the CFHTLenS catalog.

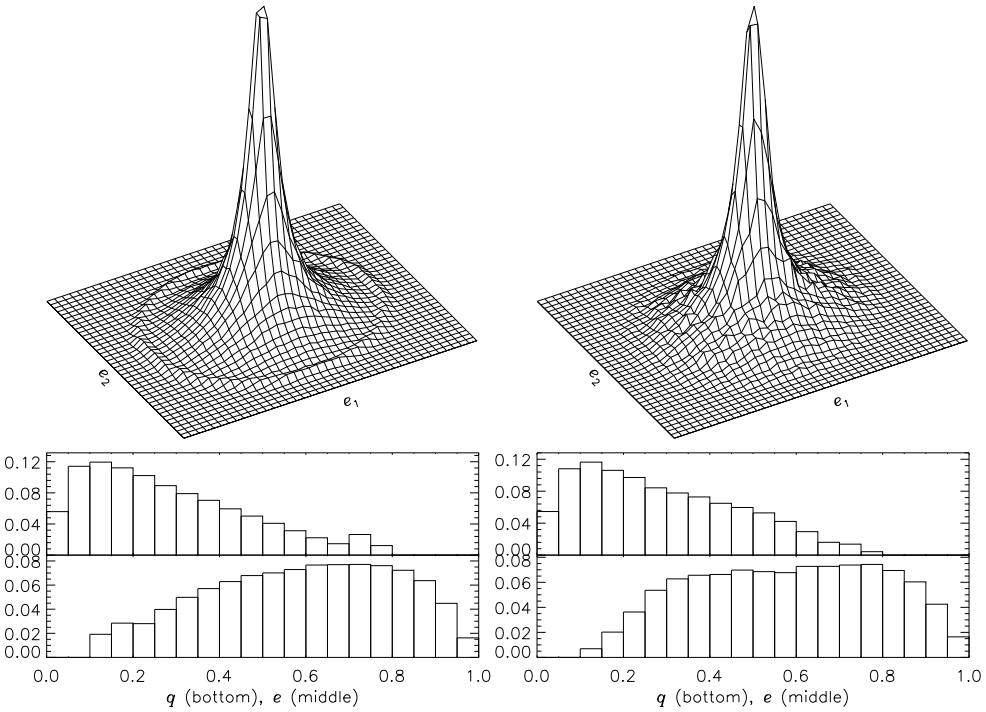


Figure 4.9: Ellipticity and axis ratio distributions for CFHTLenS blue (left), red (right). Top: a 2D histogram of ellipticities. Note that the ring-like feature at $e \approx 0.8$ is due to noisy outliers forced to a maximum e by the shape measurement pipeline, but see also Figure 4.8. Middle: histogram of the absolute ellipticity $|e|$. Bottom: histogram of the ellipse axis ratio q.

4.5 Results

4.5.1 Central value estimation

For each sample type, we produced 10^4 random samples of 100 ellipticities, which we distorted by an absolute reduced shear of $g = 0.2$, and determined relative efficiencies and possible biases. We then assessed the effect of varying the shear and the sample size.

Asymmetry and bias

Ideally, an estimator should be unbiased in the absence of noise. For the mean, this is the case (Seitz & Schneider 1997), but since the effect of shear on intrinsic ellipticities in non-linear, the resulting, observed ellipticity distribution $P(\tilde{e})$ is asymmetric, or skewed, which can lead to mean-biases for various estimators.

In Figure 4.10, we show this effect on the CHP estimator for $g = 0.3$ in two directions. The distribution of the CHP estimator is clearly skewed, as shown by the convex hulls plotted, when the coverage within the current hull is equal to approximately[13] 38.3%, 68.3%, 86.6%,

[13] CHP is a discrete and not a continuous process, but this effect is negligible for 10^4 estimates.

and 95.4%. We note that this leads to a mean-biasedness, according to definition, but the center of the estimator distribution $P(\hat{e}_{\rm CHP})$ seems significantly less biased. In other words, the CHP estimator seems 'CHP-unbiased'.

A solution to this skewness in the estimator distribution, in the absence of noise, is iteratively improving estimates by correcting the observed ellipticities $P(\tilde{e})$ by the estimated shear, using Equation 4.5, and then determining the updated residuals. We call this process of iteratively correcting the sample by the current estimate 'de-shearing' (or 'de-g'). Figure 4.10 shows how this symmetrized the estimator distribution $P(\hat{e})$, and slightly improved the efficiency as well (see section 4.5.1). The latter seemed to be the case even for the mean $\hat{e}_\mu$ as estimator, but the difference was not statistically significant.

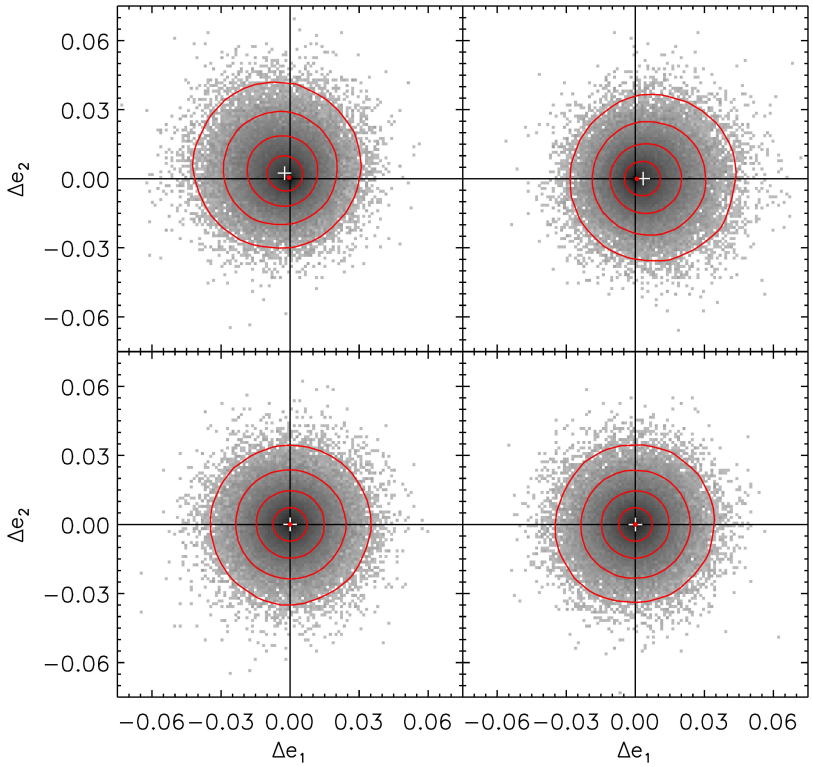


Figure 4.10: The skewed $\hat{e}_{\rm CHP}$ distribution as an example of the effect of asymmetry in a sheared ellipticity distribution. Plotted are the estimation biases Δe for 10^5 simulation runs, shown as a density in grayscale. Over-plotted are the convex hulls at approximately 38.3%, 68.3%, 86.6%, and 95.4% coverage. The mean of the distribution is shown as a white plus. Top: estimation biases $\Delta e_{\rm CHP}$ for samples with an underlying shear of $g = -0.21 + 0.21i$ (left) and $g = 0.3$ (right). Note that these estimator distributions are effectively mean-biased, because they are skewed, but still centered around $\Delta e = 0$, as indicated by the CHP estimation of the distributions. Bottom: $\Delta e_{\rm CHP}$ for the same samples, after iteratively de-shearing the samples until the final CHP estimate vanishes. These iterations remove asymptotic mean-bias and increase efficiency.

In presence of noise, the mean is a biased estimator (Melchior & Viola 2012). Given that

in reality systematic noise is always present, a form of bias is unavoidable, since the noise distribution is different[14] from the (skewed) ellipticity distribution (See Figure 4.3). This means that our method of de-shearing would introduce a noise bias for precisely the same reason, since we would not properly correct the asymmetry in the distribution.

We compared the results for simulated projected ellipsoids with and without simulated noise in Figure 4.11 to assess the effect. In the appendix, we quantified the observed multiplicative bias in the form

$$e_{\rm fit} = (1 + m)e_{\rm in}\,, \tag{4.23}$$

where e stands for $e_{1,2}$, and summarize the results in Table 4.2.

Without de-shearing, only the mean is a mean-unbiased estimator. We noted that all estimation methods could be made mean-unbiased in the noise-free case, when including de-shearing, but showed a mean-bias in the presence of noise, as expected. For the biweight estimator $\hat{e}_{\rm BI}$, this was (within statistical significance) the same bias as for the mean. For the LAD and CHP estimators $\hat{e}_{\rm LAD}$ and $\hat{e}_{\rm CHP}$, the biases were significantly reduced, up to $\sim 30\%$, to below percent level for realistic weak shear.

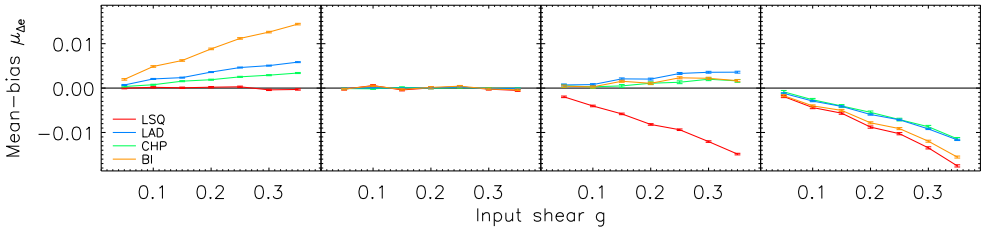


Figure 4.11: Estimator mean-bias as a function of input shear for realistic combinations of simulated disk and elliptical samples, using projected ellipsoids. From left to right: all estimators without noise, without noise after iteratively de-shearing the samples, all estimators with noise, and with noise and after iteratively de-shearing the samples. Color coding: $\hat{e}_{\mu}$ (red), $\hat{e}_{\rm LAD}$ (blue), $\hat{e}_{\rm CHP}$ (green) and $\hat{e}_{\rm BI}$ (yellow).

This decrease in bias can be explained by realizing that the observed, sheared ellipticity distribution is skewed, but the location of the central peak of intrinsically round background sources is still an unbiased estimator of the underlying shear (which can be deduced from Equation 4.5 and Figure 4.3). It is the bias in determining the location of this peak that introduces the bias in the shear estimate. Likewise, the effect of noise changes the observed ellipticity distribution, but does not affect the location of that peak. Estimators that are more sensitive to a central cusp or peak in the distribution and less to high ellipticities in the tail, such as $\hat{e}_{\rm LAD}$ and $\hat{e}_{\rm CHP}$, will therefore introduce a lower mean-bias.

We compared these results to the mean-bias in the upper panels of Figure 4.10 and the observation that the central peak of the estimator distribution is in fact located at $\Delta e \approx 0$. We found that the mean-bias arose due to the asymmetry in the estimator distribution and the CHP-bias vanished, unaffected by noise.

[14] Intrinsically, the effect of noise is symmetric, but the effect on a sample of sheared ellipticities depends on the shape measurement pipeline, as noted in Melchior & Viola (2012).

Estimator efficiencies

In Tables 4.3, 4.4, and 4.5 in the appendix, we summarized the full results for the relative efficiencies of each estimator. We applied de-shearing and note that this improves the efficiencies marginally at a similar marginal cost to the bias. We determined relative efficiencies for coverages of 25%, 50%, and 75%, corresponding to the MAD and the first and third quartiles, and 38.3%, 68.3%, 86.6%, and 95.4%, which would correspond to steps of 0.5σ in case of a Gaussian distribution with variance σ^2.

In Figure 4.12, we plot these results for a few distributions, namely Gaussian, uniform q, a combination of disk and elliptical projections and the conservative CFHTLenS catalog samples. We also plot the results for the samples with added noise and the full CFHTLenS samples.

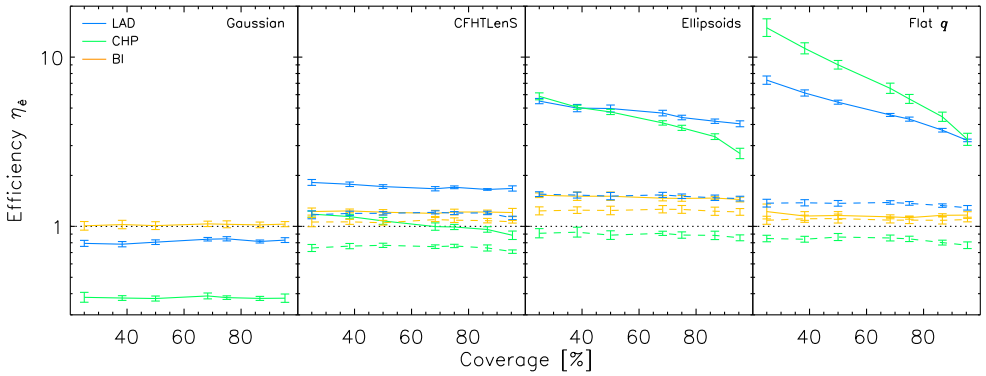


Figure 4.12: Relative efficiencies each estimator plotted at different coverages. From left to right: Relative efficiencies in case of a Gaussian $P(e)$ distribution, the CFHTLenS catalog $P(e)$ distribution, a combination of disk and elliptical distributions using projected ellipsoids, and a uniform q distribution. Color coding: relative efficiencies for $\hat{e}_{\rm LAD}$ (blue), $\hat{e}_{\rm BI}$ (yellow), and $\hat{e}_{\rm CHP}$ (green). Solid lines: simulated samples without noise or using the CFHTLenS conservative subset. Dashed lines: including noise or using the complete CFHTLenS set.

Not all estimators reached asymptotic normality. Especially CHP converged slower toward normality in the tails of the distribution, that is, at higher coverage. For LAD, this is noticeable mostly for the uniform q distribution.

The biweight is the most robust, as its relative efficiency doesn't vary much across distributions. The biweight relative efficiency is however quite low, which means that this estimator offers little improvement. Even when $P(e)$ follows a Gaussian distribution, η_{BI} is not significantly better or worse than the traditional mean.

Our results show that estimator efficiency is independent of input shear. This is the case, when we define the individual estimate biases similarly to the residuals, as noted in Section 4.3.1, that is, not as the difference $\hat{e} - g$, but as the extra shear needed over the input shear g to reach this difference, as determined by Equation 4.5:

$$\Delta e = \frac{\hat{e} - g}{1 - g^* \hat{e}}\,, \tag{4.24}$$

with g^* the complex conjugate of the input shear g of the simulations. Using that definition,

this independence is demonstrated Figure 4.13.

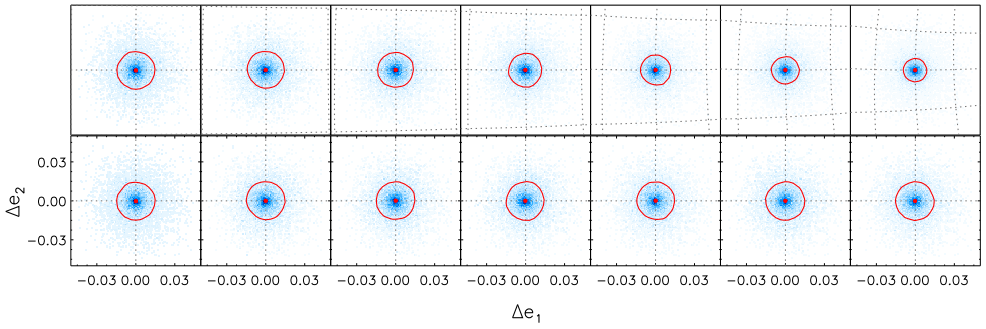


Figure 4.13: Efficiency for an arbitrary estimator and sample type versus input shear, ranging from $g = 0.05$ (left) to $g = 0.35$. Upper: simple difference $\hat{e} - g$ between estimates and input shear, with $s_{68.3}$ over-plotted. Lower: Δe, as defined by Equation 4.24, with $s_{68.3}$ over-plotted.

As an aside: since the mean of the CHP estimator is displaced from the center, this necessarily increases the distribution scale. A more proper way to compare the scale with symmetric distributions would be comparing the surface within the convex hull at a certain coverage, as s^2 is a measure of the (circular) surface around $\hat{e}$ inside that scale. In this sense, efficiency is a figure of merit. We have not done so in this paper, which means the $\eta_{\rm CHP}$ are slightly underestimated, but not significantly.

In Figure 4.14, we show the results for different samples sizes. In Table 4.6 in the appendix, we summarize the quantitative results. In the limit of very small sample sizes, the difference between the various estimators is expected to vanish. We note that a potential improvement over the mean estimator remains even for a sample size of $N = 10$.

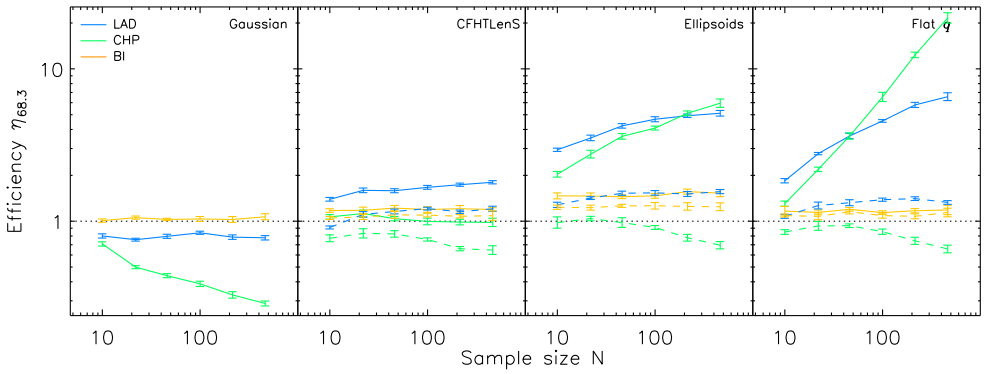


Figure 4.14: Relative efficiencies $\eta_{68.3}$ plotted against sample sizes. From left to right: Relative efficiencies in case of a Gaussian $P(e)$ distribution, the CFHTLenS catalog $P(e)$ distribution, a combination of disk and elliptical distributions using projected ellipsoids, and a uniform q distribution. Color coding: relative efficiencies for $\hat{e}_{\rm LAD}$ (blue), $\hat{e}_{\rm BI}$ (yellow), and $\hat{e}_{\rm CHP}$ (green). Solid lines: simulated samples without noise or using the CFHTLenS conservative subset. Dashed lines: including noise or using the complete CFHTLenS set.

4.5.2 Fourier mode fitting

For samples of a combination of disk and elliptical distributions using projected ellipsoids, we produced 10^3 random square fields with 10^3 simulated ellipticities. For comparison, the average number of selected sources in a CFHTLenS field is roughly $2.5 \cdot 10^4$, ranging from 9525 to 37767, or $5.3 \cdot 10^3$, ranging from 2111 to 9525 for the more conservative sample.

Using Equation 4.5, we distorted these intrinsic ellipticities by the total shear pattern of one or more full modes (as defined in Equation 4.14), then applied simulated measurement noise (as described in Section 4.4.3) as a final step.

We fitted amplitudes per individual wave using LSQ, LAD and CHP, and per mode using LSQ and LAD by simultaneously fitting all four amplitudes. We then determined relative efficiencies and possible biases of the recovered amplitudes in the same way as in Section 4.5.1.

In Figure 4.15, we show the fitted shear field for a single realization, using in this case 10^4 simulated ellipticities. We fitted 16 different modes individually, using LSQ and LAD, and 64 individual amplitudes using LSQ, LAD, and CHP, and found the amplitude residuals, $\left(\mathcal{O}\left(10^{-3}\right)\right)$, to be two orders of magnitude less than the input values, which were constrained to $g \leq 0.25$ for peak values at positive interference. Residuals in $|e|$ for this realization varied between ±0.075 for LSQ,±0.066 for LAD and ±0.14 for CHP.

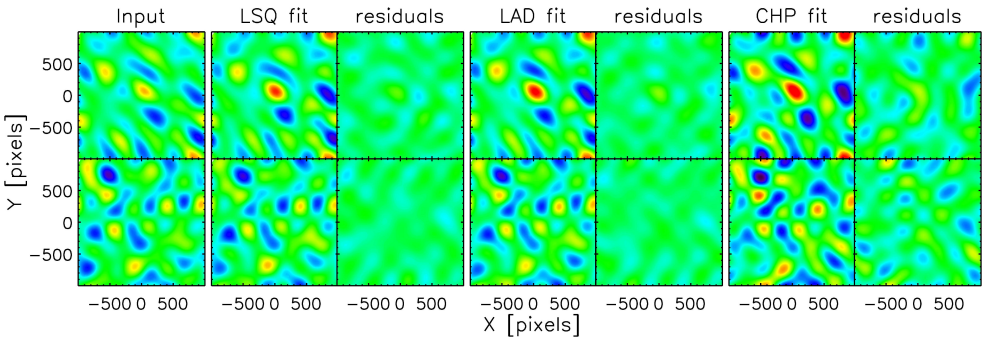


Figure 4.15: Shear field residuals when fitting 16 different modes, using simulated projected ellipsoids as intrinsic shapes, and including additional Gaussian noise. From left to right: input shear, LSQ fit and residuals, LAD fit and residuals, CHP it and residuals. Upper and lower row show e_1 and e_2 respectively. The color scale is the same in all plots for comparison and ranges between $-0.247 \leq e_{1,2} \leq 0.247$. Residuals for this realization vary between ±0.075,±0.066 and ±0.14, respectively.

Bias and efficiency

The results from Sections 4.5.1 and 4.5.1 carry over to estimates of Fourier amplitudes for LSQ and LAD. We found fitted values with standard deviations of the order of 10^{-3} for individual amplitudes. In Figure 4.16, we show the consistency of the fitted values.

Over plotted in Figure 4.16 are the best-fitting mean-bias, defined similar to Equation 4.23 as

$$a_{\rm est} = (1 + m)a_{\rm in}\,, \tag{4.25}$$

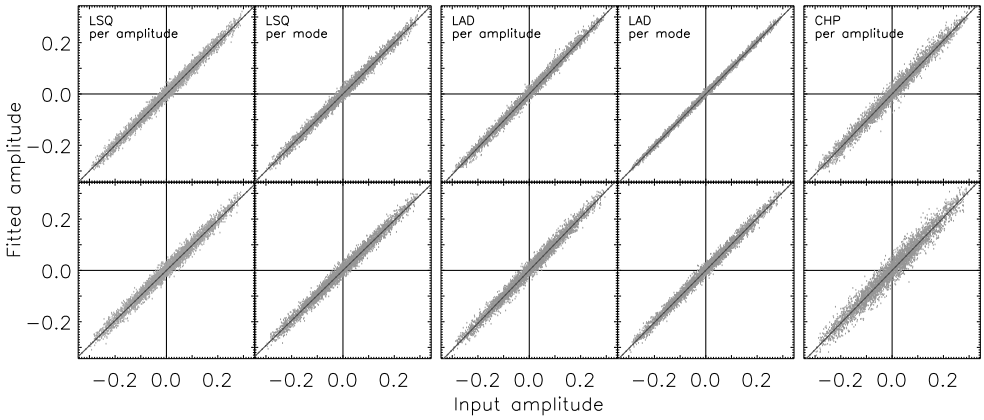


Figure 4.16: Consistency of the estimated Fourier amplitudes as defined in Equation 4.14. Plotted are the input amplitudes a_{mn}, b_{mn}, c_{mn}, d_{mn} versus their estimates (Section 4.3.3) for LSQ (left), LAD (middle) AND CHP (right). Top row: only simulated intrinsic shapes based on projected ellipsoids. Bottom row: the same, with added noise. Over plotted are the best-fitting mean-bias, as defined in Equation 4.25.

where a stands for a_{mn}, b_{mn}, c_{mn}, and d_{mn} as defined in Equation 4.14. The uncertainties are too small to be visible. In Table 4.7 in the appendix, we give the quantitative results.

Similar to the results shown in Figure 4.11, LSQ underestimates Fourier amplitudes by a few percent in the presence of noise. For LAD, we found an improvement on bias by $\sim 20\%$ in the presence of noise, when iteratively de-shearing the sample. Likewise, fitting for LAD without de-shearing slightly overestimated the amplitudes, again comparable to Figure 4.11.

We note that in this case, adding noise did not seem to have a significant effect when fitting per mode. In most cases, we did notice a significant increase in bias when fitting per single amplitude. We did not see a change in bias between LSQ per mode and per amplitude.

We also found a slightly higher relative efficiency of $\eta_{68.3} = 1.09 \pm 0.07$ for LSQ and $\eta_{68.3} = 1.47 \pm 0.09$ for LAD, when fitting per mode, with or without added noise. It is not surprising that a model with four parameters (amplitudes) fits the estimates better than a model with one parameter, but the difference of this effect between LSQ and LAD is noteworthy.

Since CHP doesn't fit a model to the data, but rather orders the (e_1, e_2) data points, there is no straightforward way to fit four amplitudes simultaneously with the necessary weighting (Section 4.3.2). We have not explored this option further in this paper.

The CHP estimator performs consistently, that is, convergent around the input values, but with a significant lower efficiency than for central value estimation of a cloud of (e_1, e_2) data points (Section 4.5.1). This is to be expected, since CHP is particularly sensitive to a (central) cusp in the distribution of data points. By shearing the intrinsic ellipticities by a model that varies over the field of view, as shown in Figure 4.6, this peak will be smeared out, decreasing the effectiveness of CHP.

In conclusion, CHP is consistently the most sensitive to the central cuspiness of a distribution. The results of this section do serve as a proof of concept for applying alternative statistics to an observed field of weak shear measurements.

4.6 Conclusions and summary

4.6.1 Optimal estimators

Our main conclusion is that to evaluate a statistical estimator, one must be willing to look beyond the canonical terms of mean-bias and the Gaussian variance as efficiency. We have shown that these commonly used meta-analysis instruments do not always properly reflect how well weak shear estimator values are constrained around the true underlying shear values.

By discussing the statistical definitions and observing the behavior of estimators for various ellipticity distributions, we have proposed ways of comparing various estimators motivated by statistical theory. The conclusions of that comparison are as follows:

Since the central peak of the intrinsic ellipticity distribution $P(e)$ is an unbiased tracer of the underlying shear, we find that the LAD and CHP estimators are less biased and more efficient than the standard mean.

When iteratively de-shearing the ellipticity sample by the estimated shear, the LAD estimator can reach a sub-percent bias for typical weak shear values, including noise. LAD is generally the most efficient of all estimators considered, potentially reducing uncertainties by more than 50% for samples simulated using a model of projected triaxial ellipsoids.

The CHP estimator is in terms of its mean-bias less affected by noise, as compared to the mean and, to a lesser extend, LAD. In fact, since the estimator distribution $P(\hat{e})$ is not symmetric, the actual center of that distribution, as opposed to the mean of that distribution, is unbiased in the presence of noise, within statistical significance. This makes CHP an important consideration, but it is less straightforward for adaptation for regression and requires careful assessment of uncertainties. Furthermore, CHP is computationally more demanding. In the presence of Gaussian noise, CHP is slightly less efficient than the mean (Figure 4.12, panels 3 and 4), but defining efficiency in terms of a figure of merit can reduce this drawback compared to the gain in bias, as proposed in Section 4.5.1.

4.6.2 Direct Fourier mode fitting

Applying different statistics to fitting individual Fourier modes to the shear field directly, we found results consistent with our previous conclusions.

We have shown that the Fourier amplitudes can be recovered with sub-percent accuracy and a minimal bias, which is an important proof of concept. Since the periodic variations in underlying shear effectively smooth the central peak of the intrinsic ellipticity distribution, the gain in efficiency is slightly less for LAD and significantly less for CHP. It is possible that an alternative to our weighting scheme for FMF with CHP could improve results. At this point, the method of CHP seems more applicable to samples of expected (roughly) constant shear, for example when measuring tangential shear around a gravitational lens candidate in bins of distance.

We have also shown that the shear field can be recovered reliably, with residuals an the order of magnitude less than the variations of the shear over the field of view for LSQ and LAD, using 10^3 sources, which is conservative compared to a typical single CFHTLenS field.

4.6.3 Future considerations and possible applications

We have discussed alternative statistics for inference of shear from samples of background sources with various intrinsic ellipticity distributions, proposing methods that could improve biases and uncertainties arising from the shape noise. It is important to consider our results within the broader context of other sources of systematics, as mentioned in our introduction.

Firstly, our results for shape noise assume trustworthy shape measurements, not only per source, but also considering the effect of systematics in the shape measurement pipelines on the reproduced ellipticity distribution as a whole: the recovery of a central peak, the distribution of outliers, among others. Examples are the effect of constraining ellipticities to a certain 'physical' maximum (e.g., $e_{\rm max} = 0.804$ for *lensfit*, Miller et al. 2013), as we see in Figure 4.2, or conversely, the unphysical outliers with $|e| > 1.0$ arising from dividing two noisy quantities (often when correcting for the point spread function, or PSF), affecting the tails of the distribution. Any features in the recovered shape distribution could affect bias and efficiency of the statistic used. Optimizing statistics will place more stringent demands on shape measurements than performing excellent 'on average'. Even methods that avoid individual shape measurements (Bernstein & Armstrong 2014), an ensemble inferred reduced shear could improve by considering the intrinsic shape distribution. Secondly, even with an accurately measured shape distribution, there will remain sources of systematic error in other steps of a cosmological analysis, as noted in our introduction. These effects still form a necessary part in a weak lensing analysis, but leave our statistical conclusions unaffected.

As survey sizes and image qualities increase, so will the demands on constraining systematic effects to a sub-dominant level, as described in for example Kuijken et al. (2015) and Mandelbaum et al. (2018) for the KiDS DR2 and HSC DR1, respectively. At the same time, it will be interesting to see measured ellipticity distributions converge as more sources are observed with higher signal-to-noise and measured with higher fidelity, due to increased depth of imaging, image quality and PSF control.

For now, we have given a proof of concept for alternative statistics in two cases: a sample of ellipticities with one underlying shear and the recovery of individual Fourier modes of the shear variation over a field of view. The first part has important applications when inferring a shear profile around lenses, both in recovering an accurate, less biased estimate and smaller error bars or confidence intervals. For the second part: since the amplitudes are well constrained by fitting individual Fourier modes, this provides a possible method toward estimation of the power spectrum. Furthermore, the shear field can be recovered in terms of its Fourier amplitudes, providing a powerful analytic model for mass reconstruction, without the need for smoothed gridding and incorporating variations in background source densities and estimated measurement uncertainties.

Acknowledgements

We thank the anonymous referee for useful comments and suggestions, which improved readability and reproducibility of this paper and its results. MS acknowledges support from the Netherlands Organization for Scientific Research (NWO).

This work is 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 (CNRS) of France, and the University of Hawaii. This research used the facilities of the Canadian Astronomy Data Centre operated by the National Research Council of Canada with the support of the Canadian Space Agency. CFHTLenS data processing was made possible thanks to significant computing support from the NSERC Research Tools and Instruments grant program.

4.A Bias estimations

In Table 4.2 are given the mean-bias and CHP-bias for each estimator, with and without de-shearing. This multiplicative bias m is defined by Equation 4.23 as $e_{\rm fit} = (1 + m)e_{\rm in}$.

4.B Efficiency estimations

In Tables 4.3, 4.4, and 4.5, we summarize the full results for the relative efficiencies of each estimator. We determine relative efficiencies for coverages of 25%, 50%, and 75%, corresponding to the MAD and the first and third quartiles, and 38.3%, 68.3%, 86.6%, and 95.4%, which would correspond to steps of 0.5σ in case of a Gaussian distribution with variance σ^2.

For easy reference, we also indicate how much the scale of the estimator distribution would improve, in percentages of the scale of the distribution of the mean estimator,

$$\Delta s_p = \frac{s_{\hat{e},p}}{s_\mu} - 1 \quad \text{(in \%)} \tag{4.26}$$

Since a higher efficiency means a smaller scale and therefore a more 'trustworthy' estimate, this is an intuitive, albeit rough indication of the change in error bars.

4.C Estimations from Fourier mode fitting

In Table 4.7 are given the mean-bias and efficiencies of the LSQ estimator, per mode and per individual amplitude, LAD estimator with and without de-shearing, per mode and per individual amplitude, and the CHP estimator, per individual amplitude. The mean-bias is again given in terms of a multiplicative component m as defined in Equation 4.25.

Table 4.2: Results for bias estimations for each estimator, with and without de-shearing. We have used simulated projected ellipsoids as intrinsic ellipticities, with and without added noise. Estimation bias is given in terms of a multiplicative component m as defined in Equation 4.23. Numbers in parentheses reflect the standard uncertainty in the last digit.

Estim.	m_μ	m_{CHP}
Simulated ellipticities		
Mean	-0.0003(4)	-0.0007(7)
de-g	0.0006(6)	0.0004(4)
LAD	0.0172(4)	0.0154(5)
de-g	-0.0000(2)	-0.0005(2)
CHP	0.0097(2)	0.0052(5)
de-g	0.0000(2)	0.0001(2)
BI	0.0426(7)	0.0421(9)
de-g	-0.0002(6)	0.0002(5)
Added noise		
Mean	-0.0404(7)	-0.040(1)
de-g	-0.045(2)	-0.045(2)
LAD	0.0114(6)	0.011(1)
de-g	-0.0308(8)	-0.031(1)
CHP	0.0053(5)	0.004(1)
de-g	-0.030(1)	-0.028(2)
BI	0.0066(8)	0.007(1)
de-g	-0.040(1)	-0.040(2)

Table 4.3: Results for scales of fixed coverage for LAD. For each sample distribution, the relative efficiencies η are given first, and the (more intuitive) relative change in estimator distribution scale is given second, in percentages of the distribution scale of the mean. Numbers in parentheses reflect the standard uncertainty in the last digit.

Distribution	η_{25} $\Delta s_{25}(\%)$	$\eta_{38.3}$ $\Delta s_{38.3}(\%)$	η_{50} $\Delta s_{50}(\%)$	$\eta_{68.3}$ $\Delta s_{68.3}(\%)$	η_{75} $\Delta s_{75}(\%)$	$\eta_{86.6}$ $\Delta s_{86.6}(\%)$	$\eta_{95.4}$ $\Delta s_{95.4}(\%)$
			Simulated ellipticities				
Gaussian	0.79(3)	0.79(3)	0.81(3)	0.83(2)	0.84(3)	0.81(2)	0.83(3)
	+12(2)	+12(2)	+11(2)	+9(1)	+9(2)	+11(1)	+10(2)
Uniform q	7.2(4)	6.1(3)	5.4(1)	4.55(9)	4.3(1)	3.69(9)	3.22(6)
	-63(1)	-59.4(8)	-56.9(5)	-53.1(5)	-51.7(7)	-48.0(6)	-44.3(5)
Elliptical	2.7(1)	2.5(1)	2.4(1)	2.26(8)	2.20(9)	2.07(8)	1.88(7)
	-39(2)	-36(2)	-35(1)	-34(1)	-33(1)	-31(1)	-27(1)
Disk	6.3(3)	5.8(2)	5.5(2)	5.1(2)	4.9(1)	4.7(1)	4.3(1)
	-60.1(9)	-58.5(8)	-57.3(6)	-55.6(7)	-54.7(7)	-53.9(7)	-52.0(7)
Combined	5.4(2)	5.0(3)	4.9(2)	4.6(2)	4.4(1)	4.1(1)	4.0(2)
	-57.2(8)	-55(1)	-55(1)	-53.6(8)	-52.2(7)	-50.8(7)	-50(1)
			Added noise				
Uniform q	1.37(7)	1.38(5)	1.37(5)	1.38(3)	1.37(3)	1.33(3)	1.29(4)
	-15(2)	-15(2)	-14(2)	-15(1)	-14(1)	-13(1)	-12(1)
Elliptical	1.19(6)	1.18(5)	1.20(6)	1.17(3)	1.18(4)	1.19(3)	1.16(4)
	-8(2)	-8(2)	-9(2)	-8(1)	-8(1)	-8(1)	-7(2)
Disk	1.48(8)	1.49(7)	1.54(7)	1.55(6)	1.57(5)	1.59(5)	1.52(4)
	-18(2)	-18(2)	-19(2)	-20(1)	-20(1)	-21(1)	-19(1)
Combined	1.55(5)	1.53(6)	1.51(8)	1.53(5)	1.51(5)	1.48(6)	1.46(5)
	-20(1)	-19(2)	-19(2)	-19(1)	-19(1)	-18(20	-17(1)
			Full CFHTLenS data				
All	1.17(6)	1.19(4)	1.19(3)	1.21(3)	1.20(2)	1.20(3)	1.12(3)
	-8(3)	-8(1)	-8(1)	-9(1)	-8.6(9)	-9(1)	-6(1)
Red	1.32(7)	1.33(6)	1.35(5)	1.26(4)	1.24(5)	1.19(5)	1.19(8)
	-13(2)	-13(2)	-14(1)	-11(1)	-10(2)	-8(2)	-8(3)
Blue	1.28(4)	1.24(5)	1.24(4)	1.20(3)	1.18(3)	1.19(5)	1.19(8)
	-12(1)	-10(2)	-10(1)	-9(1)	-8(1)	-8(2)	-8(3)
Sbc	1.23(5)	1.19(3)	1.19(5)	1.18(2)	1.17(3)	1.18(3)	1.13(6)
	-10(2)	-8(1)	-9(2)	-7.8(6)	-7.4(9)	-8(1)	-6(2)
Scd	1.21(4)	1.20(4)	1.21(4)	1.18(4)	1.16(3)	1.19(6)	1.12(7)
	-9(2)	-9(2)	-9(2)	-8(1)	-7(1)	-8(2)	-6(3)
			Conservative selection of CFHTLenS data				
All	1.82(7)	1.77(6)	1.72(4)	1.67(5)	1.70(3)	1.65(2)	1.67(6)
	-26(1)	-25(1)	-24(1)	-23(1)	-23.3(7)	-22.2(4)	-23(1)
Red	1.8(1)	1.68(9)	1.66(7)	1.64(6)	1.64(6)	1.66(6)	1.58(6)
	-26(3)	-23(2)	-22(2)	-22(1)	-22(1)	-22(1)	-21(2)
Blue	1.73(4)	1.63(5)	1.62(5)	1.65(4)	1.64(6)	1.64(5)	1.61(9)
	-24.0(8)	-22(1)	-21(1)	-22(1)	-22(1)	-22(1)	-21(2)
Sbc	1.63(8)	1.6(1)	1.59(6)	1.55(4)	1.52(4)	1.49(4)	1.47(7)
	-22(2)	-22(2)	-21(2)	-20(1)	-19(1)	-18(1)	-18(2)
Scd	1.8(1)	1.75(9)	1.70(6)	1.67(4)	1.6(4)	1.65(3)	1.55(5)
	-25(3)	-24(2)	-23(1)	-22.6(9)	-21.9(9)	-22.1(7)	-20(1)

Table 4.4: Results for scales of fixed coverage for the biweight. For each sample distribution, the relative efficiencies η are given first, and the (more intuitive) relative change in estimator distribution scale is given second, in percentages of the distribution scale of the mean. Numbers in parentheses reflect the standard uncertainty in the last digit.

Distribution	η_{25} $\Delta s_{25}(\%)$	$\eta_{38.3}$ $\Delta s_{38.3}(\%)$	η_{50} $\Delta s_{50}(\%)$	$\eta_{68.3}$ $\Delta s_{68.3}(\%)$	η_{75} $\Delta s_{75}(\%)$	$\eta_{86.6}$ $\Delta s_{86.6}(\%)$	$\eta_{95.4}$ $\Delta s_{95.4}(\%)$
			Simulated ellipticities				
Gaussian	0.99(6) +0(3)	1.02(6) -1(3)	1.00(4) -0(2)	1.03(4) -1(2)	1.02(5) -1(2)	1.03(4) -1(2)	1.05(4) -2(2)
Uniform q	1.2(1) -10(3)	1.16(5) -7(2)	1.19(2) -8.2(8)	1.14(3) -6(1)	1.12(2) -6(1)	1.14(5) -6(2)	1.17(6) -7(2)
Elliptical	1.4(2) -14(5)	1.3(1) -13(2)	1.31(8) -13(3)	1.31(8) -13(2)	1.30(6) -12(2)	1.34(9) -14(3)	1.32(3) -13(1)
Disk	1.6(1) -20(3)	1.60(7) -21(2)	1.52(6) -19(2)	1.49(6) -18(2)	1.49(6) -18(2)	1.49(5) -18(1)	1.46(4) -17(1)
Combined	1.54(8) -19(2)	1.5(1) -19(3)	1.5(1) -18(3)	1.46(5) -17(2)	1.47(6) -17(2)	1.47(5) -17(1)	1.45(8) -17(2)
			Added noise				
Uniform q	1.09(6) -4(2)	1.11(5) -5(2)	1.11(7) -5(3)	1.09(4) -4(2)	1.09(5) -4(2)	1.09(6) -4(2)	1.10(4) -5(2)
Elliptical	1.23(8) -10(3)	1.26(9) -11(3)	1.23(9) -10(3)	1.248) -10(3)	1.23(8) -10(3)	1.24(6) -10(2)	1.22(4) -9(2)
Disk	1.25(6) -10(2)	1.24(9) -10(3)	1.24(7) -10(2)	1.23(4) -10(1)	1.23(4) -10(2)	1.23(6) -10(2)	1.19(4) -8(1)
Combined	1.23(7) -10(2)	1.25(6) -10(2)	1.24(7) -10(3)	1.26(6) -11(2)	1.26(7) -11(2)	1.2397) -10(3)	1.22(6) -9(2)
			Full CFHTLenS data				
All	1.06(7) -3(3)	1.06(4) -3(2)	1.05(3) -2(2)	1.10(3) -5(1)	1.09(3) -4(2)	1.08(3) -4(1)	1.06(2) -3.1(8)
Red	1.1(1) -5(6)	1.10(8) -5(3)	1.09(5) -4(2)	1.08(5) -4(2)	1.08(6) -4(3)	1.08(6) -4(3)	1.1(1) -4(5)
Blue	1.09(3) -4(1)	1.06(4) -3(2)	1.05(7) -3(3)	1.08(5) -4(2)	1.10(5) -5(2)	1.10(7) -5(3)	1.07(8) -3(4)
Sbc	1.12(9) -5(4)	1.08(7) -4(3)	1.07(6) -3(3)	1.06(3) -3(1)	1.07(4) -3(2)	1.06(4) -3(2)	1.03(5) -2(2)
Scd	1.02(4) -1(2)	1.03(4) -1(2)	1.05(4) -3(2)	1.05(4) -2(2)	1.04(5) -2(2)	1.06(8) -3(4)	1.02(9) -1(4)
			Conservative selection of CFHTLenS data				
All	1.22(6) -10(2)	1.23(3) -10(1)	1.21(4) -9(2)	1.19(5) -8(2)	1.22(5) -9(2)	1.22(3) -9(1)	1.21(7) -9(3)
Red	1.12(5) -6(2)	1.10(5) -5(2)	1.12(5) -5(2)	1.11(4) -5(2)	1.11(4) -5(2)	1.15(6) -7(2)	1.14(50 -6(2)
Blue	1.27(8) -11(3)	1.22(7) -9(3)	1.23(8) -10(3)	1.23(7) -10(3)	1.23(5) -10(2)	1.23(5) -10(2)	1.21(6) -9(2)
Sbc	1.14(6) -6(2)	1.2(1) -7(5)	1.14(7) -6(3)	1.13(3) -6(1)	1.13(4) -6(2)	1.15(7) -7(3)	1.15(4) -7(1)
Scd	1.3(1) -14(4)	1.3(1) -14(4)	1.30(9) -12(3)	1.28(8) -12(3)	1.25(8) -10(3)	1.26(6) -11(2)	1.28(4) -11(1)

Table 4.5: Results for scales of fixed coverage for CHP. For each sample distribution, the relative efficiencies η are given first, and the (more intuitive) relative change in estimator distribution scale is given second, in percentages of the distribution scale of the mean. Numbers in parentheses reflect the standard uncertainty in the last digit.

Distribution	η_{25} $\Delta s_{25}(\%)$	$\eta_{38.3}$ $\Delta s_{38.3}(\%)$	η_{50} $\Delta s_{50}(\%)$	$\eta_{68.3}$ $\Delta s_{68.3}(\%)$	η_{75} $\Delta s_{75}(\%)$	$\eta_{86.6}$ $\Delta s_{86.6}(\%)$	$\eta_{95.4}$ $\Delta s_{95.4}(\%)$
			Simulated ellipticities				
Gaussian	0.38(3)	0.38(1)	0.38(1)	0.39(2)	0.38(1)	0.37(1)	0.38(2)
	+62(6)	+62(3)	+63(3)	+61(3)	+63(2)	+63(2)	+63(5)
Uniform q	15(2)	11.1(9)	8.9(5)	6.5(5)	5.6(4)	4.4(3)	3.3(3)
	-74(2)	-70(1)	-66.6(9)	-61(1)	-58(1)	-53(1)	-45(2)
Elliptical	2.3(2)	2.03(8)	1.80(7)	1.58(5)	1.51(6)	1.31(8)	1.13(8)
	-33(2)	-30(1)	-25(1)	-21(1)	-19(2)	-13(3)	-6(3)
Disk	6.8(7)	6.0(5)	5.4(2)	4.6(1)	4.3(1)	3.8(2)	3.2(2)
	-62(2)	-59(2)	-57.1(9)	-53.5(6)	-51.6(8)	-49(1)	-44(2)
Combined	5.8(3)	5.0(1)	4.7(2)	4.1(1)	3.8(1)	3.4(1)	2.7(2)
	-59(1)	-55.4(6)	-53.8(8)	-50.4(7)	-48.8(9)	-45(1)	-39(2)
			Added noise				
Uniform q	0.85(4)	0.84(3)	0.86(4)	0.85(4)	0.84(3)	0.80(3)	0.77(40
	+7(2)	+9(2)	+8(3)	+8(2)	+9(2)	+12(2)	+14(3)
Elliptical	0.65(5)	0.65(6)	0.64(4)	0.66(3)	0.64(3)	0.64(3)	0.64(4)
	+24(5)	+24(6)	+25(4)	+24(3)	+25(3)	+25(3)	+25(4)
Disk	0.88(7)	0.89(6)	0.88(4)	0.90(3)	0.90(3)	0.88(4)	0.87(6)
	+7(4)	+6(4)	+6(3)	+6(2)	+5(1)	+6(2)	+7(4)
Combined	0.91(6)	0.92(6)	0.89(5)	0.91(3)	0.89(4)	0.88(5)	0.86(4)
	+5(3)	+4(4)	+6(3)	+5(2)	+6(2)	+6(3)	+8(2)
			Full CFHTLenS data				
All	0.75(4)	0.76(3)	0.77(2)	0.76(2)	0.77(2)	0.75(3)	0.71(2)
	+16(3)	+14(2)	+14(2)	+15(1)	+14(1)	+16(2)	+19(1)
Red	0.81(7)	0.84(7)	0.84(4)	0.77(4)	0.76(5)	0.72(4)	0.72(4)
	+11(5)	+9(4)	+9(2)	+14(3)	+15(4)	+18(3)	+18(3)
Blue	0.81(3)	0.79(3)	0.77(3)	0.75(4)	0.76(3)	0.73(3)	0.68(6)
	+11(2)	+13(2)	+14(2)	+15(3)	+14(3)	+17(2)	+21(5)
Sbc	0.80(4)	0.78(3)	0.77(3)	0.76(3)	0.74(2)	0.71(3)	0.68(4)
	+12(3)	+14(2)	+14(2)	+15(2)	+16(2)	+18(2)	+21(4)
Scd	0.79(5)	0.75(3)	0.74(2)	0.73(3)	0.72(4)	0.72(3)	0.67(5)
	+12(3)	+15(2)	+16(2)	+17(2)	+18(3)	+18(3)	+22(5)
			Conservative selection of CFHTLenS data				
All	1.18(5)	1.14(4)	1.08(5)	1.00(5)	0.99(4)	0.96(4)	0.89(5)
	-8(2)	-6(1)	-4(2)	0(3)	0(2)	+2(2)	+6(3)
Red	1.3(1)	1.21(8)	1.15(5)	1.05(5)	1.02(5)	0.99(5)	0.89(8)
	-13(3)	-9(3)	-7(2)	-2(2)	-1(2)	0(2)	+6(5)
Blue	1.09(5)	1.02(6)	1.03(7)	0.99(5)	0.96(5)	0.97(5)	0.93(9)
	-4(2)	-1(3)	-1(3)	+1(3)	+2(3)	+2(3)	+4(5)
Sbc	1.1(1)	1.1(1)	1.0(1)	0.98(5)	0.95(5)	0.91(4)	0.88(4)
	-4(5)	-3(5)	-1(5)	+1(2)	+3(3)	+5(2)	+7(2)
Scd	1.11(9)	1.11(4)	1.06(6)	1.00(6)	0.97(5)	0.94(5)	0.85(6)
	-5(4)	-5(2)	-3(3)	0(3)	+1(3)	+3(2)	+8(3)

Table 4.6: Results for $\eta_{68.3}$ for different sample sizes. Numbers in parentheses reflect the standard uncertainty in the last digit.

Distribution	Estimator	$N = 10$	$N = 22$	$N = 46$	$N = 100$	$N = 215$	$N = 464$
				Simulated ellipticities			
Gaussian	LAD	0.80(3)	0.76(2)	0.80(3)	0.83(2)	0.79(3)	0.78(3)
	BI	1.01(3)	1.06(3)	1.03(2)	1.03(4)	1.03(5)	1.07(5)
	CHP	0.71(2)	0.50(1)	0.44(1)	0.39(2)	0.33(2)	0.29(1)
Uniform q	LAD	1.84(6)	2.78(5)	3.6(1)	4.55(9)	5.8(2)	6.6(4)
	BI	1.16(9)	1.14(4)	1.20(3)	1.14(3)	1.17(4)	1.18(9)
	CHP	1.30(5)	2.19(7)	3.6(2)	6.5(5)	12.3(5)	22(2)
Combined	LAD	2.94(7)	3.5(2)	4.2(1)	4.6(2)	4.9(1)	5.1(2)
	BI	1.47(7)	1.46(7)	1.45(4)	1.46(5)	1.56(6)	1.53(8)
	CHP	2.03(9)	2.8(2)	3.6(2)	4.1(1)	5.1(2)	6.0(4)
				Added noise			
Uniform q	LAD	1.08(3)	1.27(7)	1.32(6)	1.38(3)	1.41(4)	1.33(4)
	BI	1.09(3)	1.08(7)	1.15(4)	1.09(4)	1.08(4)	1.14(8)
	CHP	0.9(8)	0.93(6)	0.93(3)	0.85(4)	0.74(4)	0.66(4)
Combined	LAD	1.28(5)	1.42(3)	1.53(5)	1.53(5)	1.51(6)	1.56(5)
	BI	1.24(4)	1.23(5)	1.26(3)	1.26(6)	1.25(7)	1.24(8)
	CHP	0.98(9)	1.04(4)	0.98(8)	0.91(3)	0.78(4)	0.70(4)
				CFHTLenS data			
All	LAD	0.91(2)	1.11(5)	1.16(4)	1.21(3)	1.15(3)	1.21(5)
	BI	1.05(2)	1.07(4)	1.10(6)	1.10(3)	1.08(3)	1.09(8)
	CHP	0.77(4)	0.83(6)	0.82(4)	0.76(2)	0.66(2)	0.65(4)
Subset	LAD	1.39(4)	1.59(6)	1.59(5)	1.67(5)	1.74(4)	1.80(4)
	BI	1.17(3)	1.18(6)	1.22(5)	1.19(5)	1.21(6)	1.19(3)
	CHP	1.07(4)	1.12(6)	1.04(3)	1.00(5)	0.99(4)	0.98(6)

Table 4.7: Results for amplitude estimations for FMF, using different estimators (LSQ, LAD, CHP) and models (per mode or per amplitude). Estimation bias is given in terms of a multiplicative component m as defined in Equation 4.25. Efficiencies are determined relative to LSQ per individual amplitude. Numbers in parentheses reflect the standard uncertainty in the last digit.

Estimator		m	$\eta_{68.3}$
	Simulated ellipticities		
LSQ	per mode	0.001(1)	1.14(6)
	amplitude	0.000(1)	N.A.
LAD	per mode	0.0166(6)	4.6(3)
	amplitude	0.046(1)	1.31(4)
	de-g per mode	0.0006(6)	5.4(4)
	amplitude	0.022(1)	1.61(5)
CHP	per amplitude	0.043(2)	0.83(5)
	Added noise		
LSQ	per mode	-0.024(1)	1.09(7)
	amplitude	-0.023(2)	N.A.
LAD	per mode	0.014(1)	1.47(9)
	amplitude	0.032(2)	0.98(7)
	de-g per mode	-0.018(3)	1.55(9)
	amplitude	-0.007(2)	1.14(8)
CHP	per amplitude	0.015(2)	0.41(3)

5

AMICO galaxy clusters in KiDS-DR3: The impact of estimator statistics on the luminosity-mass scaling relation

As modern-day precision cosmology aims for statistical uncertainties of the percent level or lower, it becomes increasingly important to reconsider estimator assumptions at each step of the process, along with their consequences on the statistical variability of the scientific results.

We compare L^1 regression statistics to the weighted mean, the canonical L^2 method based on Gaussian assumptions, to infer the weak gravitational shear signal from a catalog of background ellipticity measurements around a sample of clusters, which has been a standard step in the processes of many recent analyses.

We use the shape measurements of background sources around 6925 AMICO clusters detected in the KiDS third data release. We investigate the robustness of our results and the dependence of uncertainties on the signal-to-noise ratios of the background source detections. Using a halo model approach, we derive lensing masses from the estimated excess surface density profiles.

The highly significant shear signal allows us to study the scaling relation between the r-band cluster luminosity, L_{200}, and the derived lensing mass, M_{200}. We show the results of the scaling relations derived in 13 bins in L_{200}, with a tightly constrained power-law slope of $\sim 1.24 \pm 0.08$. We observe a small, but significant, relative bias of a few percent in the recovered excess surface density profiles between the two regression methods, which translates to a 1σ difference in M_{200}. The efficiency of L^1 is at least that of the weighted mean and increases with higher signal-to-noise shape measurements.

Our results indicate the relevance of optimizing the estimator for inferring the gravitational shear from a distribution of background ellipticities. The interpretation of measured relative biases can be gauged by deeper observations, and the increased computation times remain feasible.

M. Smit, A. Dvornik, M. Radovich, K. Kuijken, M. Maturi, L. Moscardini, and M. Sereno
Astronomy & Astrophysics, Accepted for publication (2021)

5.1 Introduction

Statistics is an essential part of astronomy (Heck et al. 1985, Feigelson 1988, 2009, Feigelson & Babu 2013). The field relies on inferring physical properties, which cannot be determined directly, from observable quantities, which in turn need to be corrected for systematic effects as well as instrumental and observational biases. The key question that always needs to be answered when interpreting observations and results – before discussing how accurately these results can be constrained – is what one is actually seeing.

Weak gravitational lensing, caused by the deflection of light rays by density variations along the traveled path, has been a case in point for the last three decades. Gravitational lensing is a convex focusing effect that can magnify and shear affected background sources. The observed shapes and number counts can conversely yield information about these density variations but need to be disentangled statistically from the unknown intrinsic properties of background sources, such as distance, size (and luminosity), and shape.

The first detections of coherent alignments of galaxy shapes were observed in the background of clusters (Tyson et al. 1990), and subsequently in the emerging fields of galaxy-galaxy lensing (where the lensing "structure" is itself an ensemble of lenses; Brainerd et al. 1996) and cosmic shear (the weak lensing induced by large-scale structure; Wittman et al. 2000, Bacon et al. 2000, Kaiser et al. 2000, Van Waerbeke et al. 2000). Since then, techniques have progressed rapidly, and demands on accuracy have become increasingly stringent.

This is the second in a set of papers wherein we focus on the statistical aspects of inferring the lensing signal from the intrinsic shapes and the estimated lensing geometry, which depends on the distances between the observer, the moment of deflection, and the background sources. Assuming the cosmological principle, the intrinsic shapes of a sample[1] of background galaxies, including their orientation, are random, and the intrinsic galaxy shapes should average out from a sufficiently large sample, leaving the weak lensing signal as a net ellipticity. The common approach has been to take a weighted mean of galaxy ellipticities, which has computational and analytical advantages and, most importantly, is an unbiased estimator of the shear in the absence of pixel noise in the galaxy images (Seitz & Schneider 1997).

In practice, however, there are many sources of noise and the mean is known to be biased, underestimating the underlying shear signal (Melchior & Viola 2012, Viola et al. 2014, Sellentin et al. 2018, Mandelbaum 2018). The distribution of intrinsic galaxy shapes is well known to be non-Gaussian (Lambas et al. 1992, Rodríguez & Padilla 2013) and, in fact, centrally peaked. In Smit & Kuijken (2018, hereafter Paper I), we explored alternative estimators besides the mean that could potentially be better suited for such a cuspy distribution. It was found, using realistic simulated distributions and resampling of Canada-France-Hawaii Lensing Survey (CFHTLenS) shape measurements (Heymans et al. 2012b), that L^1 norm regression, also known as Least Absolute Deviations (LAD), reduces bias from between $\sim -4\%$ and $\sim -4.5\%$ to between $\sim +1\%$ and $\sim -3\%$, while at the same time reducing uncertainty by $\sim 9\%$ to $\sim 23\%$.

In this paper we extend this study by applying these statistics to a weak lensing analysis of 6925 galaxy clusters in the Adaptive Matched Identifier of Clustered Objects (AMICO) cluster catalog (Bellagamba et al. 2011, Radovich et al. 2017, Bellagamba et al. 2018, Maturi

[1] There are several considerations involved in the proper selection of such a sample, as explained in Sects. 5.2 and 5.3.

et al. 2019, Bellagamba et al. 2019) of the third data release of the Kilo-Degree Survey (KiDS-450; de Jong et al. 2017). As opposed to Paper I, in this case the true lensing signal (here in the form of the excess surface density of the clusters) is unknown. We therefore study the relative biases and uncertainties between LAD and the mean, and we compare results to our findings in Paper I.

An important application is then to study the relation between the observable properties of clusters and groups and the physical quantities derived from the lensing signal (i.e., the matter distribution) to better our understanding of galaxy and cluster formation and cosmological models (e.g., Kautsch et al. 2008, Leauthaud et al. 2010, Lesci et al. 2020). We calculate halo masses from the obtained lensing signals and derive a scaling relation between the observed r-band luminosity and the lensing mass, investigating the impact of estimator choice on the resulting constraints.

The order of magnitude of this estimated bias in the weak lensing results can be dominant compared to other sources of uncertainty in the process. Developments in the field have led to current constraints of the multiplicative bias in shape measurements on the order of $\sim 1\%$ (Bernstein & Jarvis 2002, Hirata & Seljak 2003, Heymans et al. 2006, Massey et al. 2007, Miller et al. 2007, Kitching et al. 2008, Bridle et al. 2010, Voigt & Bridle 2010, Bernstein 2010, Kitching et al. 2012, Kacprzak et al. 2012, Melchior & Viola 2012, Refregier et al. 2012, Heymans et al. 2012a, Mandelbaum et al. 2015, Viola et al. 2015, Fenech Conti et al. 2017). The uncertainty in the lensing geometry between the observer, lens, and background sources, introduced by the estimation of the photometric redshift probability distributions, can be a few percent (Hildebrandt et al. 2017, Bellagamba et al. 2019, and Appendix 5.A.1). The broad category of selection biases, for example those introduced by intrinsic alignments, contamination of the background sample by cluster member galaxies, blending, detection, and subsequent selection effects, typically accumulate up to a few percent (Miyatake et al. 2015, van Uitert et al. 2017, Bellagamba et al. 2019) for cluster weak lensing. For instance, estimations on background selection yield a foreground contamination on the order of 2%, which can be partly corrected for, but does increase the uncertainty (Dvornik et al. 2017, Bellagamba et al. 2019, and Appendix 5.A.2). In this study we investigate the usability of background sources to radii smaller than in Bellagamba et al. (2019).

These demands on accuracy and precision become higher as the data yield, and therefore the statistical power of surveys, increases dramatically (Mandelbaum 2018), as achieved by COSMOS[2] (Leauthaud et al. 2007), CFHTLenS[3] (Heymans et al. 2012b), RCSLenS[4] (Hildebrandt et al. 2016), KiDS[5] (de Jong et al. 2013), and DES[6] (Dark Energy Survey Collaboration et al. 2016), and foreseen for future surveys such as LSST[7] (Ivezić et al. 2019) and Euclid[8] (Laureijs et al. 2011). While these two future surveys will require constraints on systematic uncertainty of order $\leq 2 \times 10^{-3}$ (Mandelbaum 2018), we show that, even for weak lensing analyses in the last decade, the bias in shear inference can dominate other sources, such as the aforementioned multiplicative shape measurement bias that is commonly corrected for, as in Viola et al. (2015), Dvornik et al. (2017), and Bellagamba et al. (2019).

[2] http://cosmos.astro.caltech.edu/
[3] http://www.cfhtlens.org
[4] http://www.rcslens.org/
[5] http://kids.strw.leidenuniv.nl/
[6] http://www.darkenergysurvey.org/
[7] https://www.lsst.org/
[8] http://www.euclid-ec.org/

Several other approaches have been made to address this, including analytic modeling of the bias (e.g., Viola et al. 2014), weight corrections and priors (Bonnet & Mellier 1995, Van Waerbeke et al. 2000, Bernstein & Jarvis 2002), or nulling techniques (Herbonnet et al. 2017). The calculation of the main observable, the shapes of lensed background sources, itself relies on statistical methods. These are based mainly on surface brightness moments (Kaiser et al. 1995, Rhodes et al. 2000) or model fitting (Kuijken 1999, Bernstein & Jarvis 2002, Hirata & Seljak 2003, Refregier & Bacon 2003, Kuijken 2006, Miller et al. 2007, Kitching et al. 2008). This means the most common approaches are corrections on a statistic that remains fundamentally skewed (Sellentin et al. 2018, Mandelbaum 2018).

Promising alternative approaches by Bernstein & Armstrong (2014) and Schneider et al. (2015) do not reproduce individual background shapes, but directly determine the underlying shear field from ensembles of background sources, reconsidering these steps in the chain of statistical inference. While future lensing surveys will require innovative improvements, these methods and their priors need to be gauged by deep observations of high signal-to-noise, and it is of fundamental importance that these calibrations are well constrained and do not suffer from even subtle systematic biases. In other words, the comparison of several perspectives is paramount in determining what we actually see.

The remainder of this paper is organized as follows. We introduce the definitions of galaxy shapes and the weak lensing formalism in Sect. 2 and relate these to our statistical approach. Data, analysis methods, and selection criteria are described in Sect. 3, while Sect. 4 states our results and analysis. Section 5 gives a summary of our conclusions.

Throughout this paper we assume a Planck (Planck Collaboration et al. 2014) cosmology with $\Omega_{\rm M} = 0.315$, $\Omega_\Lambda = 0.685$, and $H_0 = 100.0\,h$ km s^{-1} Mpc^{-1}. All measurements are in co-moving units, unless specifically noted otherwise, such as in Sect. 5.2.

5.2 Weak gravitational lensing statistics

We briefly review the principles of weak gravitational lensing and relate the central concepts to our statistical approach, introducing the terminology and notation conventions used in this paper. We refer the reader to excellent reviews, such as Bartelmann & Schneider (2001), Schneider (2006), Hoekstra & Jain (2008), and Bartelmann & Maturi (2017), for more in-depth approaches.

5.2.1 Principles of weak lensing

Rays of light are deflected by the curvature or space-time due to mass inhomogeneities along their path. A mass overdensity acts as a convex lens on the light rays from distant sources behind that lens to an observer. In this section, we use $D_{\rm l}$ to denote angular-diameter distances from the observer to the lens, $D_{\rm ls}$ from the lens to the background source, and $D_{\rm s}$ from the observer to the background source (see Fig. 5.1), and in the remainder of this paper we translate quantities to co-moving units where necessary.

For the purposes of this work, the extent of the lensing mass along the line of sight, compared to the distances from the observer to the lens and from the lens to the background source, can be considered negligible. In this so-called thin-lens approximation, the deflection of light rays by a deflection angle, $\hat{\vec{\alpha}}$, leads to an effective angular displacement (again, see

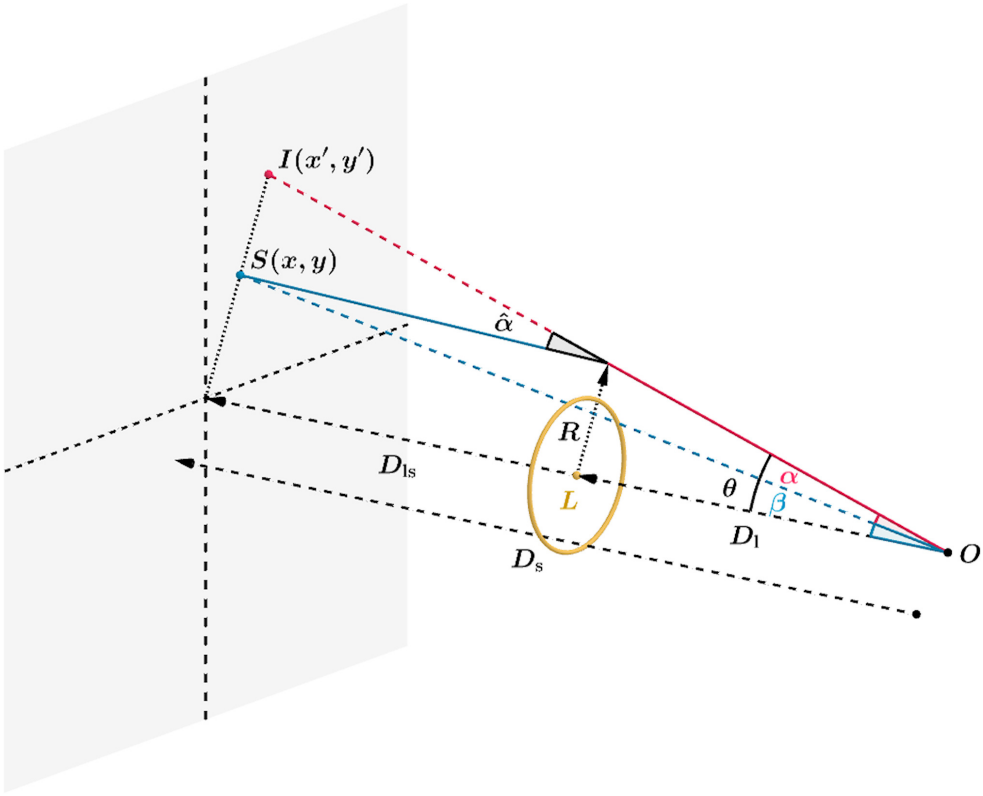


Figure 5.1: Representation of a gravitational lens system, showing the displacement of a source at position $S(x, y)$ to an image at position $I(x', y')$, where we take the origin of the source plane to be collinear with the position of the lens, L, and the observer, O.

Fig. 5.1),

$$\vec{\alpha} = -\frac{D_{ls}}{D_s}\hat{\vec{\alpha}}, \tag{5.1}$$

also called the reduced deflection angle, which relates the observed position, $\vec{\theta}$, of a distant point source to its unlensed position, $\vec{\beta}$, by the lens equation

$$\vec{\beta} = \vec{\theta} - \vec{\alpha}. \tag{5.2}$$

It can be shown through the relation between $\hat{\vec{\alpha}}$ and the three-dimensional gravitational potential, Φ, that this displacement is then described by $\vec{\alpha} = \vec{\nabla}_\theta \psi$, where

$$\psi = \frac{2}{c^2}\frac{D_{ls}}{D_l D_s}\int \Phi \, dz \tag{5.3}$$

is called the (two-dimensional) lensing potential.

The differential effect of the deflection of light on the images, $I(x, y)$, of extended background sources can to first order be described as a coordinate transformation by taking the

derivatives in the lens equation (Eq. 5.2) of the original angular position, β, with respect to the observed position, θ. Substituting $\vec{\nabla}_\theta \psi$ for $\vec{\alpha}$, we obtain the Jacobian matrix,

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1-\psi_{11} & -\psi_{12} \\ -\psi_{21} & 1-\psi_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \tag{5.4}$$

with

$$\psi_{ij} = \frac{\partial^2 \psi}{\partial\theta_i \partial\theta_j}, \tag{5.5}$$

resulting in the lensed image $I(x', y')$, which is the key observable in our weak lensing study.

Critical surface mass density

To interpret the effect on the source image, we note that such a transformation can be decomposed into three parts, namely the identity ($\boldsymbol{I}$), an isotropic part that describes a multiplication, and an anisotropic traceless part that describes a shearing of the image:

$$\boldsymbol{I} - \frac{1}{2}(\psi_{11}+\psi_{22})\boldsymbol{I} + \begin{pmatrix} -\frac{1}{2}(\psi_{11}-\psi_{22}) & -\psi_{12} \\ -\psi_{21} & \frac{1}{2}(\psi_{11}-\psi_{22}) \end{pmatrix}. \tag{5.6}$$

To relate ψ_{ij} with the density of the lensing mass, we start with the isotropic term, which is half the Laplacian of the lensing potential: $\frac{1}{2}(\psi_{11}+\psi_{22}) = \frac{1}{2}\nabla^2_\theta \psi$. From Eq. 5.3, we obtain

$$\frac{1}{2}\nabla^2_\theta \psi = \frac{1}{c^2}\frac{D_\mathrm{l} D_\mathrm{ls}}{D_\mathrm{s}} \int 4\pi G\rho \, dz, \tag{5.7}$$

which is a dimensionless quantity. Defining the surface mass density as

$$\Sigma \equiv \int \rho \, dz \tag{5.8}$$

and gathering the rest of the right-hand side into

$$\frac{4\pi G}{c^2}\frac{D_\mathrm{l} D_\mathrm{ls}}{D_\mathrm{s}} \equiv \Sigma_\mathrm{cr}^{-1}, \tag{5.9}$$

with Σ_cr being the critical surface mass density, we find that the isotropic term can be written as

$$\kappa \equiv \frac{1}{2}\nabla^2_\theta \psi = \frac{\Sigma}{\Sigma_\mathrm{cr}}, \tag{5.10}$$

with κ a normalized dimensionless surface mass density. Recognizing that $\nabla^2_\theta \psi = \vec{\nabla} \cdot \vec{\alpha}$ is the divergence of the deflection of the light rays (i.e., the manner in which those light rays converge due to the lensing effect), κ is simply the convergence.

Shear and intrinsic ellipticity

The shear matrix in Eq. 5.6 has two independent components, simply called the shear $\gamma = \gamma_1 + i\gamma_2$, with $\gamma_1 = \frac{1}{2}(\psi_{11} - \psi_{22})$ and $\gamma_2 = \psi_{12} = \psi_{21}$. Equation 5.4 then becomes

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{5.11}$$

This transformation leads to the magnification and distortion of the light distribution of background sources. In this work, we focus on the most commonly used net distortion or reduced shear $g = g_1 + ig_2 \equiv (\gamma_1 + i\gamma_2)/(1-\kappa)$,

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = (1-\kappa) \begin{pmatrix} 1-g_1 & -g_2 \\ -g_2 & 1+g_1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \tag{5.12}$$

where the transformation is written as a multiplication of $(1-\kappa)$ and a distortion matrix describing the alignment of lensed sources in the foreground potential.

The effect on a circular source is a shearing into an ellipse with axis ratio $q = \frac{b}{a}$ as

$$q = \frac{1-|g|}{1+|g|} \quad \Leftrightarrow \quad |g| = \frac{1-q}{1+q} = \frac{a-b}{a+b} \tag{5.13}$$

and position angle φ via

$$g = |g|\,(\cos 2\varphi + i \sin 2\varphi)\,. \tag{5.14}$$

As mentioned before, we do not measure this gravitational distortion directly. Background sources have an intrinsic shape distribution, and we effectively measure the combined effect of their intrinsic shape and a weak lensing distortion. It is adequate to describe images by their quadrupole brightness moments or their ellipticities as well as the respective response to weak shear distortions. It is straightforward to use the common definition[9] of ellipticity, defined as the reduced shear needed to create the intrinsic shape $\epsilon = \epsilon_1 + i\epsilon_2$ of a source from an image with circular isophotes (Bernstein & Jarvis 2002, Kuijken 2006). The resulting ellipticity, ϵ, after transforming an image with intrinsic[10] ellipticity ϵ^I by a distortion, g, is then given by (Seitz & Schneider 1997)

$$\epsilon = \frac{\epsilon^I + g}{1 + g^*\epsilon^I} \quad \text{for} \quad |g| \leq 1\,, \tag{5.15}$$

with g^* the complex conjugate of g.

The intrinsic shape distribution is called the shape noise and, assuming no preferred direction on the sky, should average to zero: $\left\langle \epsilon^I \right\rangle = 0$. This way, each background shape measurement, ϵ, is then an independent estimate of the underlying reduced shear, g.

In this paper we make use of the fact that the lensing signal is weak (i.e., $\kappa \ll 1$) and assume $g \approx \gamma$.

[9] An alternative definition of ellipticity is often denoted as $|\chi| = \frac{1-q^2}{1+q^2}$, related to the geometrical eccentricity, and called polarization (e.g., Seitz & Schneider 1995, Viola et al. 2014).

[10] We note that our notation differs from Paper I. Here, the measured ellipticity is denoted as ϵ, instead of $\tilde{e}$ (Paper I), and the intrinsic ellipticity is denoted as ϵ^I, instead of e.

5.2.2 Estimation of the surface density profile

The shear induced by gravitational lensing is sensitive to the density contrast. For an axisymmetric lens, we can write $|\gamma|(R) = \overline{\kappa}(\leq R) - \kappa(R)$, where $\overline{\kappa}$ is the average convergence within radius R. In fact, this relation holds for other mass distributions if we average azimuthally around the lens. In this work, we study the stacked signal of many lenses and assume a net axisymmetry (see, e.g., Evans & Bridle 2009, Oguri et al. 2010, Clampitt & Jain 2016, van Uitert et al. 2017, for weak lensing studies on elliptical lenses).

Since it can be seen from Fig. 5.1 that the gravitational shear acts in the radial direction, we define the tangential and cross components of the shear as

$$\begin{pmatrix} \gamma_+ \\ \gamma_\times \end{pmatrix} = \begin{pmatrix} -\cos(2\phi) & -\sin(2\phi) \\ \sin(2\phi) & -\cos(2\phi) \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix}, \tag{5.16}$$

with ϕ the counterclockwise angle between the positive x axis[11] and the vector from lens to source. This gives

$$\overline{\Sigma}(\leq R) - \Sigma(R) \equiv \Delta\Sigma(R) = \gamma_+(R)\,\Sigma_{\rm cr}\,, \tag{5.17}$$

with $\Delta\Sigma(R)$ the excess surface density (ESD) at a radius R around the lensing mass. In axisymmetric lenses, the cross component of the shear cannot arise from gravitational lensing and should average to zero, if only produced by intrinsic source orientations, and can therefore be used as an indication of systematic effects, such as imperfect corrections for the point-spread function (PSF; Schneider 2003, and Appendix 5.A.3).

The ESD is then estimated using the observed ellipticities of an ensemble of sources around the lens

$$\Delta\Sigma(R) = \left\langle \epsilon_+\,\Sigma_{\rm cr,ls} \right\rangle (R)\,, \tag{5.18}$$

with each $\epsilon_+\Sigma_{\rm cr,ls}$ an independent, albeit noisy, estimate of the ESD. Here, $\langle\cdot\rangle$ denotes a weighted average, with weights to be specified.

The $\Sigma_{\rm cr,ls}$ behaves as a geometric scaling factor, indicating the lensing efficiency for each lens-source combination. Since the variance of the noise in $\Delta\Sigma$ is then affected by $\Sigma^2_{\rm cr,ls}$, the relative precision, or inverse variance, carried by each ϵ_+ scales as $\Sigma^{-2}_{\rm cr,ls}$.

In this paper we study the ESD profile in co-moving radial bins, and we therefore use the co-moving critical surface density

$$\Sigma_{\rm cr,com} = (1 + z_{\rm l})^2\,\Sigma_{\rm cr,prop}. \tag{5.19}$$

In practice, the distance to each background source is not known exactly and is estimated by its redshift probability distribution, $p\,(z_{\rm s})$. Taking this into account, we estimate the co-moving critical surface density via

$$\left\langle \Sigma^{-1}_{\rm cr,ls} \right\rangle = \frac{4\pi G}{c^2} D(z_{\rm l})\,(1 + z_{\rm l})^2 \int \frac{D(z_{\rm l}, z_{\rm s})}{D(z_{\rm s})} p(z_{\rm s}) dz_{\rm s}. \tag{5.20}$$

[11] Of the coordinate system in which γ_1 and γ_2 are defined.

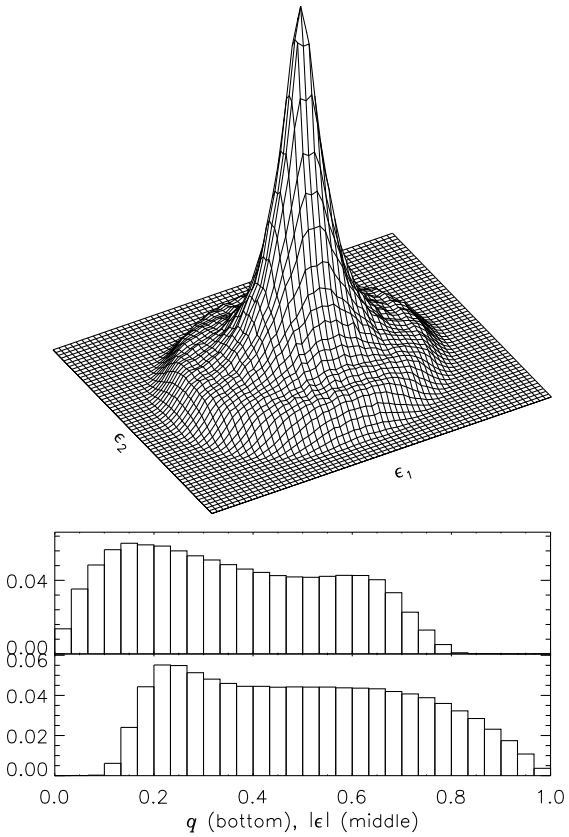


Figure 5.2: Estimated ellipticity and axis ratio distributions of sources in the KiDS-450 catalog. Top: Two-dimensional histogram of ellipticities. Middle: Histogram of the absolute ellipticity, $|\epsilon|$. Bottom: Histogram of the ellipse axis ratio, q.

5.2.3 Statistical framework

In this section we discuss the estimation of $\langle \epsilon_+ \, \Sigma_{\rm cr,ls} \rangle$. We refer to Paper I for a complementary discussion.

Important aspects of a good estimator, $\hat{\epsilon}$, are: (i) minimal bias, defined as the difference between the expected value of the estimator, $\langle \hat{\epsilon} \rangle$, and the value of the quantity being estimated, for instance the shear (γ) or, in this case, $\Delta\Sigma$; (ii) high efficiency, proportional to the inverse variance of the estimator, $\sigma_{\hat{\epsilon}}^{-2}$; and (iii) robustness, meaning the estimator retains these properties for a sufficient range of likely parameter distributions.

Bias

Even though the measured ellipticity, ϵ, is not a linear combination of the intrinsic shape, ϵ^I, and the shear, γ, it can be shown (Seitz & Schneider 1997) that, in the absence of further uncertainties, the mean $\mu(\epsilon)$ is an unbiased estimator for the underlying shear and that this

is independent of the intrinsic ellipticity distribution, $P\left(\epsilon^I\right)$. In the canonical approach, the ESD is therefore estimated as a weighted mean of an ensemble of lens-source combinations,

$$\Delta\Sigma(R) = \frac{\sum_{\rm ls} w_{\rm ls}\epsilon_{+,{\rm ls}}\Sigma_{\rm cr,ls}}{\sum_{\rm ls} w_{\rm ls}}\,, \tag{5.21}$$

where we use

$$w_{\rm ls} = w_{\rm s}\left\langle\Sigma^{-1}_{\rm cr,ls}\right\rangle^2\,. \tag{5.22}$$

Here, the weight $w_{\rm s}$ is assigned to each measured ellipticity, scaled by the estimated lensing efficiency, as explained in the previous section (see, e.g., Viola et al. 2015, Dvornik et al. 2017, Bellagamba et al. 2019).

In practice, there are various sources of uncertainty at each step of the process, such as source selection bias, distortion by the PSF, and biases due to the measurement pipeline. These lead to convolutions of the ellipticity distribution, before and after the gravitational lensing effect. The result is a bias in the mean as an estimate of the ESD (Melchior & Viola 2012, Refregier et al. 2012, Kacprzak et al. 2012, Viola et al. 2014, Kacprzak et al. 2014). In this case, the intrinsic shape distribution will play a role.

The weighted mean, μ, is a statistic that, for a set of measurements ϵ_i with weights w_i, finds the estimate of γ that minimizes the loss function

$$S_\mu = \sum_i w_i\left[(\epsilon_{i,1}-\gamma_1)^2 + (\epsilon_{i,2}-\gamma_2)^2\right]\,, \tag{5.23}$$

that is, it is a least squares (LSQ) or L^2 norm regression method and arises naturally as the optimal estimator for Gaussian distributions.

Figure 5.2 shows that the measured ellipticity distribution, $P(\epsilon)$, displays crucial differences with a Gaussian distribution, showing a sharp peak and a slower decline, including a higher number of high ellipticities, $|\epsilon|$. This central peak is an unbiased tracer of the underlying shear (Paper I). By Eq. 5.23, the mean is sensitive to outliers and therefore not robust when inferring the shear.

In contrast, the LAD or L^1 norm regression minimizes the loss function,

$$S_{\rm LAD} = \sum_i w_i\sqrt{(\epsilon_{i,1}-\gamma_1)^2 + (\epsilon_{i,2}-\gamma_2)^2}\,. \tag{5.24}$$

The LAD estimate is also known as the median in one dimension or the spatial median in higher dimensions. LAD is more sensitive to the peak and less sensitive to high ellipticity outliers. Where the mean is expected to be biased low (Melchior & Viola 2012, Refregier et al. 2012, Kacprzak et al. 2012, Viola et al. 2014, Kacprzak et al. 2014), we expect this to be less so for the LAD (Paper I).

Efficiency

The formal definition of efficiency, $\tilde{\eta}$, relates the inherent (Fisher) information, $\mathcal{I}$, of a sample to the statistical variability around the expected value of the estimator, usually taken to be the variance, $\sigma^2_{\hat{\epsilon}}$, of the estimator:

$$\tilde{\eta} = \frac{1}{\mathcal{I}\cdot\sigma^2_{\hat{\epsilon}}}\,. \tag{5.25}$$

Since the variance of an unbiased estimator cannot be less than the reciprocal of the information, $\mathcal{I}^{-1} \leq \sigma^2_{\hat{\epsilon}}$, we have $0 \leq \tilde{\eta} \leq 1$ (Rao 1945, Cramér 1946).

As we are comparing two estimators with unknown bias, it is appropriate to use the relative efficiency,

$$\eta = \frac{\sigma^2_{\mu}}{\sigma^2_{\rm LAD}}\,, \tag{5.26}$$

with $\eta < 1$ indicating a higher efficiency for the mean, and vice versa.

Paper I showed that the LAD consistently performed better than the mean, with both higher efficiency and less bias, for various cusped intrinsic ellipticity distributions, including the shear catalog from CFHTLenS. In what follows, we take the CFHTLenS shear distribution shape to be representative of KiDS data as well since both surveys were processed with the THELI pipeline (Erben et al. 2013) and the shape measurement pipeline *lens*fit (Miller et al. 2007, Kitching et al. 2008, Miller et al. 2013).

5.2.4 Halo model

Studying the effects of estimator choice on the weak lensing signal forms the technical core of this paper. The scientific goal, however, is to assess the relevance on the inference of physical quantities, such as the derivation of a lensing halo mass from an ESD profile. Since we calculate the stacked signal for an ensemble of clusters with some common (observable) property (here a range in r-band luminosity), of interest is the scaling relation between the observable and derived lensing mass, M_{200}, where we use the definition with respect to the mean density of the universe.

To do so, we modeled the lens density profile the same way as Dvornik et al. (2017), using the halo model (Seljak 2000, Peacock & Smith 2000, Cooray & Sheth 2002, van den Bosch et al. 2013, Cacciato et al. 2013, Mead et al. 2015). The initial lens density profile is described by a Navarro-Frenk-White profile (Navarro et al. 1996). We used the mass-concentration relation given by Duffy et al. (2008) and allowed for a re-normalization factor, $f_{\rm c}$ (Viola et al. 2015).

A dominant source of systematic bias in stacked weak lensing analyses is a miscentering of the lenses, which can be due to an offset of the cluster halo with the visible distribution of galaxies (see, e.g., George et al. 2012) or the resolution of the cluster detection method (less than 0.1 Mpc h^{-1} for AMICO; see Bellagamba et al. 2018). Following Johnston et al. (2007) as well as numerous subsequent works (e.g., Oguri et al. 2010, Viola et al. 2015, Dvornik et al. 2017, Bellagamba et al. 2019, Giocoli et al. 2021), we allowed a fraction, $p_{\rm off}$, of clusters to be offset from the center of the galaxy distribution, effectively smoothing the central stacked $\Delta\Sigma$ profile with a characteristic radius, $\mathcal{R}_{\rm off}$.

At large radii, typically beyond a few megaparsecs, the clustering of dark matter halos starts to dominate the signal. This "two-halo" term depends on the halo bias, b (Dvornik et al. 2017), and is modeled following Tinker et al. (2010). At small radii, the baryonic component of central galaxies can contribute to the signal, which is adequately described by a point mass, $M_\star$, in the model (Viola et al. 2015, Dvornik et al. 2017).

In Table 5.1 we summarize these six free parameters for our halo model implementation, analogous to Dvornik et al. (2017).

In the AMICO cluster sample with 6925 lenses in 440 square degrees, many background sources are lensed by more than one cluster, contributing to the estimate of the ESD profile in

Table 5.1: Summary of the halo model fitting parameters and priors.

Parameter		Prior
f_c		[0.0, 8.0]
p_{off}		[0.0, 1.0]
$\mathcal{R}_{off}$	[h^{-1} Mpc]	[0.0, 1.0]
b		[0.0, 10.0]
$\log(M_\star)$	$\log[\, h^{-1} M_\odot\,]$	[9.5, 12.5]
$\log(M_{200})$	$\log[\, h^{-1} M_\odot\,]$	[11.0, 17.0]

various radial bins of different clusters. In the model fitting, we took the covariance between the ESD estimates into account, as described in Sect. 5.3.3 (Viola et al. 2015, Dvornik et al. 2017, Bellagamba et al. 2019).

5.3 Data and analysis

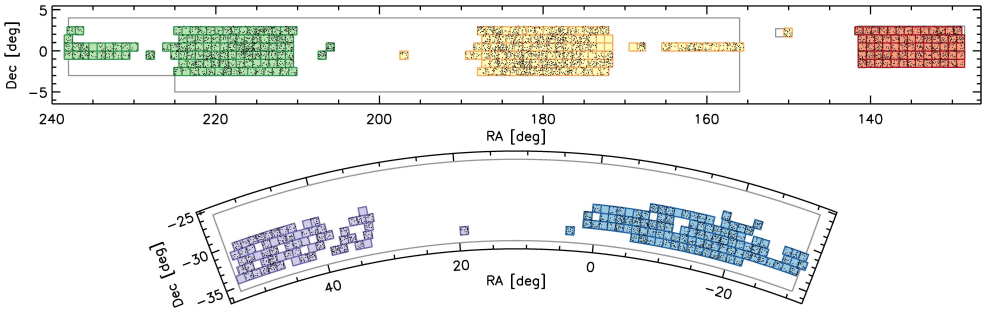


Figure 5.3: Overview of the KiDS-450 observations, with the KiDS-N (upper) and KiDS-S (lower) patches. The solid gray lines represent the planned KiDS survey area. Overplotted are the observed 1 square degree tiles, color coded with respect to their correspondence with the GAMA survey patches (G9 red, G12 yellow, G15 green, G23 blue, and GS purple; see Hildebrandt et al. 2017, for more details). The AMICO clusters analyzed in this work are represented by black dots.

In this paper we use a lensing cluster catalog and a background source catalog from KiDS-450 (de Jong et al. 2017), an optical wide-field imaging survey with OmegaCAM (Kuijken 2011) on the VLT Survey Telescope (VST; Capaccioli & Schipani 2011, de Jong et al. 2013). KiDS-450 consists of two patches, KiDS-N and KiDS-S (see Fig. 5.3), with 454 tiles of imaging data, for a total of 449.7 deg^2, in four optical filters, *ugri*. The survey was designed for lensing, ensuring a stable PSF, low seeing ($< 0.96''$, with an average of $0.66''$ in r), and good photometric redshifts (photo-z; Hildebrandt et al. 2017).

The KiDS data were reduced with ASTRO-WISE (Valentijn et al. 2007, Verdoes Kleijn et al. 2012, Begeman et al. 2013, McFarland et al. 2013), as described in de Jong et al. (2015), Hildebrandt et al. (2017). Photometric redshifts, also termed z_B, were determined

using a Bayesian photo-z estimation (BPZ; Benítez 2000, Coe et al. 2006) with PSF-matched photometry, as described in Hildebrandt et al. (2012), Kuijken et al. (2015), Hildebrandt et al. (2017).

5.3.1 Lenses

We made use of the galaxy cluster catalog derived with AMICO (Bellagamba et al. 2011, Radovich et al. 2017, Bellagamba et al. 2018), extracted from 440 tiles of KiDS-450 data and described in Maturi et al. (2019) and Bellagamba et al. (2019). For each cluster, the luminosity L_{200} is defined[12] as the sum of r-band luminosities of bright candidate member galaxies, weighted by membership probability (see Maturi et al. 2019). We selected galaxies with k-corrected r-band magnitudes brighter than $m^*(z_l) + 1$ within $R_{200}(z_l)$, where z_l is the estimated cluster redshift and R_{200} is derived from the adopted cluster model and is used in the construction of the cluster detection filter, as defined in Maturi et al. (2019). In this sense, L_{200} is defined analogously to the apparent richness, λ^*, which is a sum of membership probabilities of galaxies with $m < m^* + 1.5$, within R_{200}.

We selected clusters in the range $0.1 \leq z_l \leq 0.6$. We excluded clusters below $z = 0.1$ due to their unfavorable lensing geometry and above $z = 0.6$ due to the low density of background sources. For some clusters, no lens-source pairs were found, due to source selection criteria or masking. Our final selection comprises 6925 clusters, divided over the KiDS-450 survey area as shown in Fig 5.3 and described in Table 5.2. The redshift distribution of these clusters is shown in Fig. 5.4, with a median redshift of $z_l = 0.39$.

Table 5.2: Summary of the survey patches as described in Hildebrandt et al. (2017), with corresponding numbers of KiDS mosaic tiles and analyzed clusters.

KiDS field	Subfield	Tiles	Clusters
North	G9	65	1039
	G12	113	1778
	G15	112	1737
South	G23	101	1517
	GS	63	854

We divided the clusters into 13 bins of cluster L_{200}. The limits of these bins were chosen so that the signal-to-noise ratios of the ESD measurements were approximately the same in each bin. We give an overview of these bins, together with the estimated M_{200}, in Table 5.3.

5.3.2 Sources

We selected an initial sample of background sources using the same photometric redshift criteria as Hildebrandt et al. (2017), $0.1 < z_B \leq 0.9$, to reduce the outlier rate. We also applied the cut $z_l + \Delta z < z_B$, following Dvornik et al. (2017). Here, $\Delta z = 0.2$ is an offset between the redshift estimation, z_l, of the cluster by AMICO and the photometric redshift,

[12] We note that this does not take intracluster light into account.

z_B, of the source to sufficiently lessen the contamination of the background sample by cluster member galaxies (see also Appendix 5.A.2).

Our selection of AMICO clusters is deeper than the lenses from the Galaxy And Mass Assembly (GAMA Driver et al. 2011, Robotham et al. 2011) catalog used in Dvornik et al. (2017). As can be seen in Fig. 5.4, the redshift distributions of lenses and background sources significantly overlap, and the cut at $\Delta z = 0.2$ reduces the number density severely for clusters at higher redshift. Following Bellagamba et al. (2019), we also selected background sources using the color selection proposed by Oguri et al. (2012):

$$g - r < 0.3 \quad \vee \quad r - i > 1.3 \quad \vee \quad r - i > g - r. \tag{5.27}$$

In Fig. 5.5 we show the photometric redshift distribution of this cut in the KiDS-450 catalog and compare it to the photometric and spectroscopic redshift distribution of the same cut in the spectroscopic redshift (spec-z) catalog used in Hildebrandt et al. (2017). Based on this analysis, we additionally required $z_B \geq 0.6$ for this selection to reduce contamination by sources at low redshift and find that 98 % of the galaxies in this color selection have $z_{spec} > 0.6$.

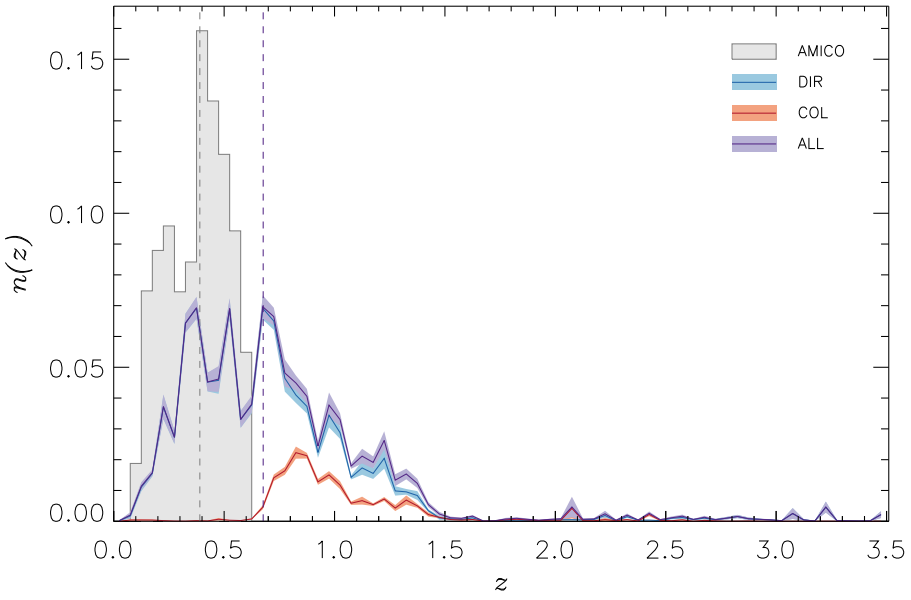


Figure 5.4: Redshift distribution of AMICO clusters (gray), with a median redshift of $z_l = 0.38$ (dashed gray), and KiDS-450 background sources (purple), with a median redshift of $z = 0.68$ (dashed purple). In blue (DIR), we show the initial selection following Hildebrandt et al. (2017) and Dvornik et al. (2017). In red, we show the estimated redshift distribution of the *gri* color selection (COL), corresponding to the bottom panel of Fig. 5.5.

Redshift distribution

To estimate the redshift distribution of background galaxies, we did not directly use the individual redshift probability distribution, $p(z_s)$, per source galaxy. Instead, we applied a

weighted direct calibration method (DIR), as motivated by Hildebrandt et al. (2017).

For each cluster, we used the spec-z catalog described in Hildebrandt et al. (2017) to select objects using the same selection criteria as described above. We then used the normalized spectroscopic redshift distribution, $n(z_s)$, of this sample to calculate the co-moving critical surface density analogous to Eq. 5.20:

$$\left\langle \Sigma_{\rm cr,l}^{-1} \right\rangle = \frac{4\pi G}{c^2} D(z_l)\,(1+z_l)^2 \int_{z_l+\Delta z}^{\infty} \frac{D(z_l, z_s)}{D(z_s)} n(z_s) dz_s. \tag{5.28}$$

The resulting redshift distribution for selected sources from the full KiDS-450 catalog is shown in Fig. 5.4.

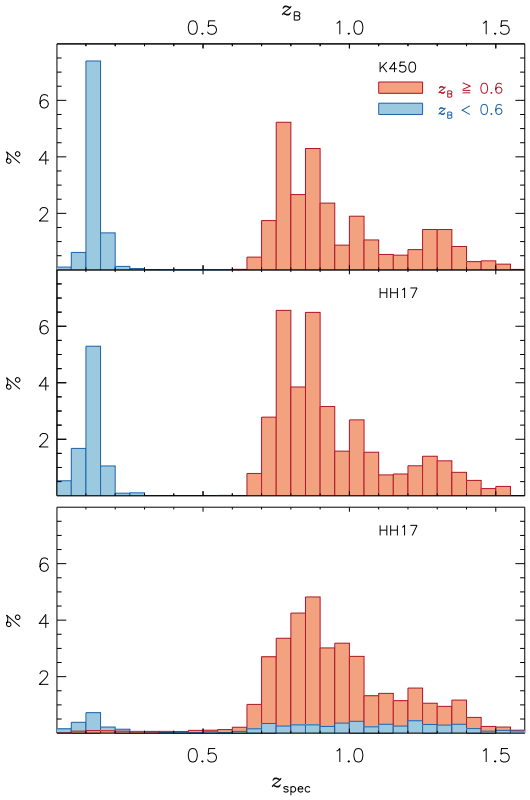


Figure 5.5: Redshift distribution of background sources selected by color. The upper panel shows the distribution of the photometric redshift, z_B, of the sources in the KiDS-450 catalog that satisfy the color cut of Eq. 5.27, of which we select the sources with $z_B \geq 0.6$ (red) and discard those with $z_B < 0.6$ (blue). The bottom two panels show the same selection applied to the spec-z catalog (Hildebrandt et al. 2017), plotted in terms of photometric redshift (middle) and spectroscopic redshift (lower). We find the contamination of sources with $z_B \geq 0.6$ and $z_{spec} < 0.6$ is $\sim 2\%$.

Shape measurements

For shape measurements, the *r*-band data were reduced using the THELI pipeline, developed to meet the requirements for weak gravitational lensing analyses (Erben et al. 2005, 2009, Schirmer 2013, Erben et al. 2013). Galaxy shapes in the KiDS-450 catalog were then measured by *lens*fit (Miller et al. 2007, Kitching et al. 2008, Miller et al. 2013, Fenech Conti et al. 2017).

For each source, *lens*fit produces the ellipticity (ϵ_1, ϵ_2), an approximately inverse-variance weight w_s (see Miller et al. 2013), and a fitting quality parameter. We excluded sources with unreliable ellipticities from our source sample, using the same *lens*fit selection criteria as described in Hildebrandt et al. (2017).

In Paper I we compared the performance of estimators for ellipticity measurements in the CFHTLenS data with a subset of that catalog, selecting sources on the signal-to-noise ratio parameter $\nu_{\rm SN}$ output by *lens*fit. We repeated that approach for a qualitative comparison here, using two subsets of the selected KiDS-450 sources. The first set selects sources with $\nu_{\rm SN} \geq 20$, similar to Paper I, retaining $\sim 30\%$ of the full background sample. The second set is a more stringent cut of the first set, additionally selecting objects with $w_s \geq 14.5$, comprising $\sim 20\%$ of the full sample.

Effective source density

The KiDS-450 catalog includes a filtering on general object detection and quality flags, for example, possibly blended sources or artifacts, as described by Kuijken et al. (2015) and Hildebrandt et al. (2017), and we discarded objects that lie in a mask. This removed approximately $\sim 12\%$ of the sources. Our final selection comprises 14124197 sources, which translates to an effective number density of $n_{\rm eff} \approx 8.23$ arcmin^{-2}, as defined in Heymans et al. (2012b),

$$n_{\rm eff} = \frac{1}{A} \frac{(\sum_i w_i)^2}{\sum_i w_i^2}\,, \tag{5.29}$$

with A the effective surface area, excluding masked regions.

5.3.3 Implementation

ESD estimation

Following Bellagamba et al. (2019), we measured the ESD in 14 logarithmic bins between 0.1 Mpc h^{-1} and 3.16 Mpc h^{-1}. Not only does this make for an easy comparison of the results, but it has several other practical advantages.

We avoided radii smaller than the AMICO detection pixel size, which has a median size of 0.1 Mpc h^{-1}, to lessen the chance of a mismodeling the halo miscentering (Sect. 5.2.4). Here, the line of sight is also most contaminated by cluster members, which can lead to an overabundance by incorrectly including ellipticity measurements that carry no lensing signal, or by an obscuring and blending of background sources, which leads to an under-abundance of sources. While these effects may partially cancel out in the number counts, the effects on the ESD measurements do not cancel out, as the first leads to a diluted signal and the second to a very poor signal-to-noise ratio (see Appendix 5.A.2 for an assessment of cluster member contamination).

At large radii, systematic additive biases can start to play a role (see, e.g., Dvornik et al. 2017, for this data set), which may differ for each KiDS survey patch (Fenech Conti et al. 2017, Hildebrandt et al. 2017). Another concern at larger separations is that the two-halo term becomes the dominant contribution to the ESD signal, which means we would need to properly constrain the halo bias, and we explain below how our approach does not fully take the clustering of dark matter halos into account.

The combination of background selection criteria from Dvornik et al. (2017) and Bellagamba et al. (2019) allows us to retain the three inner radial bins between 0.1 and 0.2 Mpc h^{-1}. We justify this inclusion in Appendix 5.A.2, where we repeat the tests of Dvornik et al. (2017).

Each lens-source pair was then assigned a combined weight of

$$w_{\rm ls} = w_{\rm s}\Sigma_{\rm cr,l}^{-2}\,, \tag{5.30}$$

as motivated in Sect. 5.2.2.

For LAD optimization, that is, the estimator that minimizes the L^1 norm (Eq. 5.24), there exists no general analytic solution. The problem can, however, be formulated as a linear optimization, which can be solved iteratively (e.g., with simplex-based methods; Barrodale & Roberts 1973). In our weak lensing analyses, we find that convergence is robust.

To derive the covariance matrices for the ESD estimates using the mean and LAD in the same way, we can therefore also not employ the analytical prescription of Viola et al. (2015) used in earlier KiDS analyses (e.g., Sifón et al. 2015a, van Uitert et al. 2016, Brouwer et al. 2016). Instead, we used a bootstrap approach.

Since the cluster bins of highest r-band luminosity, L_{200}, contain only a small number of clusters, covering only a small fraction of the KiDS-450 tiles, we cannot use the same bootstrap approach as Viola et al. (2015) and Dvornik et al. (2017) by bootstrapping 1 deg^2 tiles with replacement. Instead, we bootstrapped the source catalog, in accordance with Bellagamba et al. (2019).

This means that we are not sensitive to the clustering effect of dark matter halos, which justifies our choice of radial lens-source separation mentioned above. To assess the accuracy of these assumptions, we estimated the covariance matrix of the full 6925 cluster sample by bootstrapping the sources and by bootstrapping by KiDS-450 tiles in Appendix 5.A.4. We conclude that our bootstrapping method yields a good estimate of the covariance matrix.

Halo model fitting

Having produced the LSQ and LAD shear profiles for the stacked clusters, we fit a halo model to the results. We used the fitting procedure described in Dvornik et al. (2017), producing the full posterior probabilities by a Bayesian inference technique, via a Monte Carlo Markov chain (MCMC) maximum likelihood approach. We assumed a Gaussian likelihood and made use of the full covariance between radial bins:

$$\mathcal{L} \propto \exp\left[-\frac{1}{2}\boldsymbol{R}^{\rm T}\boldsymbol{C}^{-1}\boldsymbol{R}\right]\,, \tag{5.31}$$

where the $\boldsymbol{R}$ are the residuals and $\boldsymbol{C}$ is the covariance matrix.

We used the `emcee` Python package (Foreman-Mackey et al. 2013) for the MCMC procedure, setting flat priors for all parameters. For the evaluation of the power spectrum and the halo mass function, we used the median redshift for each cluster luminosity bin.

5.4 Results

We present our results, starting with the derived ESD profiles obtained with the mean and LAD estimators, discussing potential biases and efficiency. Then, we show the results of the halo model fitting (i.e., M_{200} for each luminosity bin) and conclude with the scaling relation between L_{200} and M_{200}. We visualize the results for the case in which all 6925 clusters are stacked together, giving the numerical results of the 13 luminosity bins in Table 5.3.

5.4.1 ESD profiles

We calculated the ESD profiles using 10^4 bootstraps with replacement. We estimated the ESD signal in the 14 radial bins, using both the mean and the LAD estimators, for each bootstrapped sample, preserving the bootstrap order of all 28 values throughout the whole process.

We find the estimator distribution to be almost perfectly normal, as expected from the central limit theorem. The correlation between the 14 bins of the full stack of clusters is shown in Fig. 5.6 and is given by

$$\rho_{ij} \equiv \frac{\mathrm{Cov}_{ij}}{\sigma_i \sigma_j}\,, \tag{5.32}$$

where i and j denote the radial bin subscripts.

The upper-left part of the matrix shows the correlation between the LAD estimates of the radial bins, and the lower-right part the shows the mean results. Although the correlation between bins is very low, it is clear that the overall trends are the same for the two estimators.

The signal-to-noise ratio of the recovered ESD profile of the full stack, which is shown in Fig. 5.7, is high enough to allow us to notice the difference between the estimators, which indicates a small relative bias. The blue points show the LAD estimates, and the red points represent the mean estimates, with error bars in both cases defined as the square root of the diagonal elements of the covariance matrices (i.e., the classical standard deviation).

Tests for systematic effects, such as the cross signal, and a test for systematic additive noise around random points were already conducted by Dvornik et al. (2017) and Bellagamba et al. (2019). In Appendix 5.A we repeat these tests for completeness since we use the KiDS-S field and an extended source selection with respect to Dvornik et al. (2017) and use a different source selection and three smaller radial bins with respect to Bellagamba et al. (2019). Our results show no residual systematic effects, in accordance with these papers.

5.4.2 Bias and efficiency

A possible bias is expected to depend on the strength of the underlying shear field since a zero lensing signal would imply no bias. In that case, the expected relevant distributions, tangential ellipticities or noise, are symmetric around zero ellipticity.

To quantify the difference between the ESD estimates, which we call the relative bias, $\Delta\Sigma_{\mathrm{Mean}} - \Delta\Sigma_{\mathrm{LAD}}$, we assumed[13] to first order

$$\Delta\Sigma_{\mathrm{Mean}} - \Delta\Sigma_{\mathrm{LAD}} = m \cdot \Delta\Sigma\,, \tag{5.33}$$

[13]This assumption is only made here to quantify the bias and is not used elsewhere in the paper.

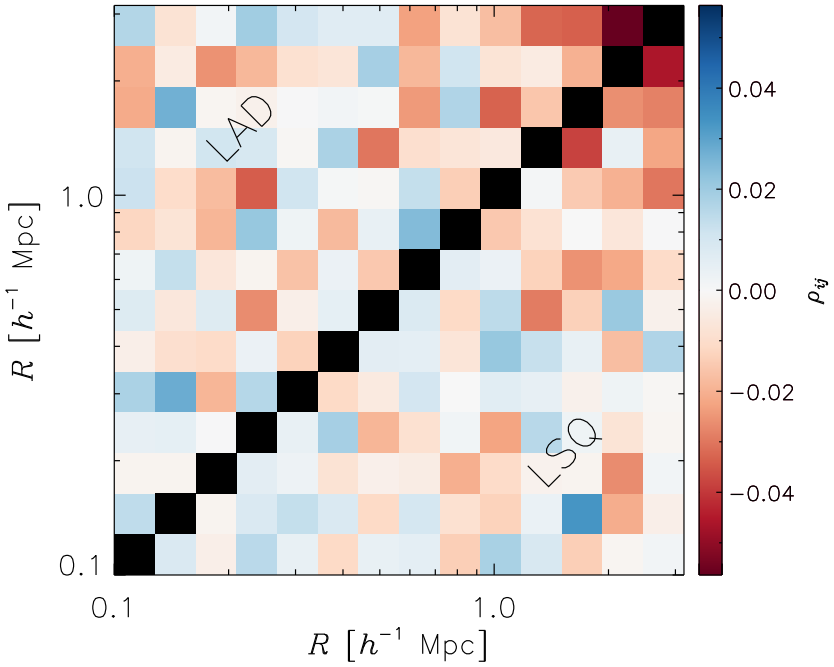


Figure 5.6: Correlation matrix between the stacked ESD signals in different radial bins, using the full AMICO cluster catalog and 10^4 bootstraps. The upper-left triangle shows the correlation, ρ_{LAD}, between the LAD estimations. The lower-right triangle shows the correlation, ρ_{mean}, between the ESD estimates using a weighted mean. We note the general similarities in the two patterns.

where we arbitrarily[14] use Σ_{LAD} for Σ.

We used the full stack for its high signal-to-noise ratio, using the 15.9th and 84.1th percentiles of the differences in all bootstrap results to calculate uncertainties for each bin. We find $m = -0.088 \pm 0.020$. In Fig. 5.8 we show this relative bias, plotting for visualization purposes

$$\frac{\Delta\Sigma_{\mathrm{Mean}} - \Delta\Sigma_{\mathrm{LAD}}}{\Delta\Sigma} \tag{5.34}$$

and a horizontal line at $m = -0.88$ to give a more intuitive impression of the relative error bars.

We reiterate that it is impossible to determine the absolute bias of each estimator as we did in Paper I, as we have no knowledge of the true ESD. However, the overall trend between the mean and LAD is similar in sign and order of magnitude, as we found for the CFHTLenS data in Paper I.

In Fig. 5.9 we show the derived relative efficiency (Eq. 5.26) $\eta = 1.047 \pm 0.006$, which is in accordance with the findings for CFHTLenS data in Paper I. The measured[15] ellipticity distribution is expected to differ for shape measurements with a higher signal-to-noise ratio.

[14] We find no qualitative difference in our results when we use $\Delta\Sigma_{\mathrm{Mean}}$ instead.

[15] i.e., the combination of the intrinsic distribution and the various effects before and after the lensing by AMICO clusters, which affects the observation and measurement of the source ellipticities.

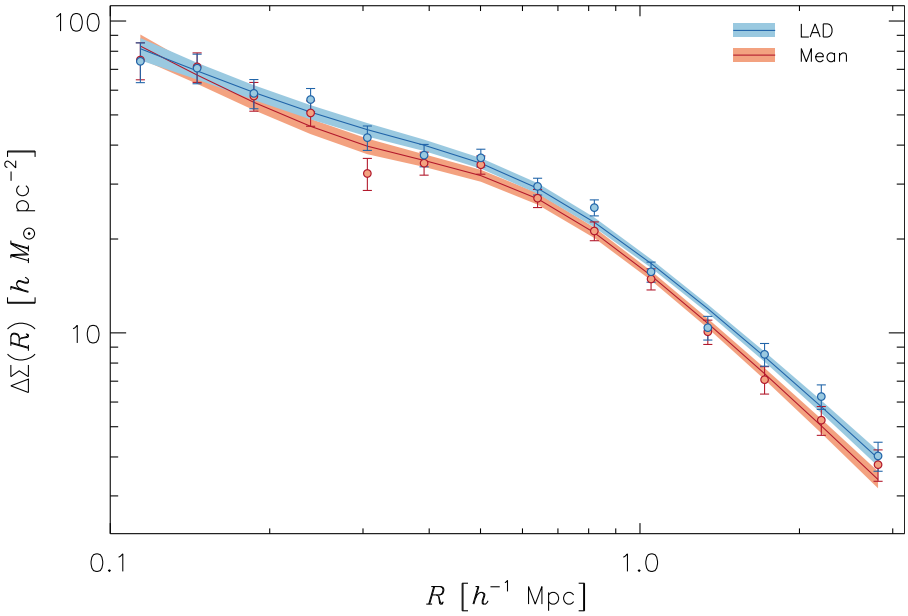


Figure 5.7: Estimated ESD profile from the full AMICO cluster catalog, using the LAD estimator (blue) and a weighted mean (red). The error bars are the square roots of the diagonal values of the respective covariance matrices. The solid lines represent the best fitting halo model obtained by the MCMC fit. The shaded regions show the 68.3% confidence bands, estimated using the 15.9th and 84.1th percentiles of the MCMC realizations.

For example, the cuspiness of the distribution shown in Fig. 5.2 can be smoothed out by noise convolutions. As the LAD estimation is more sensitive to the central peak, this will affect its precision.

In Paper I this was confirmed in the comparison of simulated data with and without noise, as well as in the results of the CFHTLenS sample with a stringent signal-to-noise selection compared with the full sample. As in Paper I, we compared the relative efficiencies for our selections with $\nu_{\rm SN} \geq 20$ and with $w_i \geq 14.5$, finding indeed a higher efficiency for less noisy shapes, namely $\eta = 1.240 \pm 0.010$ and $\eta = 1.386 \pm 0.018$, respectively.

5.4.3 Halo masses

We ran MCMC chains of 120 000 samples, using 120 walkers with 1000 steps each. The resulting chains were fully converged after the first 200 steps, so we discarded the first 24 000 samples.

We summarize the M_{200} derived from the ESD estimation of the 13 luminosity bins in Table 5.3. Reduced χ^2, estimated between 0.730 and 2.528, are fairly consistent between derived results for mean and LAD.

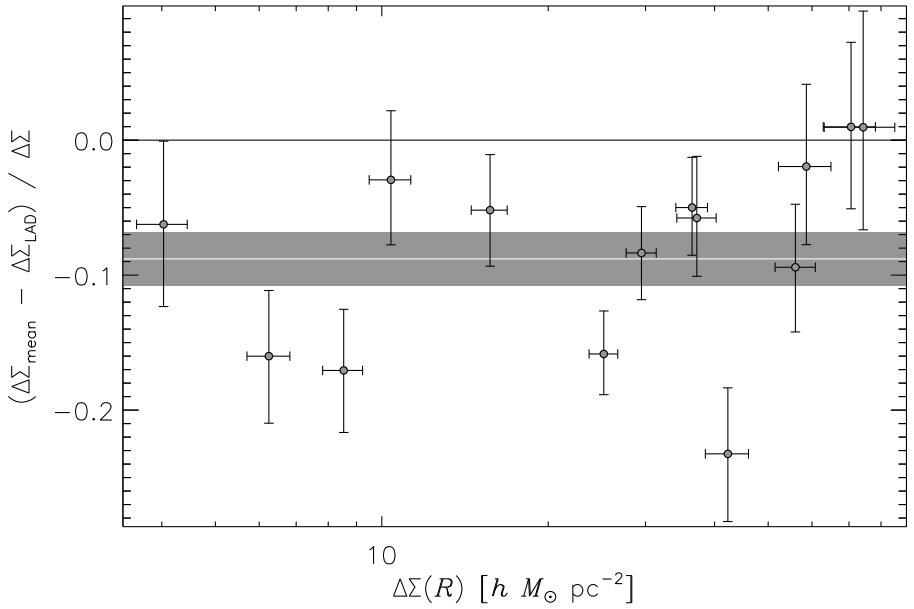


Figure 5.8: Difference between the recovered ESD signals in the radial bins, showing $(\Delta\Sigma_{\rm mean} - \Delta\Sigma_{\rm LAD})\,/\Delta\Sigma$. The solid white line represents the average difference, with the shaded region showing the formal 1σ error. As a possible bias in the recovered values is expected to increase with increasing shear, the differences are plotted versus the full ESD signal in each bin where we use $\Delta\Sigma_{\rm LAD}$, but we note that the small variations in the individual points, when plotting against $\Delta\Sigma_{\rm mean}$ instead, give the same result, within statistical significance.

For the full stack of clusters, we derived

$$M_{200} = \left(0.453^{+0.030}_{-0.030}\right) \times 10^{14} h^{-1} M_\odot\,, \quad \chi^2_\nu = 1.25 \qquad \text{(Mean)} \tag{5.35a}$$

$$M_{200} = \left(0.487^{+0.033}_{-0.036}\right) \times 10^{14} h^{-1} M_\odot\,, \quad \chi^2_\nu = 1.37 \qquad \text{(LAD).} \tag{5.35b}$$

The confidence intervals are derived from the 15.9th and 84.1th percentiles of the posterior distributions.

The best fitting ESD models are shown in Fig. 5.7. The 68.3% confidence bands overlap at some radii and are in tension at other radii. While the difference in ESD is significant, the 68.3% confidence intervals for M_{200} just touch.

5.4.4 $L_{200} - M_{200}$ scaling relation

We assumed a power-law relation between the derived halo masses and the median r-band luminosity of each cluster bin. We fit this relation in the form

$$\log\left(\frac{M_{200}}{M_{\rm piv}}\right) = a + b \log\left(\frac{L_{200}}{L_{\rm piv}}\right), \tag{5.36}$$

with a the intercept and b the slope, where $M_{\rm piv} \approx 10^{14.1} h^{-1} M_\odot$ and $L_{\rm piv} \approx 10^{11.8} h^{-2} L_\odot$ are typical pivotal values of the halo mass and luminosity, derived from the fit itself. The fit was

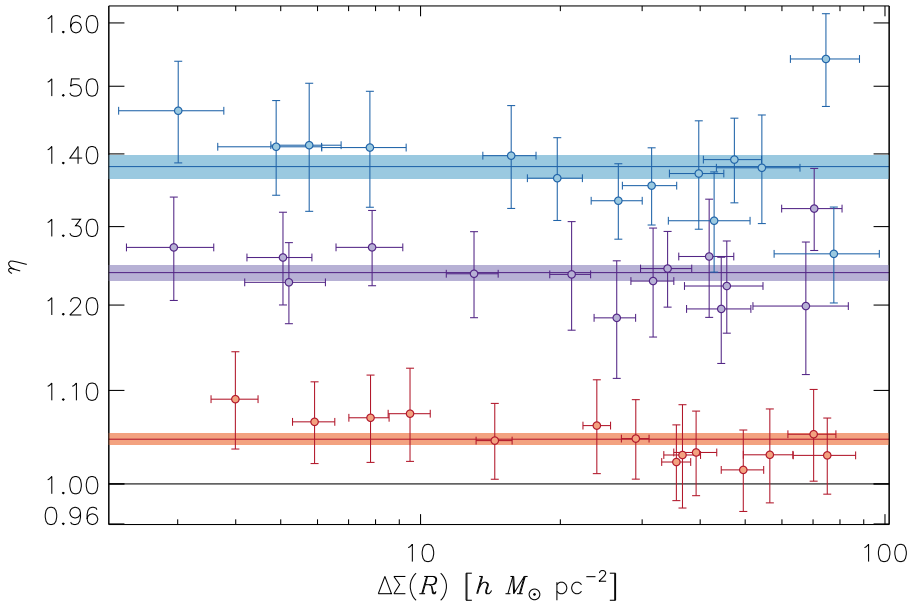


Figure 5.9: Relative efficiencies, $\eta \equiv \sigma^2_{\rm Mean}/\sigma^2_{\rm LAD}$, defined as the ratio of the diagonal elements of the covariance matrices. The relative efficiencies using the full KiDS-450 source catalog are shown in red. The purple and blue represent the higher S/N and higher *lens*fit weight selections, respectively. The solid lines represent the average η for each selection, with the shaded regions showing the 1σ errors.

done in log basis as the derived posterior distributions of the halo mass are log-normal.

We did not take a redshift dependence into account, as Bellagamba et al. (2019) showed only a marginal and not very steep dependence of the halo mass on redshift.

We obtained the scaling relations

$$\frac{M_{200}}{10^{14.1}h^{-1}M_\odot} = (0.97 \pm 0.06)\left(\frac{L_{200}}{10^{11.8}h^{-2}L_\odot}\right)^{(1.24\pm0.08)} \quad \text{(Mean)} \tag{5.37a}$$

$$\frac{M_{200}}{10^{14.1}h^{-1}M_\odot} = (1.03 \pm 0.05)\left(\frac{L_{200}}{10^{11.8}h^{-2}L_\odot}\right)^{(1.24\pm0.08)} \quad \text{(LAD)} \tag{5.37b}$$

and plot the results in Fig. 5.10. As with the derived ESD profiles and halo masses for the full stack of clusters, the 68.3% confidence bands just touch at the pivot point

$$(L_{200}, M_{200}) = \left(10^{11.8}h^{-2}L_\odot, 10^{14.1}h^{-1}M_\odot\right), \tag{5.38}$$

recognizable as the narrowest parts of the confidence bands.

5.5 Summary and conclusions

We conducted a weak shear analysis of 6925 AMICO clusters in the KiDS-450 data We derived a tightly constrained scaling relation between r-band luminosity, L_{200}, and average lensing masses, M_{200}, in concordance with earlier results in the literature.

Table 5.3: Properties and lensing results of the individual luminosity bins, with the median L_{200} and z_l. The errors on the median L_{200} and derived M_{200} in each luminosity bin are the differences with the 15.9th and 84.1th percentiles. L_{200} values are given in $\left[10^{10}h^{-2}L_\odot\right]$, M_{200} in $\left[10^{14}h^{-1}M_\odot\right]$.

Range	L_{200}	Clusters	z_l	M_{200} Mean	χ^2_ν	M_{200} LAD	χ^2_ν
$[0.4, 17.4[$	$12.3^{+3.5}_{-4.6}$	2346	0.29	$0.187^{+0.029}_{-0.026}$	1.27	$0.191^{+0.031}_{-0.026}$	1.18
$[17.4, 24.8[$	$20.8^{+2.7}_{-2.3}$	1545	0.41	$0.416^{+0.061}_{-0.076}$	2.19	$0.445^{+0.055}_{-0.061}$	1.60
$[24.8, 31.8[$	$28.0^{+2.5}_{-2.2}$	1027	0.41	$0.371^{+0.048}_{-0.046}$	0.81	$0.385^{+0.050}_{-0.046}$	0.91
$[31.8, 40.5[$	$35.2^{+3.2}_{-2.4}$	685	0.42	$0.519^{+0.078}_{-0.078}$	1.34	$0.621^{+0.076}_{-0.074}$	1.29
$[40.5, 49.0[$	$44.4^{+2.9}_{-2.7}$	457	0.42	$1.076^{+0.182}_{-0.181}$	2.18	$0.998^{+0.150}_{-0.139}$	2.53
$[49.0, 59.9[$	$54.1^{+3.6}_{-3.9}$	305	0.40	$1.387^{+0.425}_{-0.335}$	1.05	$1.462^{+0.533}_{-0.353}$	0.98
$[59.9, 72.9[$	$65.2^{+5.2}_{-3.7}$	202	0.41	$1.318^{+0.267}_{-0.234}$	0.77	$1.309^{+0.233}_{-0.203}$	0.74
$[72.9, 84.1[$	$78.6^{+3.2}_{-3.9}$	135	0.38	$1.406^{+0.344}_{-0.215}$	1.22	$1.528^{+0.409}_{-0.254}$	1.30
$[84.1, 102[$	$91.8^{+5.7}_{-4.8}$	90	0.39	$2.438^{+0.486}_{-0.443}$	0.75	$2.472^{+0.479}_{-0.425}$	0.73
$[102, 129[$	$112^{+10}_{-7.2}$	60	0.40	$2.143^{+0.618}_{-0.417}$	0.77	$1.914^{+0.680}_{-0.373}$	0.79
$[129, 160[$	$138^{+11}_{-6.9}$	40	0.395	$3.999^{+1.957}_{-0.993}$	1.35	$3.715^{+1.557}_{-0.910}$	1.64
$[160, 221[$	175^{+27}_{-11}	26	0.37	$4.207^{+0.713}_{-0.610}$	1.47	$4.786^{+0.824}_{-0.637}$	1.18
$[221, 400[$	277^{+106}_{-24}	8	0.375	$7.638^{+2.293}_{-1.613}$	1.02	$9.141^{+3.105}_{-1.913}$	0.81

We investigated the impact of estimator choice for inferring the central moment of the cusped and skewed ellipticity distribution of background galaxies, finding a relative bias on the order of a few percent, as predicted in Paper I. We find that the constraints obtained via LAD regression are tighter than those obtained via LSQ regression, and they significantly improved as the signal-to-noise ratio of the shape measurements of the background galaxies increased. Complemented by simulations from Paper I, we give an alternative perspective on the problem of inferring the central shear value from the skewed distribution of background galaxy shapes, at the minor cost of increased, but still feasible, computation times for numerical iterative regression.

5.5.1 $L_{200} - M_{200}$ relation

Since the relative bias we found in both this research and in Paper I is approximately proportional to the ESD signal, it is expected that the LAD estimator will mainly have an effect on the intercept of the $L_{200} - M_{200}$ scaling relation. This was confirmed by our results in Eq. 5.37.

The power-law index of the $L_{200}-M_{200}$ scaling relation was constrained to 1.24 ± 0.08 (Eq. 5.37), independent of estimator choice. This is in agreement with earlier work in the literature, such as Viola et al. (2015, and references therein), who cite 1.16 ± 0.13. This agreement is noteworthy since the AMICO clusters are derived from photometric redshifts, as opposed to the spectroscopically derived groups from GAMA (Driver et al. 2011, Robotham et al. 2011). The difference in confidence is explained by the increased number of lenses, which is

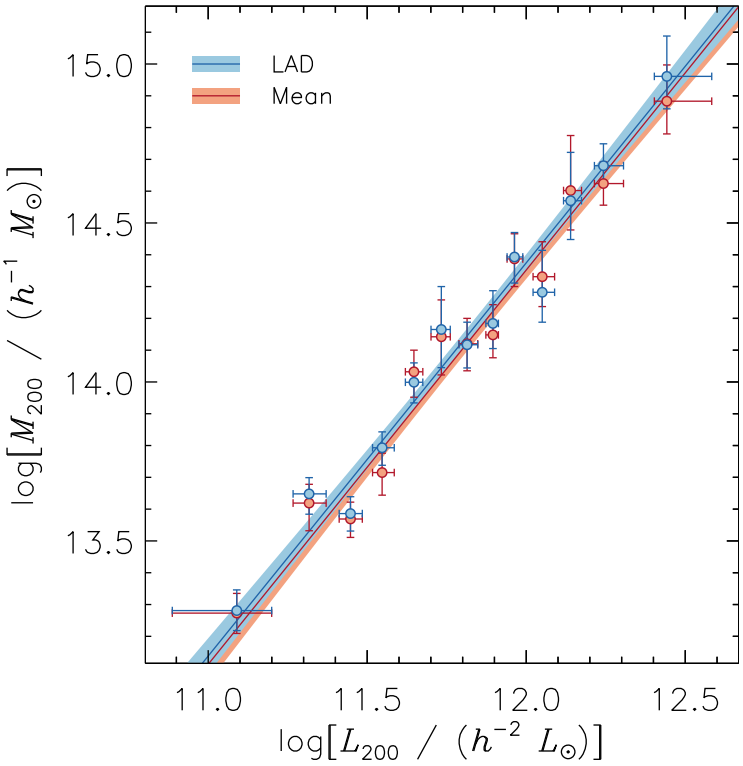


Figure 5.10: r-band luminosity-halo mass scaling relations, derived from the ESD profiles estimated using the weighted mean (red) and LAD (blue). At the pivot point, recognizable as the narrowest parts of the confidence bands, the relations just touch at the 68.3% confidence level.

not surprising given the large overlap in mass range: Viola et al. (2015) analyzed 1413 galaxy groups between $z = 0.03$ and $z = 0.33$, with r-band luminosity bin limits between 2.5×10^9 and $5.0 \times 10^{12}\ h^{-2} L_\odot$, deriving halo masses between 1.4×10^{13} and $3.5 \times 10^{14}\ h^{-1} M_\odot$.

The choice of estimator produces a difference in intercept at just the 1σ level. Using a weighted mean to derive the ESD leads to an intercept of 0.97 ± 0.06 in the scaling relation, while LAD gives an intercept of 1.03 ± 0.05. This is to be expected as the relative bias seems roughly constant, when normalized by the ESD (see Fig. 5.8), and in good agreement with the bias of $\sim 5\%$ found using simulations in Paper I.

As in Bellagamba et al. (2019), we estimate that systematic effects mostly affect the intercept. While the derived intercept is in agreement with the aforementioned papers, we note that the chosen definition[16] of L_{200}, the difference in redshift range and definitions, and the completeness of group and cluster membershipcan account for possible differences on the same order of magnitude. This would not affect our conclusions on methodology, as the results from both estimators would be similarly affected.

In a further comparison with the scaling relation between richness and mass, cited in

[16] e.g., Viola et al. (2015), where the group r-band luminosities are calculated by summing over spectroscopically confirmed group members.

Bellagamba et al. (2019), we find similar significance in constraints on the slope. We define bins in luminosity instead of AMICO detection amplitude (Radovich et al. 2017) or richness, but since the cluster luminosity is tightly correlated with the richness, λ^*, this confirms our findings.

There are a few differences to consider. We chose not to account for a possible redshift dependence. This is motivated by Bellagamba et al. (2019) finding only a shallow dependence, which they point out may be driven mainly by the highest redshift bin. Our lensing analysis employs a slightly different background selection for an increased source density, combined with a different derivation of the associated redshift distribution, $n(z)$. Another difference is the inclusion of radial bins at $0.1 \leq R < 0.2$ Mpc h^{-1}. This is expected to only have a minor effect as the contribution of the stellar mass is an order of magnitude lower than the halo term at these radii, while the contribution of miscentered halos only starts to become significant at larger radii (see also Rykoff et al. 2016, Oguri et al. 2018). In this sense, these findings are a confirmation of the robustness of the results across these papers.

5.5.2 Optimal estimators

Our results are in good agreement with Paper I, with a relative bias between the two estimators that shows the recovered lensing signal is higher with LAD, suggestive of a lower absolute bias. At the same time, LAD regression gives a small (albeit significant) gain in efficiency, giving a reduction in error bars of a few percent, and potentially up to $11\% - 18\%$ for shape measurements of a higher signal-to-noise ratio. Least absolute deviation regression comes at the cost of a higher computation time, but at a step in the analysis process that does not dominate the total computational cost.

Both simulations (Paper I) and analyses on real data (this article) cite quantitative results of significance while at the same time showing similar trends between estimators on a qualitative level. We have conducted a cautious and thorough investigation but can never exclude the unknown: biases arising due to assumptions in the simulations of Paper I or uncorrected systematic effects in this research, or, most likely, both. However, given the range and realism in simulated distributions and the similarities in findings among those simulations, this research, and other work in the literature, we are confident that the recovered differences in results between the two estimators are real.

We note some differences between the two analyses. In Paper I we analyzed the regression of a sample of ellipticities with a single underlying value of the shear and, for each type of simulation, a single intrinsic ellipticity distribution. In this research, the situation is more complex. We studied the stacked signal around samples of lenses and of samples of background sources at a range of redshifts. This means also stacking noise that has been scaled by a range in lensing geometries, quantified by $\Sigma_{\rm cr}^{-2}$. Furthermore, in each radial bin of each luminosity bin, we assume: (i) a constant lensing effect, which is in reality the stacked average of a range in L_{200}, and therefore a range in M_{200}, confounded by intrinsic scatter between these two quantities, and (ii) radial distance R from the lens, combined with a miscentering of halos.

All these effects tend to convolve the intrinsic galaxy shape distribution, which makes the level of agreement and significance between the two papers in fact remarkably robust. In conclusion:

- The combination of Paper I and this research shows that LAD regression is more naturally suited to the cusped intrinsic ellipticity distribution of background galaxies.
- Our simulations in Paper I showed a lower bias for LAD regression than for LSQ regression in the presence of noise in the background source shape measurements, while this research confirmed the same relative bias between the two estimators.
- Constraints obtained via LAD regression are comparable with or tighter than constraints obtained via LSQ regression.

An optimal estimator is, from a principled point of view, more objective and better suited than corrections to an approach, which is known to mismatch the sample distribution. More practically, LAD regression provides a robust consistency check for shear inference, which has been and still remains a major investment in the field of weak lensing. Keeping different perspectives, such as exploring these alternative statistical approaches, is fundamental for determining the way forward.

Acknowledgements

MSm acknowledges support from the Netherlands Organization for Scientific Research (NWO). AD acknowledges ERC Consolidator Grant (No. 770935). LM acknowledges the grants ASI-INAF n. 2018-23-HH.0 and PRIN-MIUR 2017 WSCC32 "Zooming into dark matter and proto-galaxies with massive lensing clusters". MSe acknowledges financial contribution from contract ASI-INAF n.2017-14-H.0 and contract INAF mainstream project 1.05.01.86.10.

Based on data products from observations made with ESO Telescopes at the La Silla Paranal Observatory under programme IDs 177.A-3016, 177.A-3017 and 177.A-3018, and on data products produced by Target/OmegaCEN, INAF-OACN, INAF-OAPD and the KiDS production team, on behalf of the KiDS consortium. OmegaCEN and the KiDS production team acknowledge support by NOVA and NWO-M grants. Members of INAF-OAPD and INAF-OACN also acknowledge the support from the Department of Physics & Astronomy of the University of Padova, and of the Department of Physics of Univ. Federico II (Naples).

5.A Tests for systematics

For completeness, we repeated some of the tests for systematics that were already carried out in Dvornik et al. (2017) and Bellagamba et al. (2019) because of our difference in sky coverage, background selection, and estimated redshift distribution compared with those two studies.

5.A.1 Photometric redshift

We used the same method as Dvornik et al. (2017) to determine the co-moving critical density. There are two important differences that could affect the uncertainty in $\Sigma_{\rm cr}$: We selected lenses at a significantly higher redshift, and we complemented our background source selection with the color selection described in Sect. 5.3.2.

We assessed the relative errors in $\Sigma_{\rm cr}$ by performing 10^4 bootstraps of the spectroscopic catalog of Hildebrandt et al. (2017). We find the median error on $\Sigma_{\rm cr}$ to be $\sim 0.5\%$, as shown in Fig. 5.11.

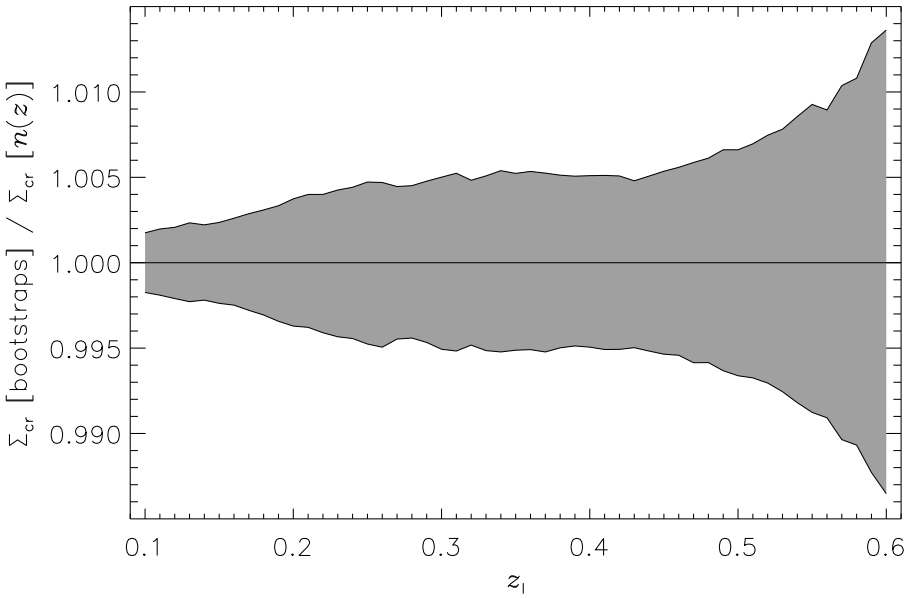


Figure 5.11: Relative errors in $\Sigma_{\rm cr}$, estimated using 10^4 bootstraps of the spectroscopic catalog of Hildebrandt et al. (2017). The relative error on the ESD is negligible.

5.A.2 Contamination of the background sample by cluster galaxies

Dvornik et al. (2017) showed that an offset of $\Delta z = 0.2$ is enough to avoid a significant contamination of the background sources by unidentified GAMA group members. For lenses at a higher redshift, this contamination increases, while at the same time the density of available background sources decreases due to the observed depth of KiDS-450.

We used the same test as Dvornik et al. (2017) to assess the source density around AMICO clusters in order to determine the necessary Δz offset between the lens and the sources. We find that $\Delta z = 0.2$ is appropriate for our cluster selection (Fig. 5.12).

5.A.3 Individual bin ESD profiles and cross signals

In Fig. 5.14 we show the ESD profiles for the 13 cluster bins, including the cross signal, which is consistent with zero. We also show the derived halo model fits and their confidence intervals, comparing the fits using the full AMICO cluster catalog from Fig. 5.7.

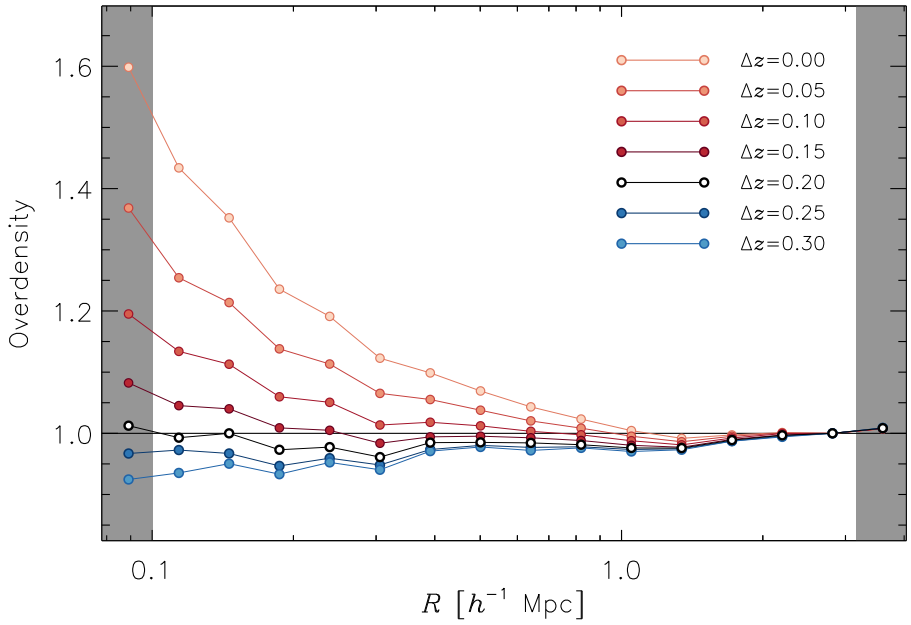


Figure 5.12: Relative source densities around AMICO clusters as a function of radius, R, and the photometric redshift offset, $z_B \geq z_l + \Delta z$, between the lens and the source. We note that some small under-densities around $R = 0.2\,h^{-1}$ Mpc may be due to the relative normalization.

5.A.4 Tile bootstrap

As described in Sect. 5.3.3, we could not use the same bootstrap approach as Dvornik et al. (2017), due to the sparsity of lenses in the highest lens luminosity bins. Since our bootstrap approach described in Sect. 5.3.3 does not account for cosmic variance and is not sensitive to the clustering effect of dark matter halos, we compare the covariances derived by the two bootstrap methods for the ESD of the whole lens selection in Fig. 5.13 and find no significant differences or pattern beyond what is expected from statistical noise. Since we expect the contribution from cosmic variance to be even lower for subsets of lenses, we conclude that our bootstrap approach yields a good estimate of the covariance.

5.B Analysis of dependence on outer data points

In Fig. 5.10 it can be seen that the distribution of clusters in the two outermost luminosity bins is not symmetric. At the lower end, this is due to the selection criterion of λ^* in the AMICO catalog. At the higher end, we have only a few clusters.

We assessed the effect these two points have on the $L_{200} - M_{200}$ scaling relation by repeating the fit without these bins. We find no difference within the statistical uncertainties, as given in Eq. 5.39:

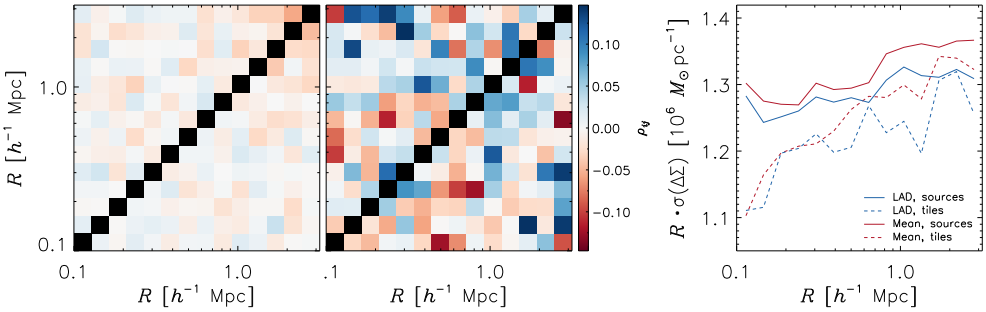


Figure 5.13: Left: Correlation derived from bootstrapping the signal by individual sources; same as Fig. 5.6, but with the color stretch adjusted to the middle plot. Middle: Correlation derived from bootstrapping the signal in 1 deg^2 tiles. The upper-left corners show the correlations from LAD regression. The lower-right corners show the correlations from using the weighted mean. Right: Comparison of the errors obtained from bootstrapping sources (LAD: solid blue; mean: solid red) and bootstrapping 1 deg^2 tiles (LAD: dashed blue; mean: dashed red).

$$\frac{M_{200}}{10^{14.1}h^{-1}M_{\odot}} = (0.98 \pm 0.06)\left(\frac{L_{200}}{10^{11.8}h^{-2}L_{\odot}}\right)^{(1.25\pm0.10)} \quad \text{(Mean)} \tag{5.39a}$$

$$\frac{M_{200}}{10^{14.1}h^{-1}M_{\odot}} = (1.02 \pm 0.06)\left(\frac{L_{200}}{10^{11.8}h^{-2}L_{\odot}}\right)^{(1.23\pm0.09)} \quad \text{(LAD)} \tag{5.39b}$$

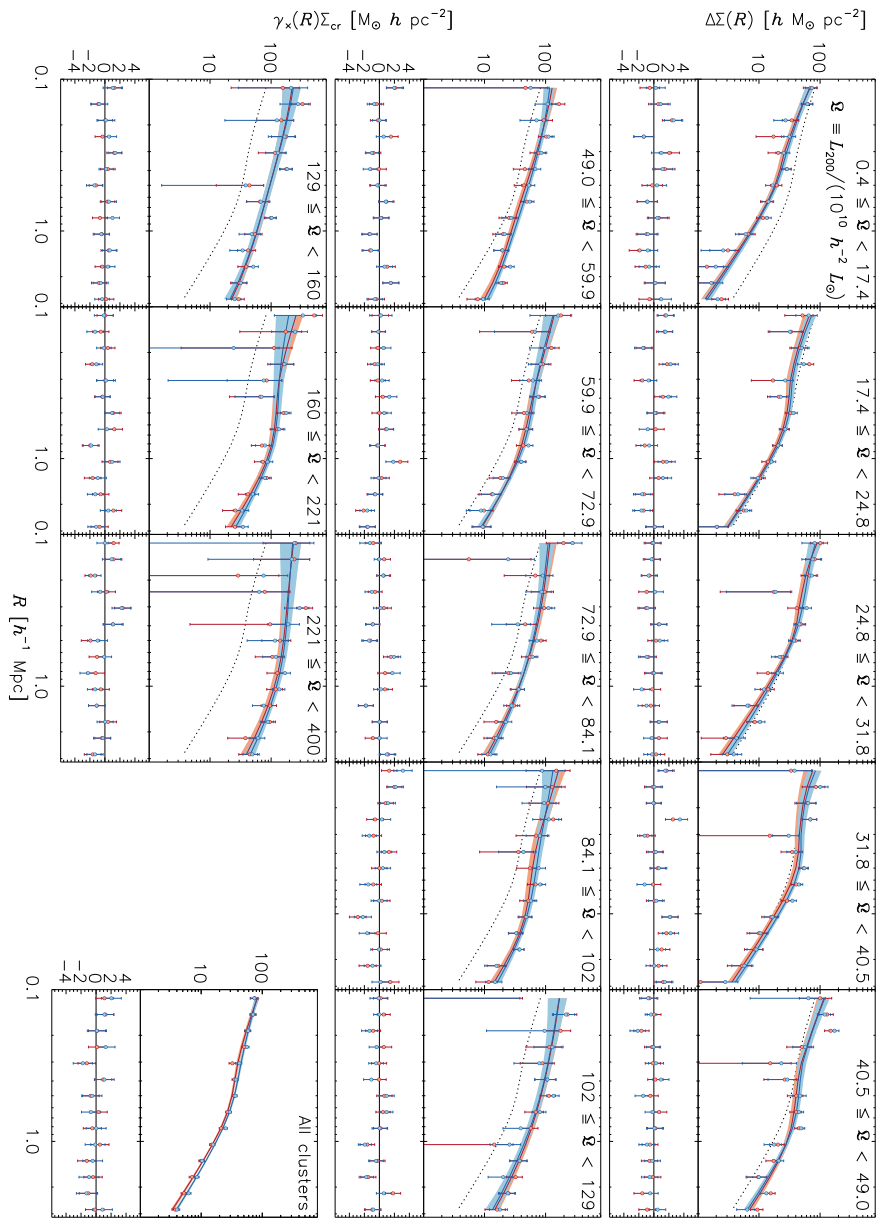


Figure 5.14: Estimated ESD profiles and cross signal of the 13 cluster bins, using the LAD estimator (blue) and a weighted mean (red). The error bars are the square roots of the diagonal values of the respective covariance matrices. The solid lines represent the best fitting halo model obtained by the MCMC fit. The shaded regions show the 68.3% confidence bands, estimated using the 15.9th and 84.1th percentiles of the MCMC realizations. Individual bins are indicated by the range in normalized luminosity, $\mathfrak{L} \equiv L_{200}\,/(10^{10}h^{-2}L_\odot)$. The lower-right plot shows the ESD profile estimated from the full AMICO cluster catalog, also shown in Fig. 5.7. The average of the two best fitting halo models from Fig. 5.7 are shown in each panel as a dotted line for easy comparison.

6

Nederlandse samenvatting

"... en zelfs, als hij laat zien dat alle dingen één zijn door deel uit te maken van eenheid en dat hetzelfde ook veel is door deel uit te maken van veelheid, lijkt me dat helemaal niet vreemd; maar als hij laat zien dat de absolute eenheid ook veel is en de absolute veelheid weer één, dan zal ik oprecht verbaasd zijn."

– Socrates tegen Zeno in Plato's dialoog Parmenides, 129B.

In zijn Ideeënleer[1] beargumenteert Plato dat onze zintuigelijke waarnemingen slechts mogelijke reflecties of gestalten zijn van een onwaarneembare en unieke universele Vorm.

Voor sterrenkundigen is dit de dagelijkse realiteit: we zijn beperkt tot de informatie die ons op aarde bereikt in de vorm van elektromagnetische straling – zoals licht – of zwaartekrachtgolven. We kunnen in onze directe omgeving – ons zonnestelsel – kijken, maar we kunnen de Melkweg niet vanaf de andere kant observeren of een zwart gat in een ander sterrenstelsel opzoeken, metingen doen en weer terugkomen. Onze blik op de hemel is een momentopname in vergelijking met de miljoenen of zelfs miljarden jaren durende evolutie van objecten in het heelal.

Dat betekent dat ons perspectief uniek is, maar ook heel beperkt. Het is onze uitdaging om uit de observaties die we kunnen doen, de indirecte informatie te halen over alles wat we niet direct kunnen waarnemen. De grote vraag is dan: hoe weten we dat onze conclusies juist zijn?

In dit proefschrift combineren we daarom het onderzoek naar de verdeling van materie in groepen en clusters van sterrenstelsels met analyses van de systematische en statistische onzekerheden van onze onderzoeksmethoden. Zien we echt wat we denken te zien?

[1] Wellicht beter genaamd "vormenleer" of "gestaltenleer"

6.1 "De hemel bij nacht"

Voor wie op een heldere nacht voldoende lang naar het firmament kijkt[2] en steeds nieuwe sterren ontdekt, terwijl de hemel langzaam voorbij draait, zal het makkelijk voor te stellen zijn hoe haar aanblik de mensheid reeds duizenden jaren heeft gefascineerd.

Reeds vanaf het begin hebben mensen geprobeerd hun waarnemingen te verklaren en betekenis te geven. Eén van de oudste voorbeelden is het verhaal van de Plejaden, het Zevengesternte, dat in de Griekse mythologie het verhaal vertelt van zeven zusters, terwijl er maar zes sterren zichtbaar zijn voor de meeste mensen. Uit oude teksten, zoals van Aratus van Soli uit de 3^{e} eeuw voor Christus, blijkt dat de Grieken ook uitgingen van zes zichtbare sterren. Het verhaal lijkt daar ook op aangepast: de verklaring voor het aantal van zes zichtbare sterren is, dat één van de zusters (Merope) zich verstopt heeft voor de jager Orion. Maar waarom de associatie van zes sterren met een verhaal over zeven zusters? Een veel gebezigde aanname is dat het hier een *cognitieve vertekening* of *cognitieve bias* betreft: namelijk dat men, door het verhaal aan te passen, als het ware naar het eigen verwachtingspatroon toe heeft geredeneerd.

Het merkwaardige feit is echter dat verschillende culturen over de hele wereld dit verhaal kennen in één of andere vorm en al deze overleveringen vasthouden aan zeven, waarvan er één verdwenen is, in plaats van een verhaal te vertellen over zes. Dit suggereert dat deze verhalen een gezamenlijke oorsprong kennen uit een tijd, zo'n 100,000 jaar geleden, waarin de zevende ster nog wel te zien was (Norris & Norris 2021). Dat men lange tijd is uitgegaan van een onvolkomenheid in de mythologieën, gebaseerd op de aanname dat onze eigen observaties van zes sterren een juiste weergave van de werkelijkheid waren, legt het probleem van cognitieve bias juist bij onze theorieën. Het laat zien hoe belangrijk het is om ons niet blind te staren op de beelden die we zien, maar ook alert te zijn op de dingen die wij, zelfs in de moderne tijd, juist *niet* zien.

6.1.1 Waarnemingen en vertekening

Wellicht het bekendste voorbeeld van hoe ons perspectief een drempel kan zijn in het vormen van juiste theorieën, is het Ptolemeïsche wereldbeeld met de aarde in het centrum, dat meer dan 1400 jaar heeft standgehouden. Toen steeds groter wordende onnauwkeurigheden de onhoudbaarheid van dit model duidelijk maakten (zie bijvoorbeeld Figuur 1.1), was Copernicus' stap naar een heliocentrisch model – met de zon in het centrum – een fundamentele verandering in het paradigma. Het was echter nog steeds niet genoeg.

Ook Copernicus bleef vasthouden aan een ideaalbeeld van volmaaktheid, namelijk bewegingen in perfecte cirkels. Dit vooroordeel leidde ertoe dat zijn voorgestelde aanpassingen niet tot overtuigende verbetering leidden. Pas toen Kepler zijn theorie van elliptische banen formuleerde – een tweede, maar net zo fundamentele stap – konden waarnemingen en theorie met elkaar in overeenstemming worden gebracht.

In dit geval betreft het een bias in het model, dat intrinsiek niet geschikt was om de werkelijkheid te beschrijven. Een andere bekende vertekening in de astronomie is de Malmquist bias (Malmquist 1925), veroorzaakt doordat heldere objecten makkelijker te detecteren zijn en dus op grotere afstand waar te nemen zijn. Objecten die makkelijker te detecteren zijn, lijken vaker voor te komen en zijn ook nauwkeuriger te bestuderen. Dit zou ertoe kunnen leiden

[2]Zoals een student sterrenkunde bij het gelijknamige waarneempracticum

dat een steekproef niet aselect is: een zogenoemde selectiebias. Een maatschappelijk voorbeeld is het probleem van de prominente focus in de geschiedenis van medisch onderzoek op vrijgezelle, witte mannen (Dresser 1992).

Het brede concept van bias, als discrepantie tussen een waarneming of interpretatie en de 'waarheid', komt regelmatig terug in dit proefschrift, waarin we kijken naar bias in de waarnemingen, in de modellen, zelfs in de uitkomsten van statistische berekeningen en onze interpretatie van die uitkomsten.

6.1.2 Donkere materie

Eén van de grootste mysteries in de sterrenkunde is het mogelijke bestaan van donkere materie: iets wat we niet direct kunnen zien, maar waarvan we wel indirect de effecten kunnen waarnemen.

Op macroscopische schaal zijn er twee soorten interacties in het universum: één gebaseerd op elektromagnetisme, die er onder andere voor zorgt dat wij kunnen waarnemen in het zichtbare licht, radiogolven of andere frequenties; en één gebaseerd op gravitatie, die onder andere terug te zien is in de bewegingen in het heelal. Op deze schaal kijken we vooral naar sterrenstelsels, verzamelingen van biljoenen sterren of meer, zoals onze eigen Melkweg. Deze sterrenstelsels hebben de neiging bij elkaar te blijven, van groepen[3] van enkele sterrenstelsels tot grote clusters van honderden stelsels of meer.

Aan het begin van de 20e eeuw kreeg men het vermoeden dat er materie moest zijn, die wel interacteerde via zwaartekracht, maar niet via elektromagnetisme. Deze materie is daardoor niet direct waar te nemen. De term 'donkere materie' is voor het eerst gebruikt door de Nederlandse sterrenkundige Kapteyn (1922). De eerste significante waarnemingen kwamen van Zwicky (1933) in het nabije Coma cluster van sterrenstelsels. Uit deze waarnemingen leken individuele sterrenstelsels door grote gravitatiekrachten rondgeslingerd te worden. De zichtbare materie in Coma had echter niet genoeg massa om deze gravitatie te veroorzaken: Zwicky concludeerde dat er meer dan 400 keer zoveel massa nodig was dan zichtbaar was, om de grote snelheden die hij observeerde te verklaren. Ook hij gebruikte hiervoor de term 'Dunkele Materie'.

Het duurde tot in de tweede helft van de 20e eeuw voordat het probleem van deze ontbrekende of onzichtbare materie systematisch en overtuigend werd aangetoond door Rubin (1983) uit de interne snelheden van sterrenstelsels. In de jaren daarna bleek dit ook uit waarnemingen van de kosmische achtergrondstraling, bijvoorbeeld met de Planck satelliet (Planck Collaboration et al. 2014). Deze achtergondstraling is een soort echo van de Big Bang. De patronen in deze gloed suggereren dat er veel meer materie nodig is om de structuur en samenhang van het universum te verklaren, dan direct kan worden waargenomen. Ook uit deze waarnemingen lijkt er een overweldigende meerderheid van zo'n 80% van alle materie te zijn die niet zichtbaar is.

Om uit waarnemingen van de hemel massa's te bepalen, zijn vaak modellen en aannamen nodig. Zo gaat men er vaak vanuit dat een cluster van sterrenstelsels in evenwicht is, om de waargenomen snelheden te vertalen naar een massaverdeling. In dit proefschrift gebruiken we een andere, onafhankelijke methode om de materie in kaart te brengen: gravitationele lenswerking.

[3]Onze eigen Melkweg vormt samen met het vergelijkbare Andromedastelsel en een aantal kleinere stelsels de zogenaamde 'Lokale Groep'.

6.2 Gravitatielenzen

Gravitationele lenswerking is voor het eerst waargenomen door Dyson et al. (1920) en is gevoelig voor alle massaverdelingen, onafhankelijk van de toestand van die massaverdeling (zoals evenwicht) of de eigenschappen van de materie (heet of koud, zichtbaar of donker). Deze methode doet geen aannamen over de dynamische evenwichtssituatie in clusters of de astrofysische processen die de waarneembare straling veroorzaken, maar is gebaseerd op de geometrie van het heelal.

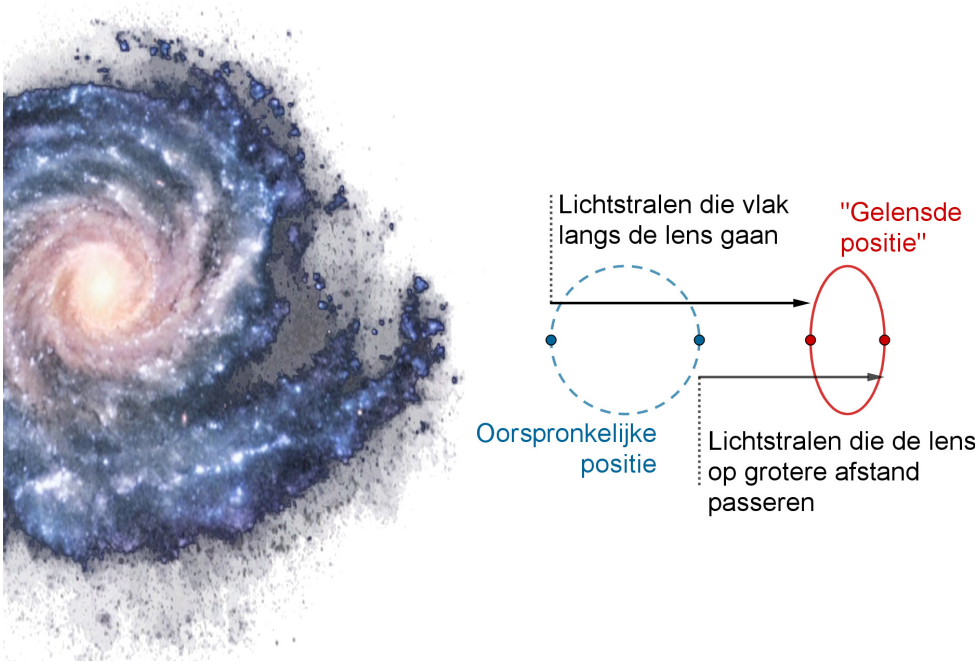
6.2.1 Kromming van ruimtetijd

Waar wij gravitatie ervaren als een 'kracht', wordt deze in de relativiteitstheorie van Einstein beschreven als een kromming van de ruimtetijd. Door deze kromming zijn er eigenlijk geen rechte lijnen in het universum. Ook lichtstralen (die geen massa hebben) worden daardoor beïnvloed door de "zwaartekracht" van objecten in de buurt van hun baan. Voor 'lenzen' wordt vaak gekeken naar sterrenstelsels of clusters van sterrenstelsels, vanwege hun grote massa en dus makkelijker meetbaar effect. Omdat wij het object dat de lichtstralen uitzond zien in de richting waar de lichtstralen vandaan leken te komen (zie ook Figuur 1.2), zorgt deze afbuiging voor een verplaatsing van objecten aan de hemel. In dit proefschrift beperken we ons tot zwakke lenswerking, waarbij dit effect zeer subtiel is en slechts statistisch waargenomen kan worden.

Bij zwakke lenswerking duwt deze subtiele verplaatsing het beeld van een bron in de achtergrond als het ware iets van het tussenliggende, 'lenzende' object vandaan (Figuur 6.1). Dit effect is logischerwijs het sterkst voor lichtstralen die dicht langs de lens gaan en minder voor lichtstralen die de lens op grotere afstand passeren. Als we naar objecten kijken die oorspronkelijk een bepaalde vorm hadden, zoals sterrenstelsels die in de achtergrond liggen, dan zien we dat het deel van het beeld dat dichter bij de lens ligt daardoor sterker verplaatst wordt en het deel dat verder van de lens ligt inhaalt: het beeld wordt als het ware iets platgedrukt, waardoor de ellipticiteit van het beeld wordt veranderd. Figuur 6.2 laat zien hoe deze vervorming aan de hemel een patroon zou kunnen veroorzaken in de ellipticiteiten van achtergrondbeelden.

De mate van deze vervorming kunnen we direct koppelen aan de massaverdeling van het lenzende object. Op deze manier kunnen we alle materie, zichtbaar en donker, in principe in kaart brengen, zonder dat we iets hoeven te weten of aan te nemen over de lens. Gravitationele lenswerking wordt daarom gezien als onafhankelijke en zeer belangrijke meetmethode voor het bestuderen van kosmologische modellen. Een beroemd voorbeeld is dat van het bullet cluster, weergegeven in Figuur 1.3. Het bullet cluster bestaat eigenlijk uit twee clusters, die met elkaar in botsing zijn gekomen. Hierdoor is het totale systeem duidelijk niet in evenwicht. Met behulp van Röntgenstraling zijn schokgolven te zien in het hete gas wat deze clusters meedroegen, maar door de botsing in het midden is achtergebleven (het linkerpaneel in Figuur 1.3). In het zichtbare licht zien we twee concentraties van sterrenstelsels, die als los zand door elkaar zijn gevlogen en dus verder van elkaar zijn verwijderd (het rechterpaneel in Figuur 1.3). Als we ons alleen op deze waarnemingen zouden baseren, zoals de aantallen sterrenstelsels en de intensiteit van het gas, zouden we de conclusie trekken dat het gas het grootste deel van de zichtbare massa vertegenwoordigt.

Door een reconstructie van de massaverdeling met behulp van zwakke lenswerking, de



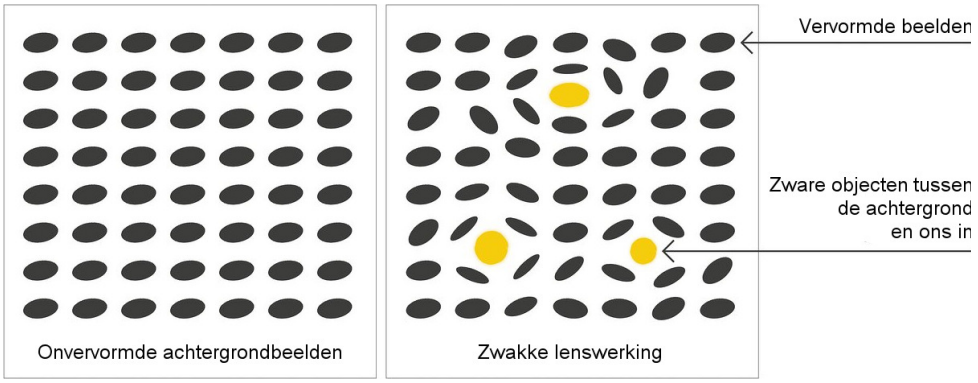
Figuur 6.1: Een voorbeeld van een lens (het grote sterrenstelsel links) met daarnaast het schematische beeld van een gelensd sterrenstelsel in de achtergrond (en daarom kleiner weergegeven). Lichtstralen die dicht langs de lens gaan worden sterker afgebogen dan lichtstralen dan lichtstralen die de lens op grotere afstand passeren. Hierdoor wordt het 'binnenste deel' van het oorspronkelijke (blauwe) beeld sterker weggeduwd dan het 'buitenste' deel. In het waargenomen, "gelensde" (rode) beeld heeft de binnenkant de buitenkant iets ingehaald, waardoor het beeld enigszins is platgedrukt.

groene contouren in Figuur 1.3, blijkt echter dat de grootste massaconcentraties verder uit elkaar liggen, niet bij het hete gas, maar juist waar de twee concentraties van sterrenstelsels zich bevinden. Er blijkt er veel meer materie in het bullet cluster te zijn dan zichtbaar is. Bovendien vertelt gravitationele lenswerking ons ook waar die materie zich bevindt en hoe deze tijdens de botsing heeft bewogen.

6.2.2 Metingen en interpretatie

Dat een model van gravitationele lenswerking geen aannamen hoeven te doen over de dynamische of astrofysische processen in de lenzende objecten, maakt het op dat punt minder gevoelig voor bijvoorbeeld model bias. Dit betekent echter niet dat gravitationele lenswerking niet haar eigen aannamen en onzekerheden kent. Bijvoorbeeld waar het de selectie en oorspronkelijke vorm van achtergrondbronnen betreft, maar ook in de statistische berekening van het effect en de interpretatie van die uitkomsten.

Om de vormen van achtergrondbronnen te meten, zijn zeer nauwkeurige beelden van hoge kwaliteit nodig. Deze beelden kennen altijd een vervorming door bijvoorbeeld de atmosfeer en de optica in de telescoop. Deze vervormingen lijken op vervormingen die veroorzaakt



Figuur 6.2: Een indicatie hoe achtergrondbeelden vervormd kunnen worden door zware objecten tussen die achtergrond en ons, de waarnemers, in. Links is een theoretische achtergrond te zien, zonder vervormingen. Rechts is geschetst hoe de ellipticiteit en de oriëntatie van die achtergrondbeelden verandert door gravitationele lenswerking. Bron: NOIRLab/NSF/AURA.

worden door gravitatielenzen en dus moet hier zorgvuldig voor gecorrigeerd worden. Bijvoorbeeld: als er meerdere beelden worden gecombineerd, kan deze correctie voor de totale combinatie wel nauwkeurig zijn, maar voor de individuele beelden een afwijking vertonen. In hoofdstuk 2 en 3 komen we hier in het bijzonder op terug.

Daarnaast moeten er aannamen worden gedaan over de oorspronkelijke vormen van achtergrondstelsels, terwijl we die nooit hebben kunnen meten. Wat dit betreft is gravitationele lenswerking een schoolvoorbeeld van het indirect waarnemen van een niet zichtbare 'ware vorm'. Om deze informatie te achterhalen, gebruiken we zogenaamde basismodellen die we vervormen en vergelijken met de waarnemingen. De benodigde vervorming is dan het signaal van de gravitationele lenswerking. We zijn dus gevoelig voor een model bias als het de achtergrondvormen betreft. Om dit zorgvuldig te ijken, vergelijken we in hoofdstuk 3 twee verschillende benaderingen hiervoor.

Omdat achtergrondbronnen bovendien al een eigen ellipticiteit hebben, combineren we een groot aantal achtergrondmetingen en nemen we daar het gemiddelde van. Het lenseffect is dan het netto signaal dat overblijft. Hierbij moeten we wel rekening houden met een complexe situatie: de combinatie van intrinsieke ellipticiteit en lenswerking is niet-lineair; de bijkomende meetruis is verschoven ten opzichte van de verdeling van de intrinsieke vormen; en de resulterende verdeling is asymmetrisch. Hierdoor wijkt dit gemiddelde van de gemeten vormen iets af van de waarde van de lenswerking, zoals we in hoofdstuk 4 en 5 aantonen.

Dit is hoe statistiek werkt: men voert een berekening uit en vervolgens koppelt men een betekenis aan de waarde die uit deze berekening volgt. Zelfs als de berekening goed wordt uitgevoerd, hoeft deze uitkomst niet *exact* te beschrijven hoe de situatie in elkaar steekt. Deze statistische bias beïnvloedt vervolgens onze reconstructie van de verdeling van materie aan de hemel. In hoofdstuk 4 en 5 bestuderen we daarom altenatieve statistische berekeningen, die dichter bij de waarheid *lijken* te liggen. In verder onderzoek zal duidelijk moeten worden of dat inderdaad zo is.

6.3 Dit proefschrift

Het onderzoek beschreven in dit proefschrift beslaat vier delen, namelijk een theoretische bespiegeling op de regressiestatistiek in de theorie van zwakke lenswerking en drie afzonderlijke analyses van zwakke lenswerking door groepen en clusters van sterrenstelsels, elk met een focus op bepaalde systematische onzekerheden in de methode.

In **hoofdstuk 2** (Smit et al. 2015) voeren we met behulp van zwakke lenswerking een analyse uit van de zogenaamde "supergroep" SG1120–1202, ook wel een protocluster genoemd (Gonzalez et al. 2005). SG1120–1202 lijkt namelijk te bestaan uit vier relatief lichte groepen van sterrenstelsels op gemiddeld 10 lichtjaren van elkaar. Voorspellingen tonen aan dat deze subgroepen over enkele miljarden lichtjaren samen zullen zijn gevoegd tot één groot cluster, vergelijkbaar met het Coma cluster. Omdat deze supergroep op 3,5 miljard lichtjaar afstand staat, zien we de ontwikkeling zoals deze 4 miljard jaar geleden was en zou deze samensmelting tot één cluster al plaats kunnen hebben gevonden.

De omvang van het gehele stelsel is qua onderlinge afstanden te vergelijken met het eerder genoemde bullet cluster, maar de subgroepen zijn van lagere massa en vertonen nog geen interactie. Ook hier is heet gas dat zichtbaar is in Röntgenstraling, maar dit vertoont geen interactie, zoals schokgolven. SG1120–1202 is onderdeel van een reeks onderzoeken, die onder andere aantonen dat de sterrenstelsels in de subgroepen het grootste deel van hun ontwikkeling al hebben doorgemaakt, hetgeen suggereert dat sterrenstelsels het grootste deel van hun tijd in kleinere groepen doormaken (Tran et al. 2008, 2009). Het vormt dus een unieke kijk op het ontstaan van structuur op kosmische schaal en toont het belang aan van het bestuderen van deze groepsfase.

Met behulp van zeer nauwkeurige waarnemingen met de Hubble Space Telescope (HST) hebben we de ellipticiteiten van achtergrondbronnen bepaald. Met behulp van aanvullende archiefdata van HST waarnemingen en de methoden uit Schrabback et al. (2010) zijn de storende vervormingen door onder andere de telescoop zelf nauwkeuriger bepaald, dan met alleen de gebruikte beelden zelf mogelijk zou zijn geweest. Gezien de onderlinge nabijheid van de subgroepen, is de gravitationele lenswerking van de supergroep in haar geheel gemodelleerd.

Uit het model bleek dat de massaverdeling in SG1120–1202 bijna hetzelfde is als die van de zichtbare sterrenstelsels en de Röntgenstraling van het hete gas in het protocluster, wat bevestigt dat de subgroepen nog niet hebben geïnteracteerd. Eén van de subgroepen ligt ongeveer 100 miljoen lichtjaar dichterbij dan de andere drie en vertoont geen verhoogde massaconcentratie. Van de andere drie hebben we onafhankelijke massabepalingen gedaan die eerdere bepalingen met andere methoden bevestigden.

Hoofdstuk 3 vormt een studie naar 79 zeer nabije en zeer lichte groepen van sterrenstelsels. Deze groepen zijn gevonden in de Two-degree-Field Galaxy Redshift Survey (Colless et al. 2001) en onderdeel van de Zürich Environmental Survey (Carollo et al. 2013), een onderzoek naar de eerder genoemde, belangrijke groepsfase in de ontwikkeling van sterrenstelsels.

Deze groepen, op slechts 600 miljoen lichtjaar afstand, variëren in massa van zeer licht (vergelijkbaar met onze eigen Melkweg) tot de (nog steeds lichte) subgroepen van de supergroep van hoofdstuk 2. Zulke hele lichte structuren zijn tot op de dag van vandaag nog steeds relatief onderbelicht. Daarnaast kon deze studie een oriëntatie vormen voor wat er met de geplande Kilo-Degree Survey (de Jong et al. 2013) mogelijk zou zijn. Dit is uiteindelijk gedaan

in Viola et al. (2015), met iets zwaardere groepen, die iets verder weg liggen.

Zowel de nabijheid als de lage massa van deze groepen vertalen zich naar een zeer zwakke lenswerking. Dit stelt extra hoge eisen aan de beheersing van systematische fouten en de kwaliteit van de meetmethoden. Een bijkomend effect is dat zulke lichte groepen moeilijker te vinden zijn, waardoor het niet altijd zeker is *dat* er een groep is of *waar* het centrum van de groep zich bevindt. Een belangrijk aspect is dus de mogelijke definitie van het centrum van zo'n groep, waarvoor we meerdere opties onderzocht hebben. In dit regime van zeer lichte groepen zijn onze methoden dus gevoeliger voor een mogelijke bias in de detectie of in de modellering van de lenzen.

Ook is er uitgebreid gekeken naar de vervormingen door de telescoop en de eerder genoemde effecten van samenvoeging van meerdere opnamen tot één beeld. Tenslotte hebben we twee methoden vergeleken om de vormen van de achtergrondbronnen te meten. De eerste methode gebruikt het meten van de helderheidsverdeling van de bronnen (KSB+, zoals beschreven in Erben et al. 2001). De andere methode vergelijkt de beelden van achtergrondbronnen met een zeer flexibele set modellen (Shapelets, zoals beschreven in Kuijken 2006).

De resultaten bleken goed met elkaar in overeenstemming. Hierdoor bleek het mogelijk om significante schattingen te maken van verschillende aspecten van de massa's van deze groepen. Deze bleken in goede overeenstemming met onder andere schattingen gebaseerd op de interne snelheden van de groepen. Die schattingen waren dus verkregen door het gebruik van eerder genoemde dynamische modellen. In het regime van (zeer) lage waarden zijn twee onafhankelijke methoden zeer belangrijk.

Hoofdstuk 4 (Smit & Kuijken 2018) is het eerste deel van een tweeluik over de statistische methoden die gebruikt worden voor regressie. Het welbekende gewogen gemiddelde werkt het best bij metingen waarvan de foutenverdelingen zich enigszins regelmatig gedragen, in het bijzonder de normale of Gaussische verdeling. Belangrijk is dat er niet een te groot aantal relatieve uitschieters in de afwijkingen voorkomt en bovendien dat die afwijkingen redelijk symmetrisch zijn verdeeld.

De natuurlijke verdeling van ellipticiteiten van sterrenstelsels vertoont een scherpe centrale piek, waar de meeste informatie te vinden is. Een methode die minder afhankelijk is van grote afwijkingen en meer van die centrale piek ligt dan voor de hand.

Daarnaast is het effect van lenswerking op die natuurlijke verdeling niet lineair en is de verdeling van meetfouten afwijkend. Dat zorgt voor een asymmetrie in de gemeten verdeling van ellipticiteiten die tot een afwijking – een statistische bias – van enkele procenten kan leiden in een gewogen gemiddelde, ten opzichte van de ware positie van de piek.

We bekijken enkele andere regressiemethoden, waarvan de belangrijkste twee de methoden van Kleinste Absolute Afwijking (Least Absolute Deviations of LAD) en Convexe Omhulling (Convex Hull Peeling of CHP) zijn. Beiden leiden in het eendimensionale geval tot de mediaan en zijn veel minder gevoelig voor sterk afwijkende waarden.

Deze regressiemethoden hebben we getest op verschillende simulaties van ellipticiteiten, die we optimaal konden aanpassen, en op de verdeling van gemeten ellipticiteiten in de catalogus van de Canada-France-Hawaii Lensing Survey (Heymans et al. 2012b). Daarbij hebben we gekeken naar de bias van de statistische methode, maar ook naar de betrouwbaarheid.

Uit ons onderzoek kwam naar voren dat de bias van CHP in theorie de kleinste kan zijn, maar dat LAD de meest betrouwbare verbetering oplevert ten opzichte van het gewogen gemiddelde, door een makkelijkere toepasbaarheid en stabielere uitkomsten.

In **hoofdstuk 5** (Smit et al. 2021) passen we de resultaten van hoofdstuk 4 toe. Hier-

voor gebruiken we nauwkeurige metingen van meer dan 10 miljoen achtergrondbronnen in 450 vierkante graden van KiDS. Als beoogde lenzen gebruiken we bijna 7000 tussenliggende clusters van sterrenstelsels, gevonden met de Adaptive Matched Identifier of Clustered Objects (AMICO, Bellagamba et al. 2018) in de KiDS waarnemingen.

Door de hoge kwaliteit en aantallen konden we het signaal nauwkeurig genoeg vaststellen, om het verschil van enkele procenten tussen het gewogen gemiddelde en LAD duidelijk zichtbaar te maken. In het bepalen van de verdeling van de dichtheid in de clusters kwam dit verschil het meest duidelijk naar voren. Dit hebben we vertaald naar het effect op de relatie tussen lichtkracht en massa van clusters van sterrenstelsels.

Ook in deze toepassing bleek LAD een iets hoger signaal te geven, met bovendien kleinere onzekerheden in de gevonden waarden.

Bibliography

Abazajian, K., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2003, AJ, 126, 2081

Appenzeller, I., Fricke, K., Fürtig, W., et al. 1998, The Messenger, 94, 1

Bacon, D. J., Refregier, A. R., & Ellis, R. S. 2000, MNRAS, 318, 625

Barnard, E. E. 1906, Popular Astronomy, 14, 579

Barrodale, I. & Roberts, F. D. K. 1973, SIAM Journal on Numerical Analysis, 10, 839

Bartelmann, M. 1996, A&A, 313, 697

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

Bartelmann, M. & Schneider, P. 2001, Phys. Rep., 340, 291

Beaton, A. E. & Tukey, J. W. 1974, Outliers in Statistical Data (New York: Wiley)

Begeman, K., Belikov, A. N., Boxhoorn, D. R., & Valentijn, E. A. 2013, Experimental Astronomy, 35, 1

Bellagamba, F., Maturi, M., Hamana, T., et al. 2011, MNRAS, 413, 1145

Bellagamba, F., Roncarelli, M., Maturi, M., & Moscardini, L. 2018, MNRAS, 473, 5221

Bellagamba, F., Sereno, M., Roncarelli, M., et al. 2019, MNRAS, 484, 1598

Benítez, N. 2000, ApJ, 536, 571

Benjamin, J., Van Waerbeke, L., Heymans, C., et al. 2013, MNRAS, 431, 1547

Berlind, A. A., Frieman, J., Weinberg, D. H., et al. 2006, ApJS, 167, 1

Bernstein, G. M. 2010, MNRAS, 406, 2793

Bernstein, G. M. & Armstrong, R. 2014, MNRAS, 438, 1880

Bernstein, G. M. & Jarvis, M. 2002, AJ, 123, 583

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

Bonnet, H. & Mellier, Y. 1995, A&A, 303, 331

Bosma, A. 1981, AJ, 86, 1825

Brainerd, T. G., Blandford, R. D., & Smail, I. 1996, ApJ, 466, 623

Bridle, S., Balan, S. T., Bethge, M., et al. 2010, MNRAS, 405, 2044

Brouwer, M. M., Cacciato, M., Dvornik, A., et al. 2016, MNRAS, 462, 4451

Burnham, R. J. 1978, Burnham's Celestial Handbook: An Observer's Guide to the Universe Beyond the Solar System, Vol. 1 (New York: Dover Publications)

Cacciato, M., van den Bosch, F. C., More, S., Mo, H., & Yang, X. 2013, MNRAS, 430, 767

Capaccioli, M. & Schipani, P. 2011, The Messenger, 146, 2

Carollo, C. M., Cibinel, A., Lilly, S. J., et al. 2013, ApJ, 776, 71

Chwolson, O. 1924, Astronomische Nachrichten, 221, 329

Clampitt, J. & Jain, B. 2016, MNRAS, 457, 4135

Clowe, D., Bradač, M., Gonzalez, A. H., et al. 2006, ApJ, 648, L109

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

Coe, D., Benítez, N., Sánchez, S. F., et al. 2006, AJ, 132, 926

Colless, M., Dalton, G., Maddox, S., et al. 2001, MNRAS, 328, 1039

Cooray, A. & Sheth, R. 2002, Phys. Rep., 372, 1

Cramér, H. 1946, Mathematical Methods of Statistics, Princeton mathematical series (Princeton University Press)

Crittenden, R. G., Natarajan, P., Pen, U.-L., & Theuns, T. 2002, ApJ, 568, 20

Curtis, H. D. 1917, PASP, 29, 206

Dark Energy Survey Collaboration, Abbott, T., Abdalla, F. B., et al. 2016, MNRAS, 460, 1270

de Jong, J. T. A., Verdoes Kleijn, G. A., Boxhoorn, D. R., et al. 2015, A&A, 582, A62

de Jong, J. T. A., Verdoes Kleijn, G. A., Erben, T., et al. 2017, A&A, 604, A134

de Jong, J. T. A., Verdoes Kleijn, G. A., Kuijken, K. H., & Valentijn, E. A. 2013, Experimental Astronomy, 35, 25

Dresser, R. 1992, The Hastings Center Report, 22, 24

Driver, S. P., Hill, D. T., Kelvin, L. S., et al. 2011, MNRAS, 413, 971

Duffy, A. R., Schaye, J., Kay, S. T., & Dalla Vecchia, C. 2008, MNRAS, 390, L64

Dvornik, A., Cacciato, M., Kuijken, K., et al. 2017, MNRAS, 468, 3251

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

Eddington, A. S. 1920, Space, time and gravitation. an outline of the general relativity theory

Einstein, A. 1917, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), 142

Eke, V. R., Baugh, C. M., Cole, S., et al. 2004, MNRAS, 348, 866

Eke, V. R., Frenk, C. S., Baugh, C. M., et al. 2004, MNRAS, 355, 769

Erben, T., Hildebrandt, H., Lerchster, M., et al. 2009, A&A, 493, 1197

Erben, T., Hildebrandt, H., Miller, L., et al. 2013, MNRAS, 433, 2545

Erben, T., Schirmer, M., Dietrich, J. P., et al. 2005, Astronomische Nachrichten, 326, 432

Erben, T., Van Waerbeke, L., Bertin, E., Mellier, Y., & Schneider, P. 2001, A&A, 366, 717

Evans, A. K. D. & Bridle, S. 2009, ApJ, 695, 1446

Falk, M. 1997, Annals of the Institute of Statistical Mathematics, 49, 615

Feigelson, E. D. 1988, Bulletin d'Information du Centre de Donnees Stellaires, 35, 197

Feigelson, E. D. 2009, arXiv e-prints, arXiv:0903.0416

Feigelson, E. D. & Babu, G. J. 2013, Statistical Methods for Astronomy, ed. T. D. Oswalt & H. E. Bond (Springer, Dordrecht), 445

Fenech Conti, I., Herbonnet, R., Hoekstra, H., et al. 2017, MNRAS, 467, 1627

Finoguenov, A., Guzzo, L., Hasinger, G., et al. 2007, ApJS, 172, 182

Ford, J., Hildebrandt, H., Van Waerbeke, L., et al. 2012, ApJ, 754, 143

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

Freeland, E., Tran, K.-V. H., Irwin, T., et al. 2011, ApJ, 742, L34

Freeman, K. C. 1970, ApJ, 160, 811

Gal, R. R., Lemaux, B. C., Lubin, L. M., Kocevski, D., & Squires, G. K. 2008, ApJ, 684, 933

George, M. R., Leauthaud, A., Bundy, K., et al. 2012, ApJ, 757, 2

Gibson, S. 2017, The Pleiades, retrieved from http://www.naic.edu/ gibson/pleiades/

Gilbank, D. G., Gladders, M. D., Yee, H. K. C., & Hsieh, B. C. 2011, AJ, 141, 94

Giocoli, C., Marulli, F., Moscardini, L., et al. 2021, A&A, 653, A19

Gladders, M. D. & Yee, H. K. C. 2005, ApJS, 157, 1

Gonzalez, A. H., Tran, K.-V. H., Conbere, M. N., & Zaritsky, D. 2005, ApJ, 624, L73

Gonzalez, A. H., Zaritsky, D., Dalcanton, J. J., & Nelson, A. 2001, ApJS, 137, 117

Hard, R. 2004, The Routledge Handbook of Greek Mythology (London: Routledge)

Heck, A., Murtagh, F., & Ponz, D. 1985, The Messenger, 41, 22

Herbonnet, R., Buddendiek, A., & Kuijken, K. 2017, A&A, 599, A73

Heymans, C., Bell, E., Rix, H.-W., et al. 2006, MNRAS, 372, 758

Heymans, C., Rowe, B., Hoekstra, H., et al. 2012a, MNRAS, 421, 381

Heymans, C., Van Waerbeke, L., Bacon, D., et al. 2006, MNRAS, 368, 1323

Heymans, C., Van Waerbeke, L., Miller, L., et al. 2012b, MNRAS, 427, 146

Hildebrandt, H., Choi, A., Heymans, C., et al. 2016, MNRAS, 463, 635

Hildebrandt, H., Erben, T., Kuijken, K., et al. 2012, MNRAS, 421, 2355

Hildebrandt, H., Muzzin, A., Erben, T., et al. 2011, ApJ, 733, L30

Hildebrandt, H., Van Waerbeke, L., & Erben, T. 2009, A&A, 507, 683

Hildebrandt, H., Viola, M., Heymans, C., et al. 2017, MNRAS, 465, 1454

Hinshaw, G., Nolta, M. R., Bennett, C. L., et al. 2007, ApJS, 170, 288

Hirata, C. & Seljak, U. 2003, MNRAS, 343, 459

Hirata, C. M., Mandelbaum, R., Seljak, U., et al. 2004, MNRAS, 353, 529

Hirata, C. M. & Seljak, U. 2004, Phys. Rev. D, 70, 063526

Hoekstra, H. 2003, MNRAS, 339, 1155

Hoekstra, H. 2004, MNRAS, 347, 1337

Hoekstra, H. 2005, in Maps of the Cosmos, ed. M. Colless, L. Staveley-Smith, & R. A. Stathakis, Vol. 216, 140

Hoekstra, H., Franx, M., Kuijken, K., et al. 2001, ApJ, 548, L5

Hoekstra, H., Franx, M., Kuijken, K., & Squires, G. 1998, ApJ, 504, 636

Hoekstra, H., Hartlap, J., Hilbert, S., & Van Uitert, E. 2011, MNRAS, 412, 2095

Hoekstra, H. & Jain, B. 2008, Annual Review of Nuclear and Particle Science, 58, 99

Hoekstra, H., Mahdavi, A., Babul, A., & Bildfell, C. 2012, MNRAS, 427, 1298

Hoekstra, H., Yee, H. K. C., & Gladders, M. D. 2004, ApJ, 606, 67

Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, ApJ, 873, 111

Jarvis, M. & Jain, B. 2004, arXiv e-prints, arXiv:astro.ph/0412234

Jarvis, M., Schechter, P., & Jain, B. 2008, arXiv e-prints, arXiv:0810.0027

Jarvis, M., Sheldon, E., Zuntz, J., et al. 2016, MNRAS, 460, 2245

Johnston, D. E., Sheldon, E. S., Wechsler, R. H., et al. 2007, arXiv e-prints, arXiv:0709.1159

Just, D. W., Zaritsky, D., Tran, K.-V. H., et al. 2011, ApJ, 740, 54

Kacprzak, T., Bridle, S., Rowe, B., et al. 2014, MNRAS, 441, 2528

Kacprzak, T., Zuntz, J., Rowe, B., et al. 2012, MNRAS, 427, 2711

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

Kaiser, N., Squires, G., & Broadhurst, T. 1995, ApJ, 449, 460

Kaiser, N., Wilson, G., & Luppino, G. A. 2000, arXiv e-prints, astro

Kalas, P., Graham, J. R., Chiang, E., et al. 2008, Science, 322, 1345

Kapteyn, J. C. 1922, ApJ, 55, 302

Kautsch, S. J., Gonzalez, A. H., Soto, C. A., et al. 2008, ApJ, 688, L5

Kelvin, L. 1904, Baltimore Lectures on Molecular Dynamics and the Wave Theory of Light (London, England: C.J. Clay and Sons)

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

Kitching, T. D., Balan, S. T., Bridle, S., et al. 2012, MNRAS, 423, 3163

Kitching, T. D., Miller, L., Heymans, C. E., Van Waerbeke, L., & Heavens, A. F. 2008, MNRAS, 390, 149

Koekemoer, A. M., Fruchter, A. S., Hook, R. N., & Hack, W. 2003, in HST Calibration Workshop : Hubble after the Installation of the ACS and the NICMOS Cooling System, ed. S. Arribas, A. Koekemoer, & B. Whitmore, 337

Krupp, E. 1991, Griffith Observer, 55, 1

Kuijken, K. 1999, A&A, 352, 355

Kuijken, K. 2006, A&A, 456, 827

Kuijken, K. 2011, The Messenger, 146, 8

Kuijken, K. & Gilmore, G. 1989, MNRAS, 239, 651

Kuijken, K., Heymans, C., Hildebrandt, H., et al. 2015, MNRAS, 454, 3500

Kyselka, W. 1993, Hawaiian Journal of History, 27, 174

Lagattuta, D. J. 2011, PhD thesis, University of California, Davis

Lambas, D. G., Maddox, S. J., & Loveday, J. 1992, MNRAS, 258, 404

Lauer, T. R., Postman, M., Weaver, H. A., et al. 2021, ApJ, 906, 77

Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, arXiv e-prints, arXiv:1110.3193

Le Fèvre, O., Guzzo, L., Meneux, B., et al. 2005, A&A, 439, 877

Le Fèvre, O., Saisse, M., Mancini, D., et al. 2003, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 4841, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, ed. M. Iye & A. F. M. Moorwood, 1670–1681

Leauthaud, A., Finoguenov, A., Kneib, J.-P., et al. 2010, ApJ, 709, 97

Leauthaud, A., Massey, R., Kneib, J.-P., et al. 2007, ApJS, 172, 219

Lee, D. T. & Schachter, B. J. 1980, International Journal of Computer & Information Sciences, 9, 219

Lee, H. 2007, in Astronomical Society of the Pacific Conference Series, Vol. 371, Statistical Challenges in Modern Astronomy IV, ed. G. J. Babu & E. D. Feigelson, 425

Lemaux, B. C., Gal, R. R., Lubin, L. M., et al. 2012, ApJ, 745, 106

Lesci, G. F., Marulli, F., Moscardini, L., et al. 2020, arXiv e-prints, arXiv:2012.12273

Malmquist, K. G. 1925, Meddelanden fran Lunds Astronomiska Observatorium Serie I, 106, 1

Mandelbaum, R. 2018, ARA&A, 56, 393

Mandelbaum, R., Hirata, C. M., Ishak, M., Seljak, U., & Brinkmann, J. 2006, MNRAS, 367, 611

Mandelbaum, R., Miyatake, H., Hamana, T., et al. 2018, PASJ, 70, S25

Mandelbaum, R., Rowe, B., Armstrong, R., et al. 2015, MNRAS, 450, 2963

Mandelbaum, R., Seljak, U., Cool, R. J., et al. 2006, MNRAS, 372, 758

Mandelbaum, R., Seljak, U., & Hirata, C. M. 2008, JCAP, 8, 6

Mandelbaum, R., Seljak, U., Kauffmann, G., Hirata, C. M., & Brinkmann, J. 2006, MNRAS, 368, 715

Margoniner, V. E., Lubin, L. M., Wittman, D. M., & Squires, G. K. 2005, AJ, 129, 20

Markevitch, M., Gonzalez, A. H., David, L., et al. 2002, ApJ, 567, L27

Massey, R., Heymans, C., Bergé, J., et al. 2007, MNRAS, 376, 13

Massey, R., Stoughton, C., Leauthaud, A., et al. 2010, MNRAS, 401, 371

Mather, J. C. 1982, Optical Engineering, 21, 769

Maturi, M., Bellagamba, F., Radovich, M., et al. 2019, MNRAS, 485, 498

McFarland, J. P., Verdoes-Kleijn, G., Sikkema, G., et al. 2013, Experimental Astronomy, 35, 45

Mead, A. J., Peacock, J. A., Heymans, C., Joudaki, S., & Heavens, A. F. 2015, MNRAS, 454, 1958

Melchior, P. & Viola, M. 2012, MNRAS, 424, 2757

Miller, L., Heymans, C., Kitching, T. D., et al. 2013, MNRAS, 429, 2858

Miller, L., Kitching, T. D., Heymans, C., Heavens, A. F., & Van Waerbeke, L. 2007, MNRAS, 382, 315

Miyatake, H., More, S., Mandelbaum, R., et al. 2015, ApJ, 806, 1

Monroe, J. T., Tran, K.-V. H., & Gonzalez, A. H. 2017, ApJ, 836, 7

Mosteller, F. & Tukey, J. W. 1977, Data Analysis and Regression: a Second Course in Statistics (Reading, MA: Addison Wesley), p. 133

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

Neto, A. F., Gao, L., Bett, P., et al. 2007, MNRAS, 381, 1450

Norris, R. P. & Norris, B. R. M. 2021, in Advancing Cultural Astronomy, ed. E. Boutsikas, S. McCluskey, & J. Steele, Historical & Cultural Astronomy (Springer, Cham)

Oguri, M., Bayliss, M. B., Dahle, H., et al. 2012, MNRAS, 420, 3213

Oguri, M., Miyazaki, S., Hikage, C., et al. 2018, PASJ, 70, S26

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

Oort, J. H. 1932, Bull. Astron. Inst. Netherlands, 6, 249

Parker, L., Hudson, M. J., Carlberg, R. G., & Hoekstra, H. 2005, ApJ, 634, 806

Parker, L. C., Hoekstra, H., Hudson, M. J., Van Waerbeke, L., & Mellier, Y. 2007, ApJ, 669, 21

Peacock, J. A. & Smith, R. E. 2000, MNRAS, 318, 1144

Peebles, P. J. E. 1970, AJ, 75, 13

Penzias, A. A. & Wilson, R. W. 1965, ApJ, 142, 419

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2014, A&A, 571, A16

Preparata, F. P. & Shamos, M. I. 1985, Computational Geometry (New York: Springer), 95–149

Radovich, M., Puddu, E., Bellagamba, F., et al. 2017, A&A, 598, A107

Rao, C. R. 1945, Bull. Calcutta Math. Soc., 37, 81

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

Rhodes, J. D., Massey, R. J., Albert, J., et al. 2007, ApJS, 172, 203

Robotham, A. S. G., Norberg, P., Driver, S. P., et al. 2011, MNRAS, 416, 2640

Rodríguez, S. & Padilla, N. D. 2013, MNRAS, 434, 2153

Rubin, V. C. 1983, Scientific American, 248, 96

Rykoff, E. S., Rozo, E., Hollowood, D., et al. 2016, ApJS, 224, 1

Schirmer, M. 2013, ApJS, 209, 21

Schneider, M. D., Hogg, D. W., Marshall, P. J., et al. 2015, ApJ, 807, 87

Schneider, P. 2003, arXiv e-prints, astro

Schneider, P. 2006, in Saas-Fee Advanced Course 33: Gravitational Lensing: Strong, Weak and Micro, ed. G. Meylan, P. Jetzer, P. North, P. Schneider, C. S. Kochanek, & J. Wambsganss, 269–451

Schneider, P. & Seitz, C. 1995, A&A, 294, 411

Schrabback, T., Erben, T., Simon, P., et al. 2007, A&A, 468, 823

Schrabback, T., Hartlap, J., Joachimi, B., et al. 2010, A&A, 516, A63

Schwarzschild, M. 1954, AJ, 59, 273

Seitz, C. & Schneider, P. 1995, A&A, 297, 287

Seitz, C. & Schneider, P. 1997, A&A, 318, 687

Seljak, U. 2000, MNRAS, 318, 203

Sellentin, E., Heymans, C., & Harnois-Déraps, J. 2018, MNRAS, 477, 4879

Sérsic, J. L. 1963, Boletin de la Asociacion Argentina de Astronomia La Plata Argentina, 6, 41

Sérsic, J. L. 1968, Atlas de Galaxias Australes

Sifón, C., Cacciato, M., Hoekstra, H., et al. 2015a, MNRAS, 454, 3938

Sifón, C., Hoekstra, H., Cacciato, M., et al. 2015b, A&A, 575, A48

Slipher, V. M. 1915, Popular Astronomy, 23, 21

Smit, M., Dvornik, A., Radovich, M., et al. 2021, arXiv e-prints, arXiv:2109.12009, *accepted*

Smit, M. & Kuijken, K. 2018, A&A, 609, A103

Smit, M., Schrabback, T., Velander, M., et al. 2015, A&A, 582, A82

Spergel, D. N., Verde, L., Peiris, H. V., et al. 2003, ApJS, 148, 175

Stark, A. A. 1977, ApJ, 213, 368

Sunyaev, R. A. & Zeldovich, Y. B. 1970, Ap&SS, 7, 3

Sunyaev, R. A. & Zeldovich, Y. B. 1972, Comments on Astrophysics and Space Physics, 4, 173

Tempel, E., Tago, E., & Liivamägi, L. J. 2012, A&A, 540, A106

Tinker, J. L., Robertson, B. E., Kravtsov, A. V., et al. 2010, ApJ, 724, 878

Tran, K.-V. H., Moustakas, J., Gonzalez, A. H., et al. 2008, ApJ, 683, L17

Tran, K.-V. H., Saintonge, A., Moustakas, J., et al. 2009, ApJ, 705, 809

Tucker, W. H., Tananbaum, H., & Remillard, R. A. 1995, ApJ, 444, 532

Tyson, J. A., Valdes, F., Jarvis, J. F., & Mills, A. P., J. 1984, ApJ, 281, L59

Tyson, J. A., Valdes, F., & Wenk, R. A. 1990, ApJ, 349, L1

Valentijn, E. A., McFarland, J. P., Snigula, J., et al. 2007, in Astronomical Society of the Pacific Conference Series, Vol. 376, Astronomical Data Analysis Software and Systems XVI, ed. R. A. Shaw, F. Hill, & D. J. Bell, 491

van den Bosch, F. C., More, S., Cacciato, M., Mo, H., & Yang, X. 2013, MNRAS, 430, 725

Van Leeuwen, F. 2009, A&A, 497, 209

van Uitert, E., Cacciato, M., Hoekstra, H., et al. 2016, MNRAS, 459, 3251

van Uitert, E., Hoekstra, H., Joachimi, B., et al. 2017, MNRAS, 467, 4131

Van Uitert, E. & Schneider, P. 2016, A&A, 595, A93

Van Waerbeke, L. 2000, MNRAS, 313, 524

Van Waerbeke, L., Hildebrandt, H., Ford, J., & Milkeraitis, M. 2010, ApJ, 723, L13

Van Waerbeke, L., Mellier, Y., Erben, T., et al. 2000, A&A, 358, 30

Velander, M., Kuijken, K., & Schrabback, T. 2011, MNRAS, 412, 2665

Verdoes Kleijn, G., de Jong, J. T. A., Valentijn, E., et al. 2012, in Astronomical Society of the Pacific Conference Series, Vol. 461, Astronomical Data Analysis Software and Systems XXI, ed. P. Ballester, D. Egret, & N. P. F. Lorente, 237

Viola, M., Cacciato, M., Brouwer, M., et al. 2015, MNRAS, 452, 3529

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

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

Von der Linden, A., Allen, M. T., Applegate, D. E., et al. 2014, MNRAS, 439, 2

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

Wright, C. O. & Brainerd, T. G. 2000, ApJ, 534, 34

Xue, Y.-J. & Wu, X.-P. 2000, ApJ, 538, 65

Zhang, J., Luo, W., & Foucaud, S. 2015, J. Cosmology Astropart. Phys., 1, 024

Zhang, J., Zhang, P., & Luo, W. 2017, ApJ, 834, 8

Zwicky, F. 1933, Helvetica Physica Acta, 6, 110

Zwicky, F. 1937, Physical Review, 51, 290

Publications

"CFHTLenS: the CanadaFranceHawaii Telescope Lensing Survey – imaging data and catalogue products"
Authors: Thomas Erben, Hendrik Hildebrandt, Lance Miller, Ludovic van Waerbeke, Catherine Heymans, Henk Hoekstra, Thomas D. Kitching, Yannick Mellier, Jonathan Benjamin, Chris Blake, Christopher Bonnett, Oliver-Mark Cordes, Jean Coupon, Liping Fu, Raphaël Gavazzi, Bryan Gillis, Emma Grocutt, Stephen D. J. Gwyn, Karianne Holhjem, Michael J. Hudson, Martin Kilbinger, Konrad Kuijken, Martha Milkeraitis, Barnaby T. P. Rowe, Tim Schrabback, Elisabetta Semboloni, Patrick Simon, Merijn Smit, Ovidiu Toader, Sanaz Vafaei, Edo van Uitert, Malin Velander
Published: 2013, MNRAS, 433, 2545
DOI: 10.1093/mnras/stt928

"Mass distribution in an assembling super galaxy group at $z = 0.37$"
Authors: Merijn Smit, Tim Schrabback, Malin Velander, Konrad Kuijken, Anthony H. Gonzalez, John Moustakas, and Kim-Vy H. Tran
Published: 2015, A&A, 582, A82
DOI: 10.1051/0004-6361/201525905

"Chasing the peak: optimal statistics for weak shear analyses"
Authors: Merijn Smit and Konrad Kuijken
Published: 2018, A&A, 609, A103
DOI: 10.1051/0004-6361/201731410

"AMICO galaxy clusters in KiDS-DR3: the impact of estimator statistics on the luminosity-mass scaling relation"
Authors: Merijn Smit, Andrej Dvornik, Mario Radovich, Konrad Kuijken, Matteo Maturi, Lauro Moscardini, and Mauro Sereno
Accepted: 2021, A&A

Curriculum Vitae

"A hero is someone who simply got too frightened to use his good sense and run away, then somehow lived through it all."

– Raymond E. Feist, *Silverthorn*

This is the part where I get to talk about myself, the good, the bad and the beautiful. I realize hardly anyone will read this part, and I have taken the liberty in using the writing of it – with a cup of coffee and a wee dram of whisky – as a summarizing reflection on my life, as far as it pertains to my professional career of course, but – inescapably intertwined – also to the scientist and the person I have become.

I was born in Haarlem, The Netherlands, on the 7th of September in 1977. I suspect that – in a way, even at that age – I was already easily distracted and momentarily lost sight of the tasks right in front of me, as I stopped breathing right away. However, be it curiosity or stubbornness, I eventually got back to it and to this day, I still foster a passion for it, akin to that for astronomy.

For as long as I can remember, I enjoy understanding. And then imagining the alternative possibilities. As soon as I could read, I was hooked and essentially didn't put the books down, until well within adulthood. During my years in primary school, I was more than a handful. As soon as I understood something, my mind had already wandered down several other paths. It's not that school didn't interest me. It did. I wanted to understand everything in school. Yet I also wanted to understand, experience, explore and imagine a lot of other things as well and I simply didn't have time to dawdle, doing the same old exercises.

When hiking through the heathlands in Drenthe, on the Veluwe or the mountains in France, I was fascinated by maps and the explanations of my father, a teacher of geography and mathematics, about how and why the world was ordered as it was. I remember picking my mother's and father's brain about the differences between religions and, during the same summer holiday evenings, I kept asking for another challenge in determining the square root of a non integer number. I was swept away by the 3D shows in the then newly opened Omniversum, when my grandmother took me there. To this day I remember the shows about astronomy, the physics behind airplanes and dizzying and exhilarating journeys over strange landscapes or through the depths of the oceans. She even gave me a booklet called 'Welke ster is dat?' by W. Widmann, full of charts and constellations for each time of the year. I already had my own simple camera, and being allowed to use my father's Canon SLR on rare occasions was a treat. Unhindered by any talent for the sport, I still managed to make it to the first soccer team of the school through stubbornness and running around a lot, and boy was I proud when we won the local championship and entered the regional school championship for the first time in years.

My high school years I attended College Hageveld in Heemstede. It is the only categorical athenaeum in The Netherlands, though I still had the opportunity to submerse myself in the classics, culminating in visiting Rome on a school trip and graduating in Latin besides Dutch, English, German, mathematics, physics, chemistry and biology. I wanted to take French,

geography and history as well, but the system wasn't flexible enough at the time. Hageveld was also a place that fostered creativity and curiosity, and I explored my first steps on stage, in writing and in music.

After extensive orienting visits to various universities all over the country, seriously contemplating philosophy, veterinary medicine, electrical or aerospace engineering, the choice for astronomy felt in hindsight as not just logical, but inevitable. I finally enjoyed a true challenge, as the sky was not the limit, but merely the start. The bigger the challenge, the stronger my interest, the higher my grades. The master courses in general relativity and gravitation, differential equations, cosmology and philosophy of science were among the highlights, some of which were the last to be taught in the classic Kamerlingh Onnes building. I also had my first taste of teaching, as I related to the struggles of students in the years below me and shared my experiences and insights gleaned during my own similar struggles. Not just sharing the knowledge, but how to change a way of thinking, how to learn. I did a minor thesis with Dr. Peter Katgert on cosmology, focusing on quintessence versus a cosmological constant. My master thesis about galactic bars in edge-on galaxies with boxy bulges, a case study of IC2531, was done under supervision of Prof. dr. Tim de Zeeuw and Dr. Martin Bureau.

Of course, besides also still devouring books, my interests still wandered all over the place, and I immersed myself in tabletop games and the first variants of online games, not only playing, but also developing and writing them. With a group of close friends, we visited the cinema at least once a week and of course several editions of the Leiden Movie Festival. I assisted the astronomy association Kaiser with a few activities over the years and participated in the trip to Northern France to see the solar eclipse of 1999 in Noyon. I joined SKC, the student volleyball club, and over the course of ten years, worked my way from the lowest to the first team, became a licensed referee and joined several committees, such as the social activities committee, the lustrum committee, the bar committee and the monthly club magazine.

Most importantly, I actively participated in De Leidsche Flesch, the student association for physics, astronomy, mathematics and computer science. Besides joining committees for first year students, the yearly almanac, movie nights and games, among others, I spent a year as president. At the time, member engagement was very low and some people wondered if the association still had a purpose, a future. With a fantastic board and a group of creative and enthusiastic first year students, we watched the association bloom with renewed social and scientific activities, planting during this year many seeds that would come to fruition in following years, like the yearly student trip abroad. It was an impressive and privileged experience to be at the front row during that change.

It will come as no surprise that at some point, my studies were taking quite a long time. A student has to eat, and I decided that after my unofficial experiences with teaching, I could dabble in the real thing. It turned out that I could all too well, and before I realized, I had three years experience in teaching almost full time in high school, at a sublocation of the Carmelcollege in Gouda. With the support at home and of Prof. dr. de Zeeuw, I quit a steady income to return home to astronomy – at least that's how it felt.

Fueled by the feeling that I had returned where I belonged, I finished my masters degree. Over the summer, I did a short research project with Dr. Richard McDermid on M32 using Oasis data in my spare time. At the same time, I applied for a PhD position with Prof. Dr. Koen Kuijken to work on weak gravitational lensing with the Kilo-Degree Survey and was ecstatic when I was accepted. In expectation of the start of KiDS, I started work-

ing on Wide-Field Imager data of nearby galaxy groups found in the 2-degree Field Galaxy Redshift Survey and observations on galaxy supergroup SG1120–1202 from the Advanced Camera for Surveys on the Hubble Space Telescope. As part of the growing lensing group in Leiden, I had the privilege to take part in the 'Dark side of the Universe through Extragalactic Lensing' (DUEL) European training network and the collaborations on Canada-France-Hawaii Legacy Survey data, evolving in name and composition from CFHT-CARS, through CFHTLS Systematics to the final CFHT Lensing Survey (CFHTLenS). I attended summer schools and workshops on statistics at Penn State, USA; astronomical image processing in Dubrovnik, Croatia; AstroWISE in Leiden, The Netherlands. I participated in scientific meetings, presented my work and gave talks and colloquia in Leiden, Ameland, and Veldhoven, The Netherlands; State College, USA; Santiago, Chile; London, and Edinburgh, UK; Paris, France; Sydney, Australia; Victoria, Canada.

At the same time, I assisted several years with the courses Introductory Astrophysics and Special Relativity. I was active for several consecutive years in the institute social committee (as chair), the computer committee and the institute council. During the same years, I worked on public outreach and gave popular lectures in, among others, Amsterdam, Leiden, Alkmaar, Hoorn, and Wormer.

During the full second year and part of the third year of my PhD, I went through a divorce that ended a six year relationship. Pursuing many interests and passions at the same time had been a characteristic of my life and that provided escape in many different directions at once. My journey made the softest landing possible, when I soon after met my wife and love of my life, but this newfound happiness occupied my attention more than the tumbleweeds blowing through the silent landscape of my PhD.

Finding strength and encouragement in each other, I returned to astronomy once more, but without funding. With savings, I could bridge a gap of several months, but to support myself, I needed to return to my secondary passion: teaching. Continuing my PhD one day per week, I completed my masters degrees in eduction of both mathematics and physics, cum laude. Since then, I have been teaching in high school for ten years, developing an interest and indeed an opinion on teacher education. I am a certified and registered video coach and have participated with great interest in working collaborations on the subjects of education of mathematics and statistics within the Dutch Association of Teachers in Mathematics (NVvW).

This has been a long reflection, but then, I am now finishing my PhD at the age of 44. I still feel that I am where I belong, and am currently living in the most important time of my life, whatever dreams and passions may lie ahead.

"Wisdom comes from experience. Experience is often a result of lack of wisdom."

– Terry Pratchett

Acknowledgements

"Every person you encounter, whom you interact with, is there to teach you something. Sometimes it may be years before you realize what each had to show you."

– Raymond E. Feist, *Rise of a Merchant Prince*

When you stick around as long as I have, perhaps the most difficult task is to not forget a single individual who helped, inspired and supported me during my time in Leiden, both as scientists and as friends. If I did, I hope you'll still accept my apologies and gratitude.

I've heard it is not customary to acknowledge directly your promotor. Yet Koen, you have been much more than a regular thesis supervisor. When I applied for the position, I followed the advice I had been given: a good supervisor is more important than an interesting project. How true that advice turned out to be, even though weak lensing has captured my heart. Your patience and availability during the full 16 years that my PhD lasted have been incredible. You never 'closed the door', as the saying goes, an example that I have taken to heart and which has in turn made me a better teacher and supervisor to my own students. You have given me an incredible opportunity and during most of our meetings you'd reiterate "go for it!" I may have needed a long time, but I have taken that opportunity. If you remember, I promised to walk back uphill, but sometimes one needs a path to be available.

I have been very fortunate by the support, guidance and friendship of the members of the Leiden lensing group. Tim, Malin, Edo, Hendrik, I fondly remember the times when we, as a small but growing group, would go to meetings and conferences together, and highlights include the hiking trip over the Berg Lake Trail in the Canadian Rocky Mountains, the tour of the New Zealand South Island including all those Lord of the Rings filming locations, the sailing trip from Dubrovnik and the drinks on the steep cliffs just outside the walls. Margot and Andrej, my time of visiting conferences and meetings may not have intersected with yours, but I treasure our talks and exchange of ideas and above all your company, and I consider you amazing persons, friends and scientists.

I have had the privilege of being a part of the DUEL and CFHTLenS groups and have learned more about lensing and what 'doing science' really is from those meetings, than any amount of courses or summer schools could offer. Besides some names from the Leiden group, I want to mention Sanaz, Emma, Karianne, Henk, Elisabetta, Barney, Tom, Martin, Liping, Jean. I want to thank especially Ludo and Catherine: I have considered the both of you inspiring scientists, mentors to many in the group, friendly and approachable in person and yet expert scientific leaders by example.

There have been many other members of the lensing group I want to thank for advice and fruitful discussions during the group meetings or over a drink during a 'borrel': Marcello, Ricardo, Christopher, Massimo. Of course my 'roommates' not yet mentioned, some of whom are also involved in the lensing group: Allesandro, Hugo, Christiaan and Vivianni.

The most constant common denominator in acknowledgements of masters and PhD theses, and rightfully so, is the amazing support at Leiden. The unequalled 'computer group', as scientists and programmers crucial to the succes of many research projects, especially

Erik, David and Aart, who have been there since my beginning. The institute management and 'secretaries', who have been so much more than that for many at the institute, such as guardian angels and personal support: Jeanne, Kirsten, Evelijn, Marjan, Monica, Alexandra.

I have experienced the same inspiration, support and inclusive environment from the observatory community as a whole. Inspiring lectures, exchange of ideas during coffee, personal advice in the 'wandelgangen', support from the PBC, or as TA. During my 16 years as PhD, I have become convinced that the better I became at teaching, the better I became at science, and vice versa. I take away a mixture of personal gratitude and inspiring teaching and science examples from: Paul, Jarle, Peter, Xander, Antony, Ignas, Tim, Joop, George, Vincent, Frank, Rudolf, and the always optimistic Jan; Gé Nienhuis, Pierre van Baal, and Robert Terwiel from the physics department; and Jo Hermans, honorary member of my student association De Leidsche Flesch.

The same goes for some memorable teachers in high school: Jos, Leo, Harry, René, Joop, Anton, en Jos. I hope I can in my own modest way contribute to students' developments in learning and science, the way you did for me. I enjoy a great group of colleagues at school and take inspiration from working together, with a special mention for the colleagues math and natural sciences. I'm grateful for being a member of the 'havo-vwo' and statistics groups of the NVvW and our group of IO's and SO's of 4OLS, both allowing me to see education in new and refreshing ways.

I've made friends at the student association De Leidsche Flesch, with whom I love to debate science, society and ethics, between tasting delicious wines, whiskies, cheese and dinners: Jasper, Marinus, Jacob, Erwin, Pim, Frédérique, Ellen, Erik, and Vera. In fact, I want to thank Jasper explicitly, for bringing the recent article from Norris & Norris (2021) to my attention, when he was writing a column about it (Financieel Dagblad, 23 januari 2021).

Volleyball has been a source of friendship, special moments and the much needed release, after which I loved to enjoy beer and 'bitterballen' or sometimes a fine whisky, with Fonz, Daaf, Yannick, Wouter, and Timon, among others. Another great way to restock on much needed energy was playing tabletop games, D&D, or 'klaverjassen': Arno, Roy, Frank, Jochem, George, Maarten, Jan-Pieter, Nathan.

Steven, no matter how long we haven't spoken, as soon as we sit down for a beer, we pick up right where we left off. Paul and Daphne have not only been the best neighbours you can imagine, but also amazing friends, companions during skiing and sailing and trying to beat Pandemic Legacy together. Jelle and Hanne have been our partners in crime for tabletop escape rooms and Chronicles of Crime, but most of all friends and confidants. Rot and Lisette, we may be journeying through Middle Earth at the moment, but we're not done exploring Terrinoth, Golarion, the House on the Hill or battling zombie hordes or mad cultists. Rot, you're my longest and best friend and I'm grateful to have you as witness to my marriage.

Natuurlijk ben ik wie ik ben, dankzij de rest van mijn familie en schoonfamilie. Anneke en Alois, voor hun steun en warmte. Mijn grootouders, voor aan de ene kant de vele dagjes uit en aan de andere kant voor misschien iets van de Wortelgenen. Mijn zus Arwen, die altijd aan mijn kant heeft gestaan, haar man Sjoerd en hun kinderen in wie ik zoveel herken. Bovenal mijn ouders, die ons altijd geleerd hebben zelf na te denken, ons de gelegenheid hebben gegeven te studeren en onze nieuwsgierigheid in onze kinderjaren altijd hebben geprikkeld. Hun nuchterheid en hun humor. Het je kunnen verliezen in verhalen en het met beide benen op de grond staan.

En tenslotte mijn eigen gezinnetje. Annelie, je bent het grootste wonder dat ik in mijn

leven heb mogen meemaken. Ik zou niet weten hoe de wereld zonder jou had kunnen bestaan. Ik gun je eindeloze nieuwsgierigheid en ontdekkingen.

Marije, je hebt alles in je wat ik nodig heb. Zorgzaamheid, liefde, een uitdagende gesprekspartner, sportiviteit en enthousiasme, veiligheid en steun, en de moeder van onze klauterkabouter. De laatste meer-dan-11 jaar heb ik dit als hobby en als passie naast een betaalde baan kunnen afmaken en dat was zonder jouw geduld en steun niet mogelijk geweest.