UNIVERSITY COLLEGE LONDON

Faculty of Mathematics and Physical Sciences

Department of Physics & Astronomy

STATISTICAL METHODS IN WEAK GRAVITATIONAL LENSING

Thesis submitted for the Degree of Doctor of Philosophy of the University of London

by

Tomasz Kacprzak

Supervisors:

Prof Sarah Bridle

Prof John Shawe-Taylor

Examiners:

February 9, 2015

To my Mum, Dad and Brother

I, Tomasz Kacprzak, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis. Specifically:

- Chapter 3 consists of the paper "Measurement and Calibration of Noise Bias in Weak Lensing Galaxy Shape Estimation" by T. Kacprzak, J. Zuntz, B. Rowe, S. Bridle A. Refregier A. Amara, L. Voigt and M. Hirsch. It was published in MNRAS, 2012, volume 427, page 2711.

- Chapter 4 is the paper "Sérsic galaxy models in weak lensing shape measurement: model bias, noise bias and their interaction" by T. Kacprzak, S. Bridle, B. Rowe, L. Voigt, J. Zuntz, M. Hirsch and N. Maccrann. This publication appeared in MNRAS volume 441, page 2528.

- Chapter 5 is a paper in preparation by T. Kacprzak, B. Rowe and C. Heymans. Working title of this paper is "Detection of weak lensing by dark matter filaments in CFHTLenS".

- Appendix A consists of my contribution to the paper "IM3SHAPE: A maximum-likelihood galaxy shear measurement code for cosmic gravitational lensing" by J. Zuntz, T. Kacprzak, L. Voigt, M. Hirsch and B. Rowe and S. Bridle. This paper appeared in MNRAS 2013, Volume 434, page 1604.

- Appendix B presents my contribution to the publication "Mass and galaxy distributions of four massive galaxy clusters from Dark Energy Survey Science Verification data" by P. Melchior, E. Suchyta, E. Huff, M. Hirsch, T. Kacprzak, E. Rykoff, D. Gruen, and 87 authors. It was submitted to MNRAS, and is available online at `http://arxiv.org/abs/1405.4285`.

- Appendix C shows my contribution to a validation of GALSIM code. A paper describing this code, titled "GALSIM: The modular galaxy image simulation toolkit" by B. Rowe, M. Jarvis, R. Mandelbaum, G. M. Bernstein, J. Bosch, M. Simet, J. E. Meyers, T. Kacprzak, R. Nakajima, J. Zuntz, H. Miyatake, J. P. Dietrich, R. Armstrong, P. Melchior, M. S. S. Gill is availble online at `http://arxiv.org/abs/1407.7676`. It was submitted to MNRAS.

Abstract

This thesis studies several topics in the area of weak gravitational lensing and addresses some key statistical problems within this subject. A large part of the thesis concerns the measurement of galaxy shapes for weak gravitational lensing and the systematics they introduce. I focused on studying two key effects, typical for model-fitting shape measurement methods. First is *noise bias*, which arises due to pixel noise on astronomical images. I measure noise bias as a function of key galaxy and image parameters and found that the results are in good agreement with theoretical predictions. I found that if the statistical power of a survey is to be fully utilised, noise bias effects have to be calibrated. The second effect is called *model bias*, which stems from using simple models to fit galaxy images, which can have more complicated morphologies. I also investigate the interaction of these two systematics. I found model bias to be small for ground-based surveys, rarely exceeding 1%. Its interaction with noise bias was found to be negligible. These results suggest that for ongoing weak lensing surveys, noise bias is the dominant effect.

Chapter 5 describes my search for a weak lensing signal from dark matter filaments in CFHTLenS fields. It presents a novel, model-fitting approach to modelling the mass distribution and combining measurements from multiple filaments. We find that CFHTLenS data does provide very good evidence for dark matter filaments, with detection significance of 3.9σ for the filament density parameter relative to mean halo density of connected halos at their R_{200}. For 19 pairs of the most massive halos, the integrated density contrast of filaments was found on a level of $1 \cdot 10^{13} \mathrm{M}_{\odot}/\mathrm{h}$.

The appendices present my contribution to three other papers. They describe practical applications of the calibration of noise bias in the GREAT08 challenge and the Dark Energy Survey. I also present the results of the validation of reconvolution and image rendering using FFTs in the GalSim toolkit.

Contents

List of Figures

This page was intentionally left blank

List of Tables

Chapter 1

Introduction

Observational evidence from multiple cosmological probes point towards a theory in which our Universe started with a hot big bang, and has been expanding ever since. Moreover, in recent times, the expansion rate seems to be accelerating. This theory, *The Standard Model of Cosmology* is called ΛCDM (Λ - Cold Dark Matter). It is built upon several pillars: (i) Einstein's theory of General Relativity, (ii) The Cosmological Principle, (iii) dark energy, manifested by the expansion of the Universe, and (iv) the existence of cold dark matter, and (v) a period of rapid expansion after the Big Bang, called Inflation. In this chapter I will briefly present observational evidence for the cosmological paradigms and then describe the basics of the theory behind the standard model.

1.1 Observational evidence

This section presents the most important observational data sets which all seem to point towards the standard model of cosmology. These include recession velocities of type Ia supernovae, abundance of light elements, cosmic microwave background, distribution of large scale structure and observational effects attributed to dark matter: galaxy rotation curves and gravitational lensing, among others. This list is by no means complete, as there are many more types of observations which can be used to study cosmology. In this chapter I outline only the most important and relevant to the work presented in this thesis.

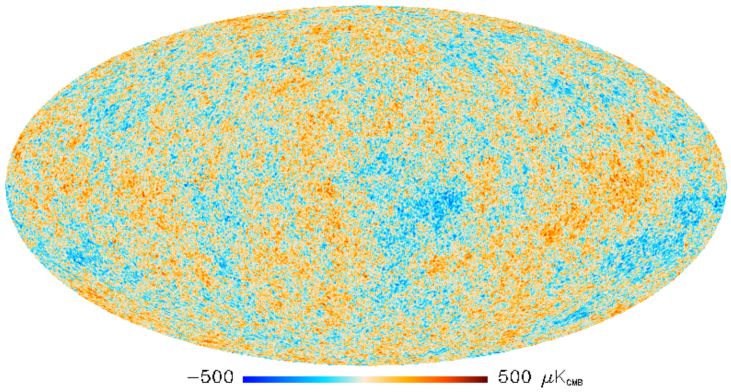

Figure 1.1. Cosmic microwave background temperature anisotropies, as seen by the Planck satellite (Planck Collaboration 2014*a*).

Inflation

Observations of the Cosmic Microwave Background (CMB) (Penzias & Wilson 1965; Mather et al. 1990; Spergel et al. 2003; Planck Collaboration 2014*a*) are a fundamental data set for cosmology. Its striking property is the fact that its temperature is constant across the sky to 1 part in 10^5. The uniformity of this temperature gives rise to the *horizon problem*: why is it that parts of the sky which could not have been causally connected before show such uniform distribution? The theory of Inflation provides a compelling answer to this question. It states that shortly after the beginning of the Universe, space underwent a phase of rapid accelerated expansion, which put previously bound plasma out of causal contact. Tiny fluctuations in the CMB temperature (Fig. 1.1) have a typical spatial scale, around 1°, and serve as *standard rulers*. These rulers measure the size of causally connected regions at the time when CMB was imprinted on the sky and have information used to constrain cosmological models.

Big Bang Nucleosynthesis

A compelling evidence that creates a foundation of Big Bang cosmology is the abundance of light elements (Olive et al. 1999; Kawasaki et al. 2005) . Until around one second after the Big Bang, matter formed of protons and neutrons was very dense. As the Universe expanded and cooled, protons and neutrons combined together and created helium isotopes H-3 and H-4 and deuterium, as well as lithium. This process is called the Big Bang nucleosynthesis. Theoretical calculations for these nuclear processes predict that about

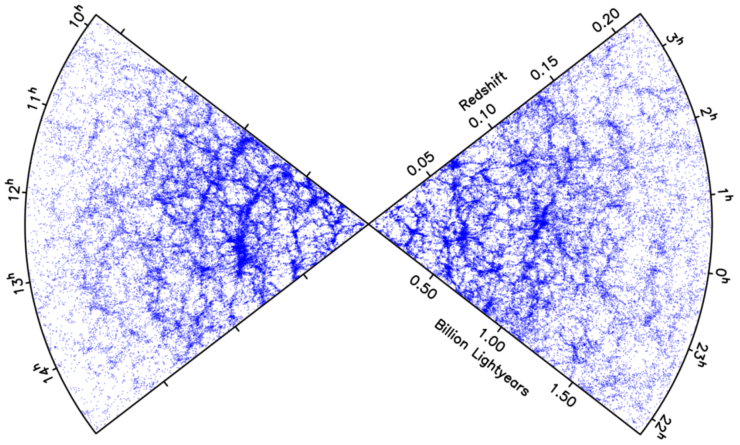

Figure 1.2. Galaxy distribution in 2dF survey (Colless 1998). Structures in this distribution are clearly visible. Statistics of these structures can be used to test cosmological models.

25% of the Universe consists of H-4. Recent stellar observations agree with theoretical predictions to great accuracy.

Structure formation

The distribution of matter in the Universe, in the visible and invisible forms, also supports the concordance model (Colless 1998; Eisenstein et al. 2005; Tegmark et al. 2004) . About 10,000 years after the Big Bang, the tiny fluctuations in matter density started to collapse under the force of gravity. The theory of this process is also well understood and can predict the statistical properties of the matter distribution. Figure 1.2 shows the distribution of galaxies as measured by the Two Degree Field survey (2dF). Large scale, web-like structures in galaxy distribution are very prominent.

Dark matter

One of the fundamental concepts in cosmology is the existence of dark matter. It was first postulated after a discovery of increased rotation velocities of galaxies at far radii from the galactic cores (e.g. Rubin et al. 1980). The baryonic content of galaxies was not enough to explain the shape of the rotation curves, and adding an invisible mass provided a good explanation. Perhaps the most convincing evidence for this invisible form of mass, which also characterises its properties, is the distribution of mass and X-ray emitting gas in the Bullet Cluster (e.g. Markevitch 2005; Clowe et al. 2006), as shown in Fig. 1.3. This system contains two clusters undergoing a merger. They have already collided once; after

Figure 1.3. The Bullet Cluster. The blue colour shows dark matter density distribution, measured using gravitational lensing. The red color is the X-ray emitting gas. Composite Credit: X-ray: NASA/CXC/CfA/ M.Markevitch et al.; Lensing Map: NASA/STScI; ESO WFI; Magellan/U.Arizona (Clowe et al. 2006). Optical: NASA/STScI; Magellan/U.Arizona (Markevitch 2005).

the collision, the gas has decelerated much faster than the dark matter due to pressure of collision, and is visible between dark matter peaks. Dark matter clumps, which do not interact by pressure, went straight through each other, and are decelerated only by gravitational force. This important property of dark matter plays a key role in modern cosmological models.

Dark energy

Evidence for an accelerating Universe is provided by observations of distant type Ia supernovae (Perlmutter et al. 1998; Riess et al. 1998), amongst others. Supernovae are treated as standard (or standardizable) candles, which emit light with the same intensity independently of their distance. This can be used to calculate the physical distance to an exploding star. Observation of its spectrum can provide information about its recessional velocity. This dependence is called Hubble's Law (Hubble 1929). Recent observation of supernovae supply evidence that the most distant ones move away from us with increasing velocity (Fig. 1.4). This recession is attributed to the expansion of space between us and the star. A cause for this expansion is called *dark energy* and its nature is one of the biggest unsolved questions in cosmology.

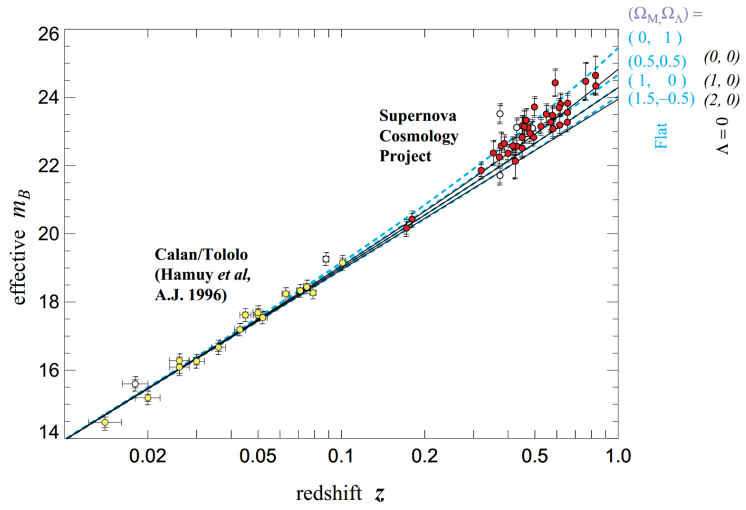

Figure 1.4. Hubble diagram for 42 high-redshift Type Ia supernovae from the Supernova Cosmology Project, and 18 low-redshift Type Ia supernovae from the Calan/Tololo Supernova Survey (Perlmutter et al. 1998). High redshift supernovae are receding from us with increasing velocity. Data is fit using several cosmological models, and prefers a dark energy dominated universe.

Evidence for dark energy also comes from CMB and large scale structure observations. Measurements of CMB anisotropies indicate that space is close to being flat. Observations of the properties of galaxy clusters suggest that, if space is indeed flat, the matter (both dark and baryonic) must consist only of $\approx 30\%$ of the total mass-energy in the Universe. Thus, to reconcile the flatness of space and low matter density today, dark energy has to be incorporated into the cosmological model.

1.2 Theory of General Relativity

Einstein's attempts to reconcile Newton's theory of gravity and his own theory of Special Relativity led to formulation of the *Theory of General Relativity*. The proposal was that gravity is not a force in a conventional sense, but rather a manifestation of the curvature of spacetime, induced by the presence of matter. This theory addressed two main discrepancies between Newton's theory and Special Relativity. Firstly, in Newton's theory, the gravitational potential Φ responds instantaneously to a disturbance in matter density ρ, which violates Special Relativity's rule that no signal should travel faster than the speed of light. Secondly, it naturally incorporated the *Equivalence Principle*, a very precisely confirmed observational rule which states the equality of inertial and gravitational mass.

Einstein's field equations describe how the curvature of spacetime relates to the distri-

bution of matter at some event. In this theory, matter and energy distribution is described as the *energy - momentum tensor*, which, for a perfect fluid, incorporates the total *energy density* ρc^2 (including all potential energy contributions from forces on particles and their kinetic energy from thermal motion), and isotropic pressure p in three-dimensional space. This tensor, for a inertial reference frame, is shown in Eqn. 1.1.

$$[T^{\mu\nu}] = \begin{bmatrix} -\rho c^2 & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{bmatrix} \tag{1.1}$$

For a general fluid, off-diagonal terms are non-zero, and incorporate viscous stresses, heat conduction and momentum of energy conducted by it.

Geometry of the spacetime affected by gravitating matter is described in terms of the metric tensor $g_{\mu\nu}$, Ricci Tensor $R_{\mu\nu}$, and curvature scalar R, so that *Einstein's field equation* is given by 1.2

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \frac{8\pi G}{c^4} T_{\mu\nu} \tag{1.2}$$

where G is the gravitational constant. This equation can be then used to find equation of motion of a particle. Note that in the absence of matter (and radiative energy and momentum), $T_{\mu\nu}$, the curvature vanishes, $R_{\mu\nu}$. Another property is that, in the weak field limit, these equations in most cases simplify to Newton's gravity. Gravitational Lensing is an interesting case where a discrepancy exists between Einstein's and Newthon's descriptions; the bend angle of light is two times larger in relativity.

1.3 Cosmological constant and dark energy

Einstein's field equations are not unique. A *cosmological constant* term can be added to Eqn. 1.2 and the field equation will still be consistent with energy conservation $\nabla_\mu T_{\mu\nu} = 0$, such that (Eqn. 1.3)

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}. \tag{1.3}$$

The cosmological constant term acts like a gravitational repulsion, whose strength increases linearly with distance (see Carroll 2000; Martin 2012 for recent reviews). It is postulated as a possible explanation for the expansion of the Universe. This is a unique 'substance' with properties never encountered before. First, its equation of state includes negative pressure $p = -\rho c^2$. Secondly, it depends only on the metric tensor $g_{\mu\nu}$, which indicates that it is a property of vacuum itself. Thirdly, it fixes the energy density of the vacuum, so that $\rho_{\text{vac}} = \frac{\Lambda c^2}{8\pi G}$.

Calculation of vacuum energy, dubbed *dark energy* for its mysterious nature, is one of the biggest unsolved problems in physics. It disagrees drastically with calculations from quantum mechanics; the sum of all zero-point energies of all fields known in nature differs from limits on Λ set from cosmological observations by 120 orders magnitude! Different phenomenological models of dark energy are postulated: a *fifth force*, where dark energy is a scalar field, which can evolve with time (see Copeland et al. 2006 for review), *modified gravity* theories state that accelerated expansion can be a result of gravity laws differing from General Relativity on large scales (see Clifton et al. 2012 for review).

1.4 The Friedman - Lemaître - Robertson - Walker geometry

The Standard Model of Cosmology is a description of the universe based on General Relativity, which includes several key assumptions based on observational evidence. The First is the *cosmological principle*. It states that the Universe is *homogeneous* on large scales (the matter distribution is uniform) and *isotropic* (there is no preferred point in space). In other words, at any time, the Universe looks statistically the same from all positions in space, and all directions are equivalent.

Assuming idealised *fundamental observers*, which have no peculiar velocity with respect to surrounding matter on large scale, a *comoving coordinate system* can be created. In this system, the proper time along each observers worldline is equal to its time coordinate ($x^0 = \tau$) and orthogonal to spatial coordinates. An expanding universe can be described in terms of a scale factor $a(t)$ depending only on cosmic time (today $a(t_0) = 1$). This scale factor describes a change in physical distance between fundamental observers now and at time t.

Another consequence of assuming the cosmological principle is a restriction on the

curvature of spacelike coordinates. Homogeneity and isotropy force this space to be a *maximally symmetric 3-space*. The three-dimensional curvature tensor R_{ijkl} for such space will be dependent only on one number - curvature K. If $K = 0$ then the space is *flat* (Euclidean), if $K > 0$, the space is *closed* (a 3-sphere), and $K < 0$ indicated the space is *open* (a hyperboloid). Recent cosmological observations suggest that our Universe is flat and this is an assumption often made in cosmology and astrophysics (see Lahav & Liddle 2014 for review on cosmological parameters).

These assumptions lead to a formulation of the Friedman-Robertson-Walker metric

$$ds^2 = c^2dt^2 - a^2(t)\left[r^2 + f(r)^2(d\theta^2 + \sin^2\theta d\phi^2)\right] \tag{1.4}$$

where r, θ, ϕ are spherical polar coordinates, and r is a comoving distance, which depends on the curvature of the spacetime in the following way

$$f(r) = \begin{cases} \frac{1}{\sqrt{K}}\sin\left(r\sqrt{K}\right) & K > 0 \\ r & K = 0 \\ \frac{1}{\sqrt{-K}}\sinh\left(r\sqrt{-K}\right) & K < 0 \end{cases} \tag{1.5}$$

Using Einstein's field equations 1.2 with the FRLW metric results in two unique *cosmological field equations*

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{Kc^2}{a^2} \tag{1.6}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{c^2}(\rho + 3p) \tag{1.7}$$

These are called *Friedman-Lemaître-Robertson-Walker equations* and describe the time evolution of the scale factor.

1.5 Cosmological redshift

In the Standard Model of Cosmology, with gravity described by General Relativity and an expanding universe with FLRW metric, there are only a few cosmological parameters which will describe the entire history of the Universe. Historically, the first parameter to be measured was the Hubble parameter, which describes the rate of expansion.

An expanding space will cause photon wavelengths to be stretched (so that frequency

$\nu_{\text{emit}} \rightarrow \nu_{\text{receive}}$), in a way related to the scale factor. Assuming the scale factor changes slowly with cosmic time, the FRW metric can be used to derive:

$$1 + z = \frac{\nu_{\text{emit}}}{\nu_{\text{receive}}} = \frac{a_{\text{receive}}}{a_{\text{emit}}} \tag{1.8}$$

Using the FRW metric, the recessional velocity v of a nearby galaxy can be directly related to its distance d via

$$v = cz = Hd \tag{1.9}$$

where the *Hubble parameter* defines the expansion rate $H = \frac{\dot{a}}{a}$, and its measurements at the present day indicate that it is close to $H_0 = 70$ km s^{-1} Mpc^{-1} (Planck Collaboration 2014*a*). The formulation of this *Hubble's law* in 1929 following the observation of Cepheids was a starting point of Big Bang Cosmology.

1.6 Cosmological fluids

Other cosmological parameters describe the content of the Universe, in terms of density of its components, called cosmological fluids. If we define the relationship between pressure and energy density for a cosmological fluid as an *equation of state*

$$p = w\rho c^2 \tag{1.10}$$

we can use the Friedman equations to establish a relationship

$$\rho \propto a^{-3(1+w)} \tag{1.11}$$

which gives the evolution of density ρ of a cosmological fluid in terms of the scale factor a. But what are these cosmological fluids and how do they behave? I will briefly describe three groups of density components: matter ρ_m, radiation ρ_r and vacuum energy ρ_Λ.

1.6.1 Matter

Matter is a term commonly referring to both baryonic and dark matter, both of these species are non - relativistic (in the Cold Dark Matter model). Current measurements indicate that Dark Matter consists of 25% of the total matter-energy of the Universe, and

Baryonic matter only 5%.

Baryonic matter is all of the 'ordinary' particles. Cosmologists refer to Baryons more widely than particle physicists, including leptons. The important property of baryons is that they are visible; they interact with light and can be observed directly. Most of the baryons in the Universe were created during the Big Bang Nucleosynthesis (BBN), which happened in the first 100 seconds after the Big Bang. The Universe is dominated by light elements created at the BBN: Hydrogen (74%), Helium (26%), and traces of metals (also produced later in stars) (Olive et al. 1999; Coc et al. 2013).

Dark matter is believed to be a particle beyond the standard model of particle physics, although it has not yet been detected directly. It has a property that it does not interact strongly with light, except via gravity, and it can only be observed indirectly, through, for example, measurements of gravitational lensing (see Sec. 2) or measurement of galactic rotation curves.

Dark matter is believed to be in a form of *dust*: its thermal energy is much smaller than its rest energy, and so it can be considered to have no pressure. Non-relativistic baryonic matter can also be approximated as pressureless. This requirement sets $w_m = 0$ and from 1.11 we calculate that its density evolves as $\rho_m \propto a^{-3}$, and scales with redshift as $(1+z)^3$.

1.6.2 Radiation

Radiation consists of a group of relativistic particles: neutrinos and photons, which have very small or zero rest mass. There is evidence of neutrinos having mass; however it is very small, close to 0.3 eV (Goobar et al. 2006; Thomas et al. 2010; Battye & Moss 2014). Neutrinos, as well as photons, have the equation of state $w_r = 1/3$. Their density scales with redshift as $(1+z)^4$, the additional factor of $(1+z)$ comes from the redshifting of each photon. Most photons and neutrinos constituting the radiation energy of the Universe come from the cosmic microwave background (CMB), with a present day temperature of 2.726K. Radiation is believed to constitute around 10^{-4} of the total density, which is a tiny fraction compared to earlier times (Planck Collaboration 2014*b*). This massive change is due to the fast decaying $(1+z)^4$ scaling factor.

	w	$\rho(a) \propto$	$\Omega_j \approx$
radiation	1/3	a^{-4}	10^{-4}
matter	0	a^{-3}	0.3
curvature	-1/3	a^{-2}	0
cosmological constant	-1	a^0	0.7

Table 1.1. Evolution of density of components with scale factor. Last column contains the fraction of critical density for each component today. Values Ω_j are approximate, but in agreement with recent cosmological observations.

1.6.3 Vacuum

As already described in Sec. 1.3, the vacuum is a perfect fluid with $p = -\rho c^2$, which indicates negative pressure. The equation of state of the vacuum is therefore $w = -1$. Recent cosmological observations estimate the contribution to total energy density as high as 70% today. The reason for this domination is that ρ_Λ does not change with time, and other forms of energy decrease their densities as the Universe expands. In alternative phenomenological dark energy models, w is allowed to change as a function of scale factor. The simplest modification is to use a linear expansion $w(a) = w_0 + (1 - a)w_a$. There is a significant effort nowadays to measure the dark energy equation of state parameters w_0 and w_a, including with weak gravitational lensing tomography (see section 2.4).

Having established the types and properties of cosmological fluids, we can now trace the composition of the density back in cosmological time. It is a common practice to work with dimensionless density parameters defined as a fraction of the *critical density*

$$\Omega_j(t) = \rho_j(t)/\rho_{cr}(t) \tag{1.12}$$

where $\rho_{cr} = \frac{3H^2(t)}{8\pi G}$ is the critical density, which can be derived from first Friedman Equation 1.6, and j stands for a component fluid. Fractional contribution of each component to the total critical density today Ω is noted as Ω_m, Ω_r, Ω_Λ, excluding the additional contribution from the curvature Ω_k. Table 1.1 shows the evolution of cosmological fluids and gives the observed values for density parameters today.

Using the measured values for these cosmological parameters, we can plot the density of each of the components as a function of scale factor, starting from the Big Bang. Figure 1.5 presents the history of the composition of the Universe. We notice that, shortly after the Big Bang, the Universe was **radiation dominated**, then it transitioned to being

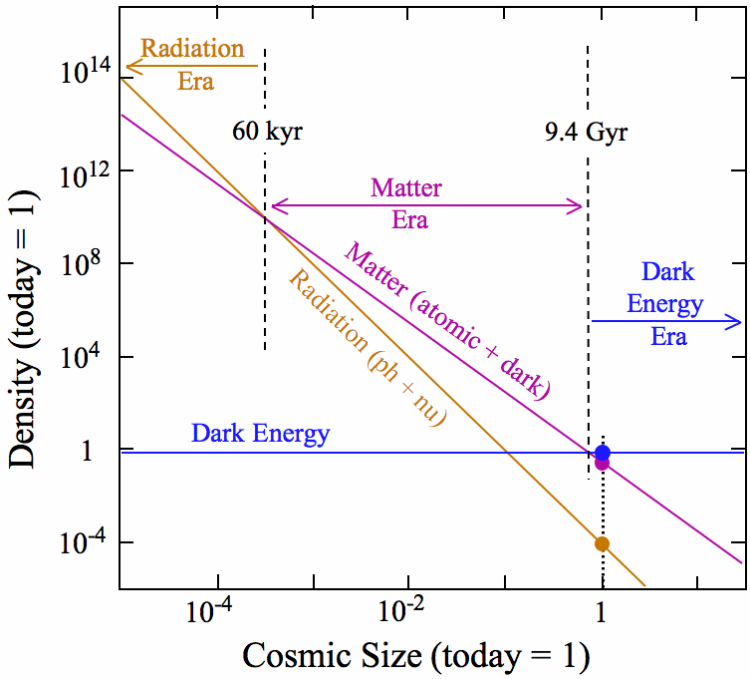

Figure 1.5. The evolution of composition of the total density of the Universe as a function of scale factor. Credit: Mark Whittle, University of Virginia.

matter dominated, and finally, only recently to being **dark energy dominated**.

Having outlined the big picture of the history of the Universe, now is a good time to take a look at the most significant cosmological events happening during these periods. Shortly after the Big Bang the Universe was in a form of hot, dense plasma. Its distribution was homogeneous, disrupted only by tiny quantum fluctuations. Photons were not able to travel freely, as they were constantly scattering electrons. No particles could form, as the bonds were immediately destroyed by photons. As the Universe expanded, the energy of photons decreased. When the scale factor was around ten orders of magnitude smaller than today, the first nuclei started to form. This *Big Bang Nucleosynthesis* created hydrogen, helium and lithium (and their isotopes). The observed abundance of these elements today matches the theoretical predictions remarkably well, and this fact is one of the main observational pillars of modern cosmology. In the radiation domination era, no structures in matter could be formed, as the gravitational collapse of inhomogeneities was counteracted by pressure from radiation. A notable time in this timeline is the *matter-radiation equality*, when $\Omega_m = \Omega_r$, which happened around redshift $z_{EQ} \sim 3100$. After this transition, when the background temperature was around 3000° K (redshift of 1000) the density fell enough so that electrons started to bind with protons to create first atoms.

This time is called the *recombination*. This was followed by *photon decoupling*, and photons to be able to stream freely. These photons reach us today at temperature of 3° K in a form of background radiation, called the cosmic microwave background (CMB). The CMB has a black-body spectrum and its temperature is uniform to one part in 10^5. Tiny anisotropies in the CMB are a result of oscillations started by quantum fluctuations at the time of Inflation. Anisotropies in this background are another very important observable, matching the theoretical predictions very well. In the matter-dominated era, when the radiation pressure decreased, the inhomogeneities started to grow under gravitational collapse. Dark matter began to collapse into halos and baryons started to create stars and galaxies. Photons travelling through these overdense regions started to be subject to gravitational lensing.

1.7 Early Universe

Until this point we assumed that the Universe was completely homogeneous and isotropic; it is clearly a great simplification as the Universe we observe does in fact have inhomogeneities. All structures: clusters, galaxies, planets, even us, are a consequence of the *perturbations* in the initial distribution of mass-energy. These perturbations come from quantum fluctuations at the time of the Big Bang, and can be described simply as

$$\delta = \frac{\delta(\mathbf{x}) - \bar{\rho}}{\bar{\rho}} \tag{1.13}$$

where $\bar{\rho}$ is the mean density in the Universe. No matter how small, these perturbations will start to grow due to gravity. An overdensity will only affect matter in a radius which is causally connected to it. This radius is called *the horizon*, and corresponds to comoving scales greater than $(aH)^{-1}$. Hubble time $H^{-1} \propto t$ corresponds to the time since the beginning of expansion, and the factor a converts to comoving coordinates. However, observational measurements of the CMB show that its temperature is uniform to one part in 10^5. This is unexpected, since separate parts of the sky could not have been in causal contact before, according to our picture until now. This is in fact called *the horizon problem*. A solution to this problem, which also gives a set of initial conditions for the density perturbations, is provided by a group of theories called *inflation*. Inflation is a period of accelerated expansion in the Early Universe, which increases the scale factor rapidly by more than 27 orders of magnitude. This process rapidly puts a previously

connected region out of causal contact, which will result in the homogeneous temperature of CMB we observe today. It is believed to be driven by a scalar field called *inflaton*, which undergoes a phase transition at the end of the inflation era and decays into other particles. Most theories of inflation predict the spectrum of initial density perturbations, called the *primordial power spectrum*, to be almost constant and in a form of power law $\mathcal{P}_{\rm s}(k) \propto k^{n_{\rm s}-1}$, where $n_{\rm s}$ is the *spectral index*. Most recent measurements of $n_{\rm s}$ are close to 1 (Planck Collaboration 2014*b*). These perturbations provide initial conditions for their further evolution. The evolution of each Fourier mode $\delta(k, \eta)$ can be traced as a function of cosmic time, and we can predict its value today.

1.8 Structure formation

Structures in the Universe can be described in terms of the isotropic *power spectrum* $P(k) = \langle|\delta_k|^2\rangle$. Using the standard model and inflation theory, the matter power spectrum today can be calculated, and compared to observations. To calculate it, its evolution has to be traced from the primordial power spectrum, through radiation- and matter-dominated eras, to the present day. This evolution process is described by two functions: *transfer function* and *growth function* at late times. Schematically, it is written as

$$\Phi(k, z) = \Phi(k)_p \times T(k) \times G(z) \tag{1.14}$$

where subscript p corresponds to primordial time, T is the transfer function and G is the function describing scale-independent, linear growth. Using linear perturbation theory and following (Liddle & Lyth 2000), this growth is described as

$$\ddot{\delta} + 2\frac{\dot{a}}{a}\dot{\delta} = 4\pi G\rho\delta. \tag{1.15}$$

Amplitude of the transfer function $T(k)$ is controlled by a normalisation parameter σ_8. It is defined as the root mean square of matter density, which is smoothed with a top-hat filter sphere with a radius of 8 Mpc/h. Normalisation of the power spectrum is an important cosmological parameter due to being degenerate with the matter density Ω_m.

As the Universe expands, scales of increasing size will begin to re-enter the horizon. The transfer function for a mode $\mathbf{k}$ will depend on the time this mode enters the horizon. Let us highlight the three most important stages in this evolution, which will create the

characteristic shape of matter power spectrum today.

First let us consider modes outside the horizon. During the radiation domination era, these modes will have experienced a fixed growth at rate $\propto a^2$. During matter domination, this rate would have been $\propto a$. If a scale enters the horizon during the radiation-dominated era, its growth is completely suppressed by the pressure from radiation. If this entry happens after matter-radiation equality, growth continues as $\propto a$. Finally, non-linear growth will be affecting small scales to a degree which cannot be ignored, starting from scales $k_{NL} \simeq 0.2\ h\ \mathrm{Mpc}^{-1}$ for most models.

Perturbation terms will also enter the geometry part of Einstein's equation 1.3. On the geometry side, the metric is decomposed in to *scalar* and *tensor* parts. The scalar part couples to the density of matter and radiation and is responsible for most of the inhomogeneities and anisotropies. Tensor perturbations, called *gravitational waves*, do not couple to density and are not responsible for evolution of large scale structure. Here, I will consider only the scalar perturbations. In the Conformal Newtonian Gauge, the scalar metric in the perturbed Universe is often written as

$$ds^2 = a^2(\eta)[(1+2\Psi)d\eta^2 + \delta_{ij}(1-2\Phi)dx_i dx_j] \tag{1.16}$$

where Ψ is the Newtonian potential (often called time-like) and Φ is the perturbation to spatial curvature (often called a space-like potential). In General Relativity, these two potentials are equal and related to the overdensity via the Poisson equation

$$k^2\Phi = 4\pi G\rho_m a^2\delta. \tag{1.17}$$

The prediction for the matter power spectrum today supported by observations is shown on Fig. 1.6. The characteristic peak of this spectrum corresponds to a mode which entered the horizon at the time of matter-radiation equality.

1.8.1 Sheets, filaments and halos

Formation of structure is more difficult to describe in the non-linear regime. Large N-body simulations are often used to calculate statistics of matter distribution on small scales. These simulations start from some initial conditions and then trace the displacement of all dark matter particles through cosmic time. Simulations reveal that the structure of

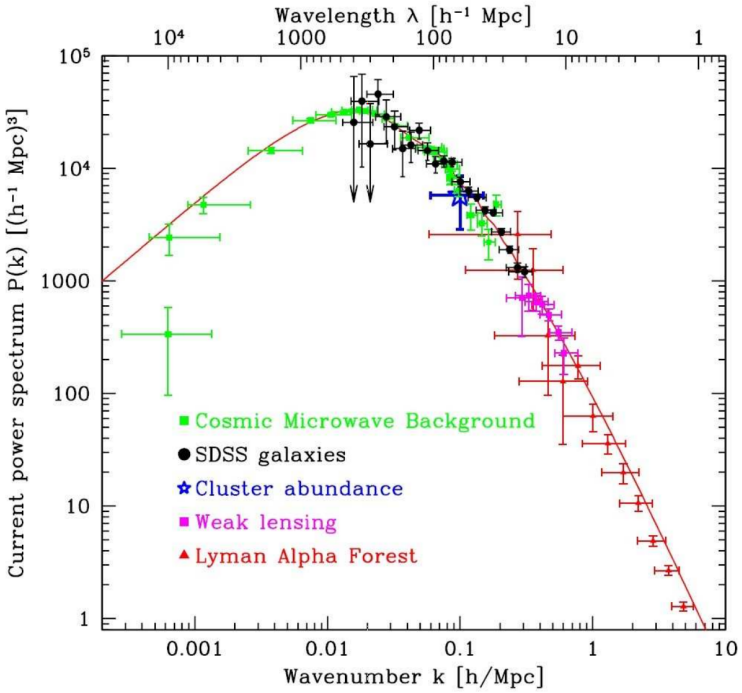

Figure 1.6. Matter power spectrum measurements from various cosmological probes. Source: Tegmark et al. (2006). The non-linear contribution, not shown in this figure, starts around $k \simeq 0.2$ and induces characteristic small 'wiggles' onto the spectrum shape. For more recent measurements from Planck and CFHTLenS, see for example (MacCrann et al. 2014).

matter resembles a web-like network, where peaks, called *halos*, are connected to each other by extended structures, called *filaments*. Fig. 1.7 shows a slice through the Millenium simulation (Springel et al. 2005), with a massive halo in the centre and several filaments connected to its neighbours.

There exist several useful approximations which allow us to understand the dynamics of structure formation without resorting to N-body simulations. The Zel'dovich approximation (Zel'dovich 1970) is an approach in which the initial displacement of dark matter particles is calculated, and its movement is assumed to continue in this initial direction. Proper coordinates $\mathbf{q}$ of the given particle are modified by a displacement field $\mathbf{f}$.

$$\mathbf{x}(t) = \mathbf{q}(t) + b(t)\mathbf{f}(\mathbf{q}) \tag{1.18}$$

where b scales the displacement function. The *deformation tensor* can be calculated as $\partial f_i/\partial q_j$. Collapse of the structure will take place first along the direction corresponding to the largest negative eigenvalue of this tensor. For a triaxial ellipsoidal distribution of

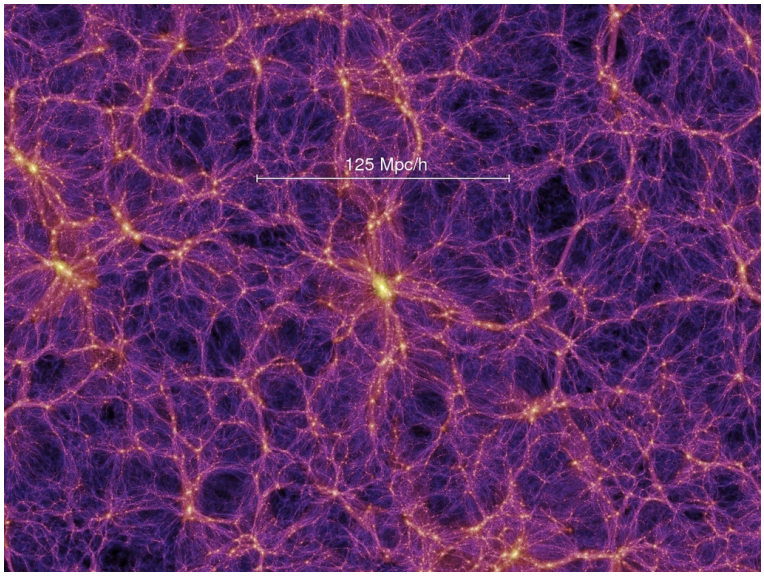

Figure 1.7. Cosmic web, consisting of dark matter halos, sheets and filaments. Image from Millenium simulation (Springel et al. 2005).

matter, the largest negative eigenvalue is aligned with the shortest axis of the ellipsoid. Such a distribution will collapse along this axis into a flattened structure called a *pancake*. Collapsing along the second largest eigenvalue will make this distribution take a form of an extended filament, with an overdensity forming in its centre. Finally, the entire structure will collapse into a concentrated *halo*. This simplified view of the formation of structure of the cosmic web seems to match well with simulations.

1.8.2 NFW halo profile

The spatial distribution of dark matter in halos is often modelled with a Navarro-Frenk-White profile (NFW) (Navarro et al. 1997). Radially averaged halos were found to have the same profile, independent of halo mass, initial density fluctuation spectrum, and the values of cosmological parameters. The radial profile of matter density follows

$$\rho(r) = \frac{\rho_0}{\frac{r}{R_s}\left(1 + \frac{r}{R_s}\right)^2} \tag{1.19}$$

where R_s is the scale radius and ρ_0 is the density, and these parameters vary among halos. The total mass of this distribution is divergent: it is customary to truncate it at a specific radius, called the *virial radius*. The virial radius is often taken to be the radius at which the average density encircled by it is 200 times the critical density of the Universe, and

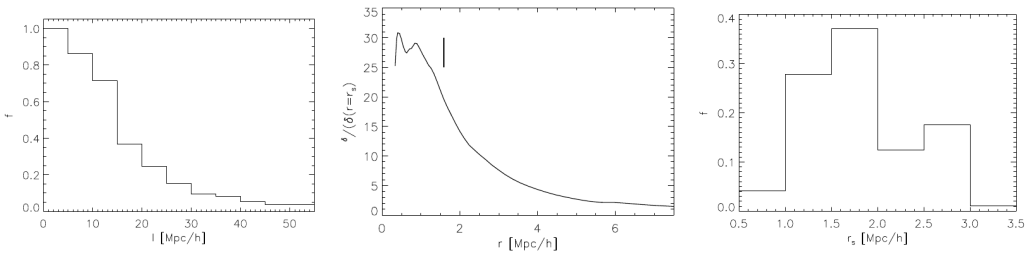

Figure 1.8. Properties of intercluster filaments, measured by Colberg et al. (2005). Left panel shows the fractional abundance of filaments as a function of their length. Middle panel presents the average radial density distribution for straight filaments. Distribution of scale radius parameters is shown in the right panel.

called R_{200}. Virial and scale radii are related via a *concentration* parameter c, so that $R_{vir} = cR_s$. The relation between the total mass of the halo (and hence R_{200}) and its concentration is studied using N-body simulations. For example, Duffy et al. (2011) uses large N-body simulations to determine this relation for WMAP cosmology. The mass-concentration relation is well described by a power law for halos with masses between $M_{200} = 10^{11}$–$10^{15}\mathrm{M}_{\odot}/\mathrm{h}$. In Chapter 5, a relation was used with the form of

$$c = \frac{5.72}{(1+z)^{0.71}}(M_{200}/10^{14})^{-0.081} \tag{1.20}$$

where z is the redshift of the cluster and M_{200} is the total mass within the virial radius.

1.8.3 Properties of filaments

N-body simulations have been used to study properties of filaments. Colberg et al. (2005) identified 228 filaments connecting clusters with masses $M_{200} > 10^{14}$. They found that halos separated by less than 5 Mpc/h are almost always connected by a dark matter filament, and their abundance falls as the distance increases. The left panel on Fig. 1.8 shows the fractional abundance of filaments for varying distance between halos. Moreover, when distance increases, filaments have an increasing probability of being wrapped; they do not connect halos in a straight line, but by a bent arch. For lengths greater than 15 Mpc/h, close to half of the filaments are wrapped.

Colberg et al. (2005) also measures the radial density profile of straight filaments (middle panel on Fig. 1.8). This profile follows an inverse square law starting from a

radius r_s. An equation which approximates this function well is

$$\rho(r) = \frac{\rho_0}{1 + r^2/r_s^2}. \tag{1.21}$$

Radius parameter r_s differs from one filament to another. The right panel on Fig. 1.8 shows the distribution of radii measured for straight filaments. It seems that filaments can vary in thickness, from 0.5-3 Mpc/h. Relatively little is known about filaments from direct observations. One of the first direct detections of dark matter filaments was made by Dietrich et al. (2012), who measured the mass distribution between A222/A223 clusters using weak gravitational lensing. For more on filaments see Chapter 5.

1.9 Recent measurements of cosmological parameters

To summarise the theory chapter, I will briefly describe the values of cosmological parameters as measured by the most recent cosmological probes. In many ways the measurements from all probes, such as supernovae, CMB, Baryon Acoustic Oscillations (BAO), and cosmic shear point towards the same region of parameter space. Measurements indicate that we are living in a Universe which is 13.8 Gyr old, and has recently transitioned from being matter-dominated to dark-energy-dominated, with cosmological parameters $\Omega_m = 0.31^{+0.16}_{-0.17}$ and $\Omega_\Lambda = 0.69 \pm 0.01$. These values come from a combination of: CMB temperature measurements from the Planck satellite, CMB polarisation analysis from the Wilkinson Microwave Anisotropy Probe (WMAP), high resolution CMB temperature measurements fromthe Atacama Cosmology Telescope (ACT) and South Pole Telescope (SPT) and BAO measurements from the Sloan Digital Sky Survey (SDSS), as reported in (Planck Collaboration 2014*b*). The left panel on Fig. 1.9 shows the remarkable agreement of multiple cosmological probes in the $\Omega_m - \Omega_\Lambda$ plane. These include slightly older measurements of CMB temperature from WMAP.

Recently, measurements of cosmic shear from the Canada - France - Hawaii Telescope Lensing Survey (CFHTLenS) has put competitive constraints on cosmological parameters. Lensing is especially useful on constraining the $\Omega_m - \sigma_8$ relation, as it breaks the degeneracy which exists between these two parameters in the CMB measurements. A combination of constraints from the CFHTLenS, the Baryon Oscillations Spectroscopic Survey (BOSS), WMAP CMB temperature and supernovae observations with the Hubble Space Telescope (HST) point towards measurements of $\sigma_8 = 0.794^{+0.016}_{-0.017}$, see middle panel on Fig. 1.9

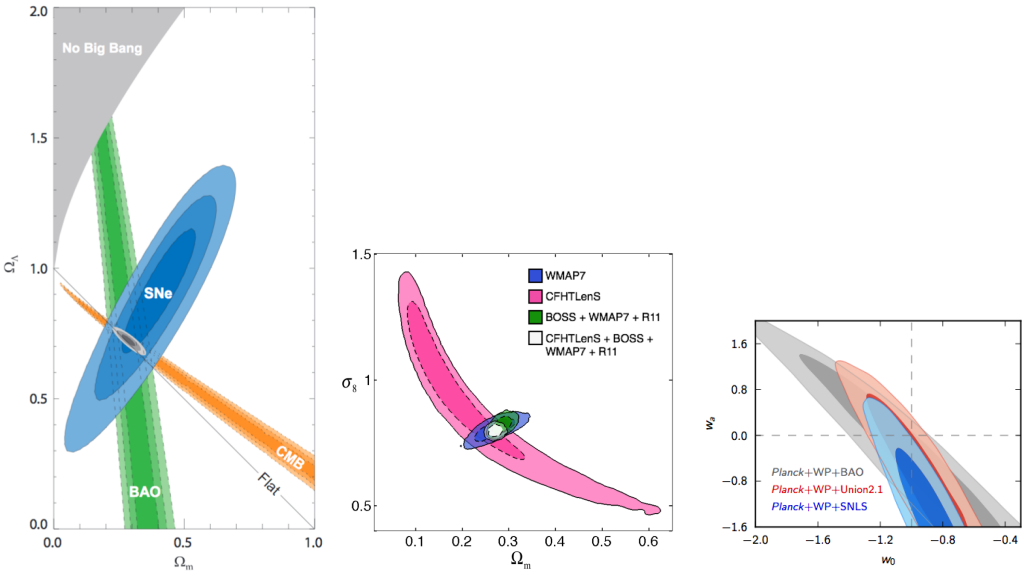

Figure 1.9. Constraints on cosmological parameters from various probes. Left panel comes from Lahav & Liddle (2014). This figure comes from 2008 and measurements of CMB temperature as measured from WMAP. Planck offers more precise measurements of CMB temperature. Middle panel comes from Heymans et al. (2013) and right panel from Planck Collaboration (2014*b*).

(Heymans et al. 2013).

Planck has also put constraints on the evolution of dark energy equation of state parameters. In the most simple parametrisation where the equation of state parameter is evolving as a function of the scale factor a as $w(a) = w_0 + w_a(1 - a)$, parameters w_0 and w_a were measured. Although there is no evidence of dark energy evolution, which corresponds to $w_0 = -1$ and $w_a = 0$, the precision on this measurement is low (Fig. 1.9, right panel). Cosmic shear is a probe with great potential for improving this measurement (Albrecht et al. 2006). The Dark Energy Survey (see section 2.8) aims to improve the quality of these constraints.

Chapter 2

Gravitational lensing

According to Einstein's theory of General Relativity, the presence of gravitating matter creates a curvature in the spacetime. In such space, photons travelling from a source to an observer will follow a curved trajectory. This results in the light ray being bent around gravitating matter. The bending of light rays can cause interesting effects on the images of galaxies emitting them: they can be distorted and magnified in size and brightness, their apparent positions can be changed. Depending on the configuration of the light source (background galaxy), gravitating mass (for example, a foreground dark matter halo) and observer positions, and the mass of the lens, the strength of the lensing effect can vary greatly, leading to three distinct types of lensing: strong lensing, weak lensing and microlensing.

The observer can register the image of the background source in multiple locations, smeared into a full or partial ring. Such phenomena are called Einstein crosses, or Einstein rings, respectively. These are examples of *strong gravitational lensing*. If the lensing effect is less pronounced, the image of the background galaxy can be slightly distorted, or *sheared*. This effect is called *weak gravitational lensing*. In fact, the majority of observed galaxies are weakly lensed, and a typical change the galaxy shape due to shear, measured by the major-to-minor axis ratio, is of order 1%. Finally, if the gravitating mass is very small (for example, a planet or a star), then gravitational lensing can be observed only in the apparent change in brightness of the source. For example, if a source passes around the lens periodically, then it will seem brighter when passing behind the lens. This effect is called

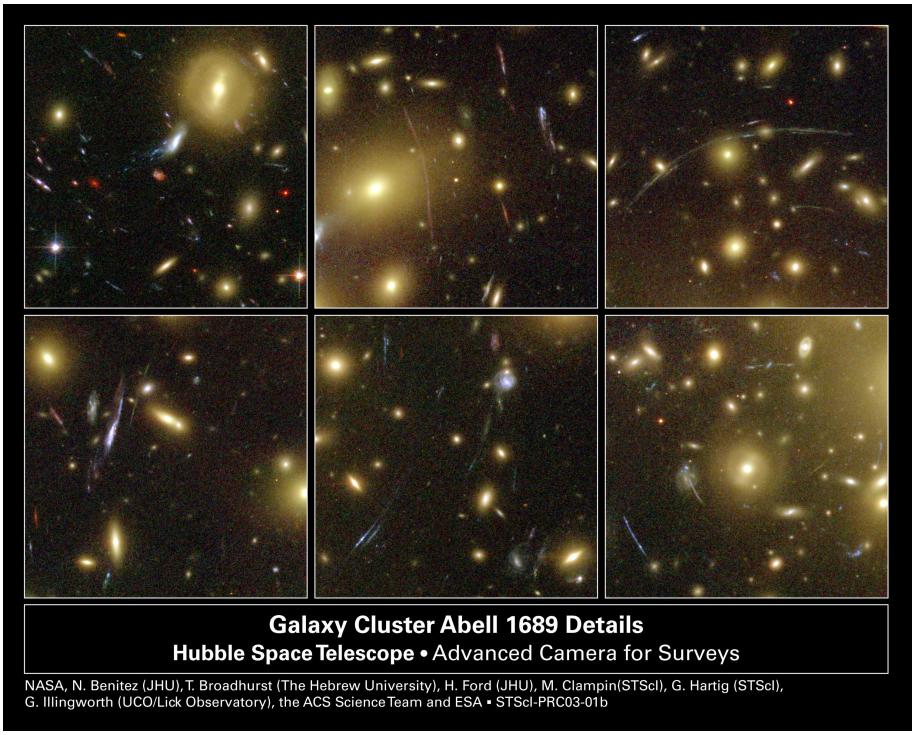

Figure 2.1. Images of gravitational lensing effects created by a massive galaxy cluster Abel 1698. Credit: NASA, N. Benitez (JHU), T. Broadhurst (The Hebrew University), H. Ford (JHU), M. Clampin(STScI), G. Hartig (STScI), G. Illingworth (UCO/Lick Observatory), the ACS Science Team and ESA.

microlensing. An example image containing both weakly and strongly lensed galaxies is shown in figure 2.1. It shows a galaxy cluster Abel 1698 which acts as a massive lens.

In this chapter we focus on weak gravitational lensing. Firstly, we describe its principles 2.1. Then we discuss the lensing observables in Section 2.3. Section 2.4 describes how to use these observables to study cosmology with two point correlation functions.

2.1 Lens equation

Let's consider a thin, spherically symmetric lens with mass density distribution $\rho(\xi, z)$, where ξ is the angle vector from centre of the lens and z is the line of sight distance. This lens will have a projected mass density, defined

$$\Sigma(\xi) = \int \rho(\xi, z) dz \tag{2.1}$$

The image of this background galaxy, located such that the impact factor is ξ, will be deflected by an angle $\tilde{\alpha}$. The deflection angle of a light ray can derived from GR by

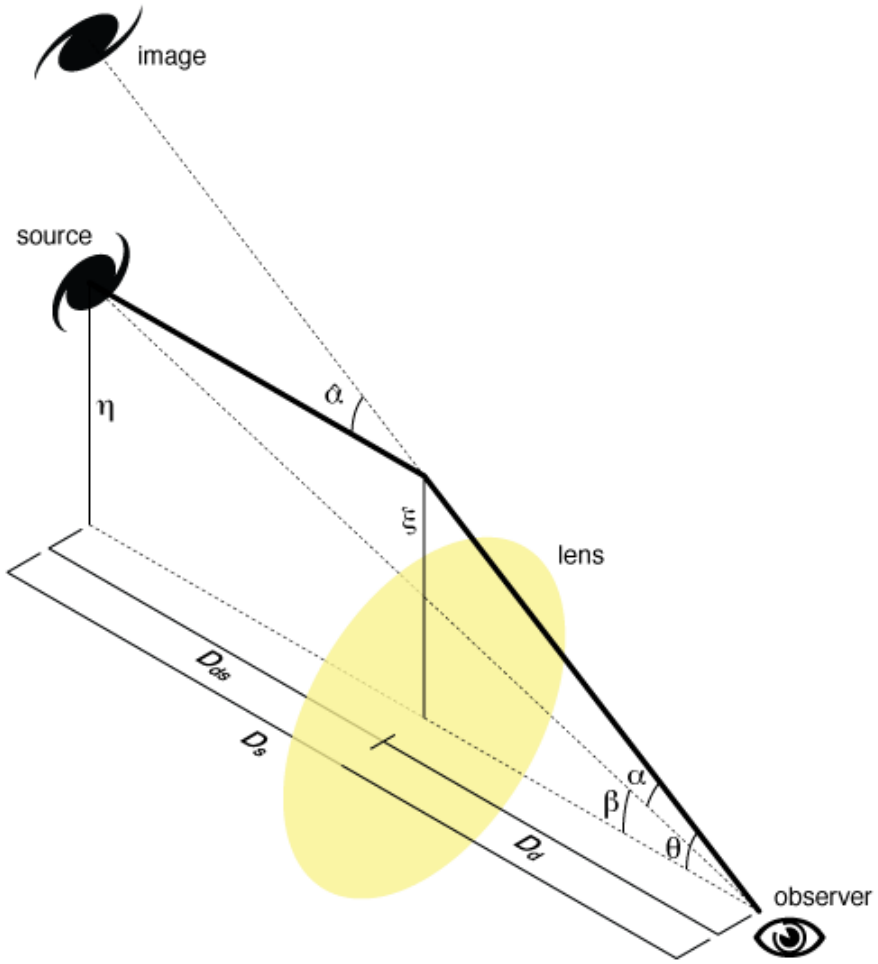

Figure 2.2. A schematic for bending of light due to gravitational lensing. Image credit: Michael Sachs (creative commons licence).

calculating a geodesic of a photon travelling in the proximity of large mass and be expressed as (see Bartelmann & Schneider 1999 for review)

$$\tilde{\alpha} = \frac{4\pi G}{c^2} \int \frac{\xi - \xi'}{|\xi - \xi'|^2} \Sigma(\xi') d^2\xi'^2 \tag{2.2}$$

where G is the Gravitational Constant. Fig. 2.2 shows a configuration for source, lens and observer positions. Given the distances from the observer to the lens D_d, observer to the source D_s and lens to source D_{ds}, it is possible to calculate the true, unlensed position angle for the source galaxy by analysing the geometry of the system presented in Fig 2.2.

$$\beta = \theta - \frac{D_{ds}}{D_s} \tilde{\alpha}(D_d \theta) \tag{2.3}$$

where $\theta = \xi/D_d$ is the angle of the observed background galaxy. Note that for fixed β there may be multiple θ which gives solutions to Eqn. 2.3. That means that the source can be multiply imaged, entering the regime of strong lensing.

2.2 Convergence and lensing potential

In this section I will discuss two very useful quantities in gravitational lensing: *convergence* and *deflection potential*. These are used later to derive lensing observables, such as shear and magnification. Let us define dimensionless convergence

$$\kappa = \frac{\Sigma(D_s\theta)}{\Sigma_{cr}} \tag{2.4}$$

where $\Sigma_{cr} = \frac{c^2 D_s}{4\pi G D_{ds} D_d}$ is the called *critical surface density*. Convergence can be used to determine between strong and weak lensing regimes. If $\kappa \ll 1$, then the galaxy is only weakly lensed and $\kappa \geq 1$ suggests a strong lensing effect. The deflection angle can be written in terms of convergence (by combining Eqns. 2.4 and 2.2)

$$\tilde{\alpha} = \frac{1}{\pi}\int_{\mathbb{R}^2} d^2\theta' \kappa(\theta') \frac{\theta - \theta'}{|\theta - \theta'|^2} \tag{2.5}$$

Deflection potential is defined as

$$\Psi(\theta) = \frac{1}{\pi}\int_{\mathbb{R}^2} d^2\theta' \kappa(\theta') \ln|\theta - \theta'|. \tag{2.6}$$

Then, the bending angle can be expressed as a gradient of the lensing potential: $\alpha = \nabla\Psi$.

2.3 Image distortions

How can we describe the way that the galaxy light is redistributed when it passed through a gravitational lens? Let us consider a galaxy image with intensity $\mathrm{I}^s[\beta]$ on the source plane, which is observed and has the intensity of $\mathrm{I}^o[\beta(\theta)]$ on the observers plane. We would like to find a mapping which will modify the light distribution from the source to the lens plane. If the size of the object is small compared to the impact factor ξ, this mapping can be locally linearised and represented as a Jacobi matrix $\mathcal{A}(\theta) = \frac{\partial\beta}{\partial\theta}$. Brightness of an image at the observer's plane and angle θ are related to source plane image in the following

way:

$$\mathrm{I}^o(\theta) = \mathrm{I}^s[\beta_0 + \mathcal{A}(\theta_0) \cdot (\theta - \theta_0)] \tag{2.7}$$

where $\beta_0 = \beta(\theta_0)$ and θ_0 is the point around which the transformation was linearised. This Jacobian is a symmetric and trace-free matrix, as it preserves surface brightness of the source galaxy. With these properties, it can be expressed as a single complex number, as there exists a one-to-one mapping between such matrices and complex numbers. Equation 2.8 shows how to relate this Jacobian to the deflection potential (Eqn. 2.6), and introduces complex *gravitational shear* γ.

$$\mathcal{A}(\theta) = \delta_{ij} - \frac{\partial^2 \Psi(\theta)}{\partial\theta_i \partial\theta_j} = \begin{bmatrix} 1 - \gamma_1 - \kappa & -\gamma_2 \\ -\gamma_2 & 1 + \gamma_1 - \kappa \end{bmatrix} \tag{2.8}$$

Another way to relate the shear to the deflection potential is

$$\gamma_1 = \frac{1}{2}(\Psi_{,11} - \Psi_{,22}) \quad , \quad \gamma_2 = \Psi_{,12}. \tag{2.9}$$

Shear can additionally be related to convergence as

$$\gamma = \frac{1}{\pi} \int_{\mathbb{R}^2} d^2\theta' \mathcal{D}(\theta - \theta')\kappa(\theta') \quad , \quad \mathcal{D} = \frac{-1}{(\theta_1 - \mathrm{i}\theta_2)^2} \tag{2.10}$$

which is a two-dimensional convolution with the shear response kernel $\mathcal{D}$.

As a complex number, shear can be expressed in terms of magnitude and angle

$$\gamma = \gamma_1 + \mathrm{i}\gamma_2 = |\gamma| \exp(2\mathrm{i}\phi). \tag{2.11}$$

Shear is a spin-2 vector, which transforms into itself after a rotation of 180°. Another property of shear γ is that its magnitude does not exceed unity: $|\gamma| < 1$.

Very commonly used quantities are the *reduced shear* $\mathbf{g}$, *conformal shear* η, and *dis-*

tortion δ, which are related to ellipse axis ratio $q = a/b$ via Eqn. 2.12

$$\begin{aligned} g &= \frac{\gamma}{1-\kappa} = \frac{a-b}{a+b} \\ e^{-\eta} &= a/b = q \\ \delta &= \frac{a^2-b^2}{a^2+b^2} = \tan(\eta). \end{aligned} \tag{2.12}$$

For more details about different shear parameterisations, see (Bernstein & Jarvis 2001). In this work I will mostly use the reduced shear parameterisation. If an isotropic luminosity distribution, which has a diagonal covariance matrix as measured by its quadrupole moments, is transformed by a shear $\mathbf{g}$, its covariance matrix $\mathbf{C}$ will become (Eqn. 2.13)

$$\mathbf{C}^{-1} = \begin{bmatrix} \frac{1+|\mathbf{g}|^2-2g_1}{1-|\mathbf{g}|^2} & \frac{-2g_2}{1-|\mathbf{g}|^2} \\ \frac{-2g_2}{1-|\mathbf{g}|^2} & \frac{1+|\mathbf{g}|^2+2g_1}{1-|\mathbf{g}|^2} \end{bmatrix} \tag{2.13}$$

Note that this covariance does not include the change in the apparent size of the galaxy induced by lensing. Gravitational lensing also induces *magnification*, which is a change in apparent galaxy size (Gaztanaga 2003; Schmidt et al. 2011). Magnification can be expressed in terms of the determinant of the Jacobian matrix (Eqn. 2.14)

$$\mu = \frac{1}{\det\mathcal{A}} = \frac{1}{(1-\kappa)^2 - |\gamma|^2} \tag{2.14}$$

Magnification and shear are the two most commonly used lensing observables. Including higher order expansion in the transformation of lensing images gives rise to *flexion* (Bacon et al. 2006; Rowe et al. 2013; Velander et al. 2011).

2.4 Cosmology with weak-gravitational lensing

Gravitational lensing is a unique probe of cosmology. Its uniqueness stems from the fact that image distortions induced by matter are an unbiased tracer of mass of this matter. Another feature of weak lensing is the fact that matter distribution in the large scale structure can be studied as a function of redshift, thus probing its evolution in time. Statistics of distribution of mass can be used to constrain cosmological parameters and test the laws of gravitation on large scales (e.g. Kaiser 1996).

The observed lensing shape is an effect of distortions induced by all intervening matter

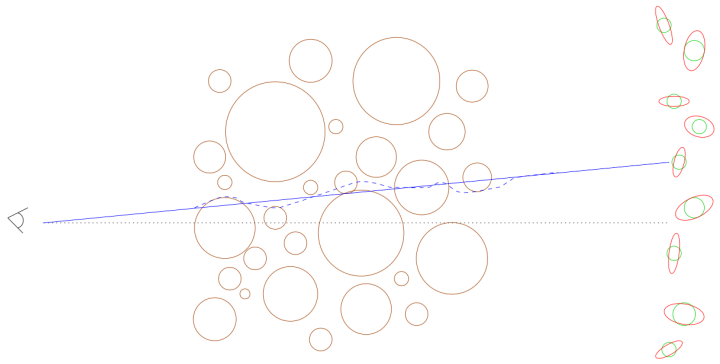

Figure 2.3. Light from a distant galaxy propagating through the large scale structure of the Universe. The most contribution to the lensing signal comes from structures half way between a source and a observer. Image from Refregier (2003).

along the line of sight. As the lensing strength depends on the redshift of the source galaxy and the lens, distortions are integrated in the radial direction with a specific kernel, called the *lensing kernel*. Lensing potential thus becomes

$$\Psi_{ij} = \int_0^{\chi_h} d\chi \partial_i \partial_j \Phi g(\chi) \tag{2.15}$$

where Φ is the Newtonian potential, χ is the comoving distance, χ_h is the distance to horizon, derivatives are performed on the sky plane, and the lensing kernel is

$$g(\chi) = 2 \int_{\chi}^{\chi_h} d\chi' n(\chi') \frac{r(\chi) r(\chi' - \chi)}{r(\chi')} \tag{2.16}$$

where $r = D_A/a$, and D_A is the angular diameter distance (Refregier 2003). The number density of galaxies $n(\chi)$ is assumed to be normalised and depends on the depth of a survey. The lensing kernel has a property that it peaks around half-way between an observer and a source. Fig. 2.3 shows a schematic of the path of a light ray from a distant galaxy to an observer through the large scale structure. In addition to shape distortion, LSS also changes the position angle at which the galaxy is observed. The integral in Eqn. 2.15 should be taken along the perturbed photon path, but the deflection is typically small, so to first order it can be integrated along a straight line. This is known as the *Born approximation*.

2.4.1 Shear two point statistics

A commonly used statistic of the shear distribution is a two-point angular correlation function, expressed as

$$\xi_{\pm}(\theta) = \langle \gamma_{+}\gamma_{+} \rangle \pm \langle \gamma_{\times}\gamma_{\times} \rangle \tag{2.17}$$

where + and × are tangential and cross component of shear defined with respect to the angle joining the two galaxies. The correlation function is related to the shear power spectrum via the Hankel transform

$$\langle \xi_{+}\xi_{+} \rangle + \langle \xi_{\times}\xi_{\times} \rangle = \int \frac{\ell d\ell}{2\pi} (C_{\ell}^{EE} + C_{\ell}^{BB}) J_0 \ell\theta \tag{2.18}$$

where C^{EE} is the E-mode and C^{BB} is the B-mode of the power spectrum. Gravitational lensing does not produce any B-mode shear, as distortions produced by clumped mass can only produce E-mode patterns. Measurements of the B-mode can be a useful diagnostic for systematics in a survey. However, the B-mode can be observed due to survey area covering only a fraction of the sky, which is the case for most surveys.

The two-dimensional, angular shear power spectrum can be related to the three-dimensional matter power spectrum, which is predicted by various cosmologies. In fact, the 2D power spectrum can be calculated as an integral of the 3D power spectrum with a weight function. This calculation is simplified greatly using the Limber approximation, which considers the fact that, for small angles, only large scale modes along the line of sight contribute to the integral. The final power spectrum is (Refregier 2003)

$$C_{\ell} = \frac{9}{16} \left(\frac{H_0}{c} \right)^4 \Omega_m^2 \int_0^{\chi_h} d\chi \left[\frac{g(\chi)}{ar(\chi)} \right]^2 P(\frac{\ell}{r}, \chi) \tag{2.19}$$

where H_0 is the Hubble parameter today, Ω_m is the matter density today, a is the scale factor. Note that the observable used for this measurement is the *reduced* shear g. The impact of using reduced shear instead of cosmic shear was quantified by Dodelson et al. (2005). Cosmic shear was detected first by four groups 15 years ago (Wittman et al. 2000; Kaiser et al. 2000; Van Waerbeke et al. 2000; Bacon et al. 2000*b*), and since then has been a growing field of research.

Cosmic shear can also be estimated at different redshifts. Using photometric estimators

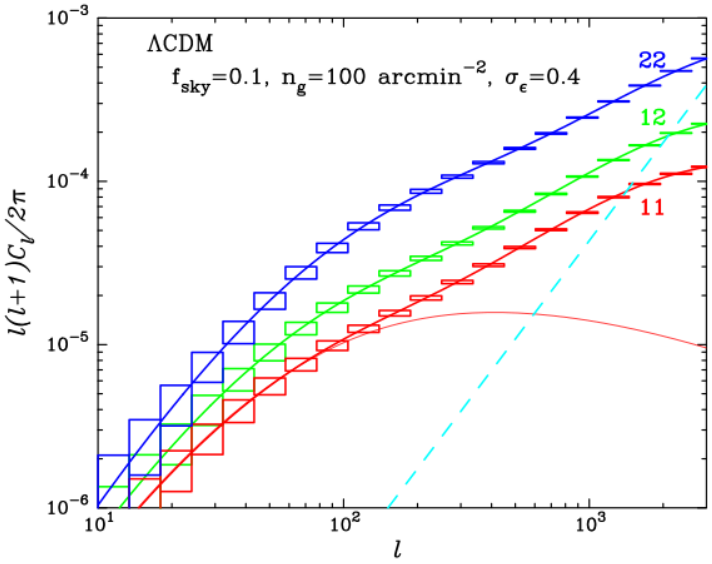

Figure 2.4. Lensing power spectra calculated by Takada & Jain (2005). Redshift bins were $z \in [0, 1.3]$ and $z \in [1., 2.2]$. Spectrum 12 is the cross-component between bins. Boxes show the measurement error from sample variance and shape noise. Linear contribution to bin 11 is shown in a solid line to give an idea about the size of non-linear effects. Shot noise contribution to the measurement is shown with the dashed curve.

of redshifts, source galaxies can be put into bins in z. This measurement is called *cosmic shear tomography* (Takada & Jain 2005). Fig. 2.4 shows the example power spectra for two redshift bins. In particular, this probe is useful for constraining the evolution of the dark energy equation of state (see Albrecht et al. 2006).

2.4.2 Requirements on biases on shear measurements

Systematic errors on measurements of shear will cause the estimated cosmological parameters to be biased. It is important to calculate how exactly these systematic errors translate to the bias on cosmological parameters. In this section I summarise the findings of Amara & Refregier (2007), who give formulas for required shear systematics as a function of the size and depth of a survey.

Forecasting the uncertainty on measured cosmological parameters for a survey is often done using a Fisher matrix formalism. An observed power spectrum of shapes of galaxies ($C_\ell^{\rm obs}$) is a sum of true lensing signal ($C_\ell^{\rm lens}$), noise contribution ($C_\ell^{\rm noise}$), and systematics ($C_\ell^{\rm sys}$): $C_\ell^{\rm obs} = C_\ell^{\rm lens} + C_\ell^{\rm noise} + C_\ell^{\rm sys}$. An estimator $\widehat{C}_\ell^{\rm lens} = C_\ell^{\rm obs} - C_\ell^{\rm noise}$ is assumed to be biased, as the systematic term is not known, or corrected to a known level of uncertainty.

This estimator then has an error of

$$\Delta C_\ell = \sqrt{\frac{1}{(2l+1)f_{sky}}} \tag{2.20}$$

where f_{sky} is the sky fraction covered by a survey. Then, the maximum likelihood of cosmological parameter $\hat{p}_i$ is found, following the ususal minimization of $\chi^2 = \sum_l \Delta C_\ell^{-2}[\widehat{C_\ell^{\rm lens}} - C_\ell^{\rm lens}(p_i)]^2$. An estimate of uncertainty on the cosmological parameter is then calculated using an inverse of the Fisher matrix

$$F_{ij} = \sum_{ij} \Delta C_\ell^{-2} \frac{dC_\ell^{\rm lens}}{dp_i} \frac{dC_\ell^{\rm lens}}{dp_j} \tag{2.21}$$

If we include the residual systematic on C_ℓ in the analysis and assume it to be small, we can calculate bias on cosmological parameters with respect to the true parameter $p_i^{\rm true}$ as in Eqn. 2.22, assuming summation convention

$$b[\hat{p}_i] = \langle \hat{p}_i \rangle - \langle p_i^{\rm true} \rangle = (F_{ij})^{-1} B_{ij} \tag{2.22}$$

where $B_j = \sum_\ell \Delta C_\ell^{-2} C_\ell^{\rm sys} \frac{dC_\ell^{\rm lens}}{dp_j}$ is the bias vector. A cosmological survey will be able to achieve its potential if systematic errors are smaller than statistical errors, and criterion 2.23 is satisfied

$$b[\hat{p}_i] \leq \sigma[\hat{p}_i] \tag{2.23}$$

where $\sigma^2[\hat{p}_i] = (F_{ii})^{-1}$ is the statistical error on a parameter. If this condition is not met, it is not useful for a survey to take more data.

The bias on shear is usually decomposed into multiplicative and additive bias, following Heymans et al. (2006): $g^{observed} = mg^{true} + c$. This decomposition is particularly useful as the PSF shape introduces mainly the additive systematic.

Amara & Refregier (2007) used the criterion 2.23 to put limits on multiplicative and additive shear biases, assuming equal systematic and statistical errors. They considered investigating a wide range of bias models. These models included additive and multiplicative biases with and without redshift evolution, as well as different dependencies on scale: (i) log-linear, (ii) with the same shape as the lensing signal and (iii) when a single parameter can mimic a small change in cosmological parameter p_α: $C_\ell^{\rm sys} = A_2 \frac{C_\ell^{ij}}{dp_\alpha}$, where

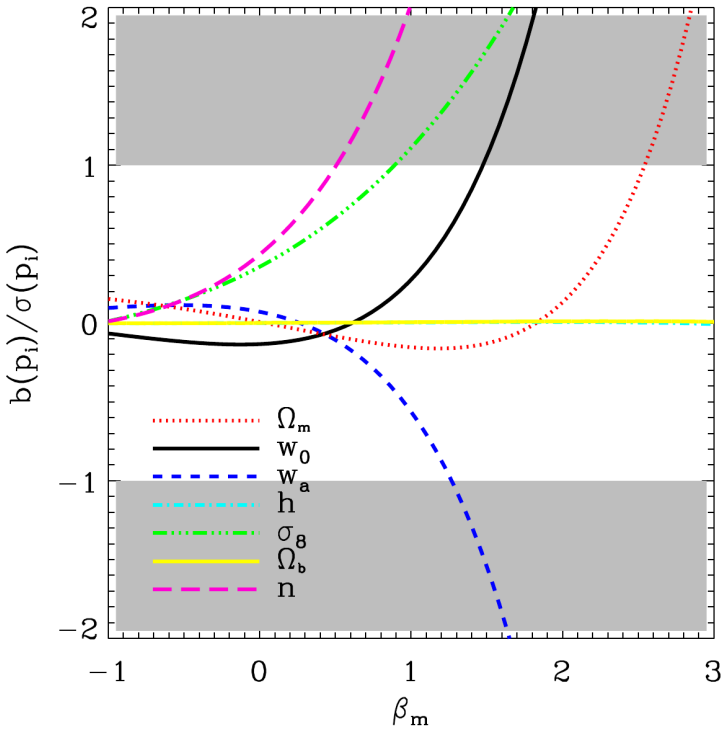

Figure 2.5. The ratio of systematic to statistical errors for the multiplicative bias for various cosmological parameters. The evolution with redshift followed a power-law 2.24, with $m_0 = 5 \cdot 10^{-3}$. This figure illustrated that for this parameter configuration the redshift dependence has to be small in order for the systematics to be sub-dominant. Figure from Amara & Refregier (2007).

A_2 is the normalisation factor.

Here I consider only the multiplicative bias model with redshift evolution. This evolution was modelled as a power law

$$m = m_0(1 + z_m)^{\beta_m}. \tag{2.24}$$

In figure 2.5 the ratio of systematic to statistical error is shown for multiplicative bias evolving with redshift. Systematic errors are sub-dominant for $\beta_m < 1.5$, and that is the set throughout the rest of their analysis.

This kind of analysis is repeated for a range of bias models. When reporting the final requirements on the bias, the approach was very conservative: the most biased out of all the cosmological parameters was driving the constraints on the shear systematic level. Moreover, the bias was allowed to change sign as it evolved in redshift. The most important conclusions were that both the additive and multiplicative signals have to have

weak dependence on redshift. The tolerance on additive and multiplicative for a custom survey with area A_s, median redshift z_m and number density of galaxies n_g is (Eqn. 2.26)

$$\sigma^2_{\rm sys} < 10^{-7} \left(\frac{A_s}{2 \cdot 10^4 \text{ deg}^2}\right)^{-0.5} \left(\frac{n_g}{35 \text{ arcmin}^{-2}}\right)^{-0.5} \left(\frac{z_m}{0.9 \text{ deg}^2}\right)^{-0.6} \tag{2.25}$$

$$m_0 < 10^{-3} \left(\frac{A_s}{2 \cdot 10^4 \text{ deg}^2}\right)^{-0.5} \left(\frac{n_g}{35 \text{ arcmin}^{-2}}\right)^{-0.5} \left(\frac{z_m}{0.9 \text{ deg}^2}\right)^{-0.6} \tag{2.26}$$

In section 3.2, in Table 4.1 I show the requirements calculated for typical survey specifications, for current, ongoing and far future surveys. Good progress has been made in the development of shear measurement methods over the last decade, and the tools currently used do satisfy these requirements (see Sec. 2.7 and Fig. 2.7). However, reaching the systematic floor of $m < 10^{-3}$ will be increasingly difficult and much more effort is needed in development of shear measurement algorithms and calibration methodologies.

2.5 Shear measurement and its systematics

In this section I describe the practical aspects of measurement of shear. I start with outlining the lensing galaxy image generation process. Next, I describe popular approaches to shear measurement and sources of systematic biases inherent to these methods. Finally, I highlight the importance of shear accuracy testing programmes, organised by the lensing community.

Galaxies in the universe are intrinsically elliptical, even in the absence of gravitational lensing. Therefore, it is not possible to measure shear from individual galaxy images; shear has to be measured statistically. The key property which allows us to make such a measurement is that, in the absence of intrinsic alignments, rotation angles of galaxies are random and unlensed galaxy ellipticities average to zero: $\langle \mathbf{e}^i \rangle = 0$. If a galaxy with perfectly elliptical isophotes and shape $\mathbf{e}^i$ is distorted by a shear $\mathbf{g}$, then the resulting ellipticity of the observed galaxy $\mathbf{e}^l$ will be (Eqn. 2.27)

$$\mathbf{e}^l = \frac{\mathbf{e}^i + \mathbf{g}}{1 + \mathbf{g}^* \mathbf{e}} \tag{2.27}$$

Equation 2.27 can be used to show that the lensed galaxy ellipticity is an unbiased shear estimator (Schneider & Seitz 1994; Seitz & Schneider 1996), to third order in shear, so that $\langle \mathbf{e}^l \rangle = \mathbf{g}$.

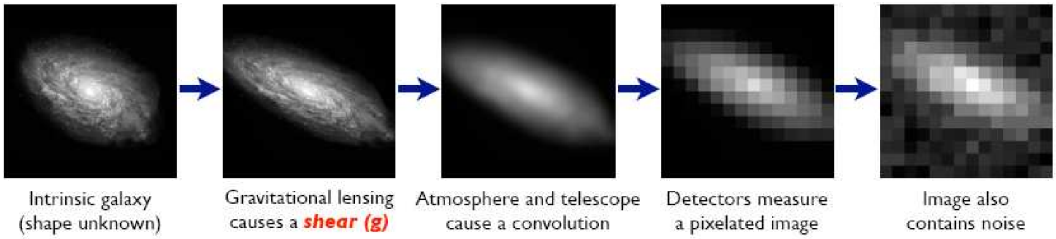

Figure 2.6. Forward process for creation of galaxy images observed by a telescope. Image credit: Bridle et al. (2009*b*)

.

In reality, however, galaxies do not have perfectly elliptical isophotes, so it is not clear what the shape of such a galaxy would be. In principle, a covariance matrix of all photons coming from the galaxy could be used. This approach is not practical, as galaxy light can extend far away from its centre and thus prevent its precise calculation in the presence of noise on the images. That is why practical shape measurement often uses ellipticity estimators, which transform as Eqn. 2.27. Multiple shear measurement methods have been proposed over the last two decades. Before I introduce them, I will first discuss other effects which affect observed galaxy images.

2.5.1 Forward process for observed galaxy images

The image of a galaxy is subject to several degrading effects before it is recorded. Figure 2.6 shows the forward process for observed galaxy images. Firstly, distortion comprises the lensing effects, including shear, magnification and higher order flexions. When the photons travel through the earth's atmosphere (in the case of ground-based observations), and telescope optics, the image is being convolved by a Point Spread Function (PSF). Then the light is integrated within pixels of the detector, and this induces further convolution with a pixel response kernel, followed by downsampling of the image. Additionally, the detectors register noise from photons in the atmosphere and inside the detector. For ground-based observations, the noise is sky-limited, and it can be well described as additive Gaussian noise. For space-based imaging, the noise is often modelled as Poisson, as the number of background photons is low. Detector noise is often assumed to be Gaussian.

A shear measurement method has to invert this process, and be able to statistically recover shear $\mathbf{g}$, including accurate treatment of PSF, pixelisation and noise effects. It is very important to account for the PSF convolution accurately, as it affects the galaxy image in a similar way that the shear can (Paulin-Henriksson et al. 2008). PSF size can

often be of similar size as the galaxy, and in principle is limited by the telescope optical system. Given the telescope aperture D and light wavelength λ, the size of the PSF scales as λ/D. Telescope optics models provide estimates of PSF properties, but in general, the PSF at the position of the galaxy is not precisely known. Under the assumption that the PSF varies slowly across the telescope field of view, one can use neighbouring stars to estimate the PSF on the position of the galaxy (e.g. Gentile et al. 2012*b*). Even when the PSF at the position of the galaxy is known very well, a shape measurement method has to account for it accurately. Many approaches to this deconvolution problem have been proposed, and I will discuss them in section 2.6.

Noise related effects can also introduce a significant bias on lensing measurements, especially if high magnitude galaxies are used. In (Refregier et al. 2012; Kacprzak et al. 2012) I studied the theoretical foundations of *noise bias*, its magnitude and properties.

Finally, it is common that a survey will take several images of the same galaxy at different times (and with different PSF). These exposures can be dithered with respect to the centre of the imaged galaxy, and taken at different angles, which can introduce distortions related to the World Coordinate System (WCS). A shape measurement method has to be able to perform well on multiply-imaged galaxies, either by using co-added pixels, or fitting the images jointly. Other effects that can affect the quality of shape measurement are discussed in Sec. 2.5.2.

2.5.2 Considerations for measurement of shapes

In Section 2.5.1, I discussed the effects that play a significant role in the creation of the sky image. However, this list is not exhaustive, and there exists a range of other effects and considerations which are important for shear estimation. In this section I will highlight some of the most important ones, but again, this list will not be exhaustive.

Complex galaxy morphologies

Morphologies of galaxies can be complicated; they can have bulges, bars, spiral arms, bright star-forming clusters and dust lanes. Despite the fact that most of the small scale variation in the brightness distribution is smeared out by the PSF, galaxy morphology can influence shear measurement. Some methods (see Sec. 2.6) use parametric galaxy models to fit the image pixels, and if the models are underfitting (not able to capture all possible galaxy morphology aspects after blurring), this can cause a bias (Bernstein 2010; Voigt &

Bridle 2009; Kacprzak et al. 2013). This bias is called *model bias*. I measured the scale of model bias for the first time using high resolution galaxy images from the Hubble Space Telescope, degraded in quality to resemble ground based observations. I found it to be sub-percent level, see Sec. 4.4, which can be important for future surveys.

Distribution of galaxy parameters, as seen by a survey

Many methods use empirical calibration factors derived from simulations. At the precision level required by ongoing and future surveys, the resulting calibration factors are strongly dependent on input to the simulation. If shear is to be calibrated correctly, the simulation input must represent the real survey data very well. Distributions of true parameters of galaxies in the survey, such as magnitudes, sizes, intrinsic ellipticities, Sérsic indices (see Sec. 2.6), etc., have to be known. Obtaining these is not an easy task, as the only parameters measured directly from survey images are noisy. An approach has been proposed by Refregier & Amara (2014) to infer these distributions, using a control loop approach. In this method, the simulations are created repetitively until the distribution of *observed* parameters match. In the future, the importance of shear calibrations using large simulations will increase, as higher precision of the measurement will be required.

Blending

The fraction of galaxies which are blended with a neighbour increases greatly with limiting magnitude. In current surveys, such as the Dark Energy Survey (DES), reaching the magnitude of 24, the fraction of blends can be as high as 10%. For the Large Synoptic Sky Survey (LSST), this fraction will increase even more, posing a fundamental limit on the statistical power of the survey if blends are to be rejected (Chang et al. 2013). Most of the current shape measurement methods are not designed to deal with blends, and this will become a major challenge in the future.

2.6 Shape measurement methods

In this section I will describe a range of shear measurement methods, highlighting their differences in the PSF deconvolution process, optimization, and general performance.

2.6.1 Quadrupole moment methods

The first methods used for lensing measurements were based on weighted quadrupole moments (e.g Kaiser et al. 1994; Luppino & Kaiser 1996; Hirata & Seljak 2003; Okura & Futamase 2010). These moments measure the covariance of the weighted galaxy luminosity around a central point and then correct for the PSF. The quadrupole moment is defined as

$$q_{xy} = \frac{\sum(x_i - \bar{x}_i)(x_j - \bar{x}_j)W(\mathbf{x} - \bar{\mathbf{x}})\mathrm{I}(\mathbf{x})}{\sum w(\mathbf{x} - \bar{\mathbf{x}})\mathrm{I}(\mathbf{x})} \tag{2.28}$$

where $i, j \in 1, 2$ are image coordinates and $\bar{\mathbf{x}}$ is the centre of the moment. The weighting function is often selected to maximise the signal to noise ratio (SNR) of the measurement and to decrease the contribution of the noise from pixels which are far from the edge of the postage stamp. The PSF correction is done after the moments are measured, analytically, using the moments measured from the PSF, under the assumption that both PSF and the galaxy have small ellipticities. Then, the ellipticity can be extracted from quadrupole moments (Eqn. 2.29)

$$\mathbf{e}^l = \frac{q_{11} - q_{22} + 2iq_{12}}{q_{11} + q_{22} + 2\sqrt{q_{11}q_{22} - q_1^2}}. \tag{2.29}$$

Erben et al. (1999); Bacon et al. (2000a) used numerical simulations to test various KSB weighting schemes and they concluded that shear values can be recovered with systematic uncertainty of around 10%. Melchior et al. (2012) presented a method in which the deconvolution of the moments can be done exactly in the absence of noise, weighting function and perfect centroiding, and then studied the impact of those effects on deconvolution in moment space. Moment-based methods often used a responsivity correction, an empirical calibration factor applied to the shear estimator. Advantages of quadrupole moments methods are that in principle they do not assume any light profile for a galaxy, and thus are free of model bias. In particular, Bernstein (2010), demonstrated Fourier Domain Null Testing (FDNT) method, which calculates unbiased estimates of shape even for galaxies which have very complicated morphologies and strong ellipticity gradients.

2.6.2 Shapelets

Methods based on shapelets (Refregier 2001; Massey & Refregier 2006; Nakajima & Bernstein 2006) use a decomposition of an image onto an orthogonal basis set. In the shapelet methods family, the basis often consists of Gauss-Hermite or Gauss-Laguerre polynomials. These are essentially Gaussian profiles multiplied by a polynomial function. This basis set is complete and able to reconstruct all images given a basis set which is big enough. The same decomposition is done to the PSF, and then the convolution can be carried out analytically using just the shapelets coefficients. Galaxy shapes can then be calculated analytically from the basis set coefficients. Point estimators are used for fitting, and solving for the best fit coefficients requires a matrix inverse. However, in practice it is not tractable to use a full shapelet coefficients expansion; a truncated basis set has to be used for both practical implementation reasons and to prevent overfitting the noise. This creates a fundamental limitation; shapelets in general resemble Gaussian intensity profiles, whereas galaxies usually follow Sérsic profiles (Sérsic 1963), which have much higher kurtosis. This makes shapelets a basis set not well suited for galaxy profiles when a finite set of basis function is used.

2.6.3 Model fitting methods

Another popular approach involves model-fitting methods. In these, a parametric model is used for galaxy and PSF images. Most often used models are generalised Gaussian profiles, in astronomy known as Sérsic profiles, which follow this function

$$\mathbf{I}(\mathbf{x}) = A \exp\left(-k[(\mathbf{x}-\mathbf{x}_0)^\intercal \mathbf{C}^{-1}(\mathbf{x}-\mathbf{x}_0)]^{0.5n}\right) \tag{2.30}$$

where A is the amplitude, $k = 1.9992n - 0.3271$, $\mathbf{x}_0$ is the centre of the profile, $\mathbf{C}$ is the covariance matrix and n is the Sérsic index of the profile, controlling its kurtosis. For an isotropic profile, diagonal covariance matrix elements will define a half-light radius of the profile, r_e, which encloses half of the total flux. Two most commonly used Sérsic indices are $n = 1$ (Exponential profile) and $n = 4$ (DeVaucouleurs profile), which are commonly used models for galaxy discs and bulges, respectively. Model fitting methods, such as (Kuijken 1999; Bridle et al. 2002; Zuntz et al. 2013; Miller et al. 2007; Gentile et al. 2012*a*) use these profiles to model galaxy light profiles. PSF modelling is often done using other parametric models. Then, convolution is done either numerically on a fine grid or

analytically after decomposing the fitted galaxy and PSF profile into a sum of Gaussians. A likelihood function is used, such as in Equation 3.7 and the final shape estimator is then derived from this likelihood function. Often used estimators include Maximum Likelihood (ML) or Mean Posterior (also known as Bayesian Estimator). For finding this estimator, various numerical optimization algorithms are used, such as numerical steepest descent (for example, LEVMAR or MINUIT), adaptive grid search, or implementations of the Markov Chain Monte Carlo methods (MCMC). Model fitting, like other methods, suffers from noise bias effects, as the resulting estimator is a non-linear function of pixel intensities affected by Gaussian noise. It is also not free from model bias, as models are simplistic and not able to characterise the complex morphologies of galaxies. Finally, the optimization process can introduce systematic effects; for example, the starting point and convergence of steepest descent algorithms can influence the shear estimator.

2.6.4 Bayesian methods

Bayesian methods are an alternative approach. Rather than aiming to produce an unbiased estimator of shear, they combine full probability distributions of shear from many galaxies. For example, Bernstein & Armstrong (2013) presented a way of combining the likelihoods of galaxies which avoids noise bias. An implementation of this method by Sheldon (2014) also shows promising results. As these methods do not produce e_1 and e_2 estimators for each galaxy but store information about the likelihood of a galaxy in some form, other steps of the weak lensing analysis have to be adjusted to accommodate for this change. Up until today, these methods are still in early stages of development.

2.6.5 Stacking methods

Another approach which does not use a shear estimator for each galaxy is based on stacking the images in pixel space. In such a stack, in the absence of applied shear, the individual galaxy intrinsic ellipticity will average out and create an image with elliptical isophotes. When the shear is constant, the ellipticity of the resulting stack should be the same as applied shear. Lewis (2009) and Hosseini & Bethge (2009) demonstrated the performance of stacking methods on GREAT08, with the latter winning the main challenge. Note that good performance was achieved in the presence of constant PSF on all images, which is not the case for real galaxy surveys. These methods are in principle immune to model bias, as according to the central limit theorem individual morphologies of galaxies should

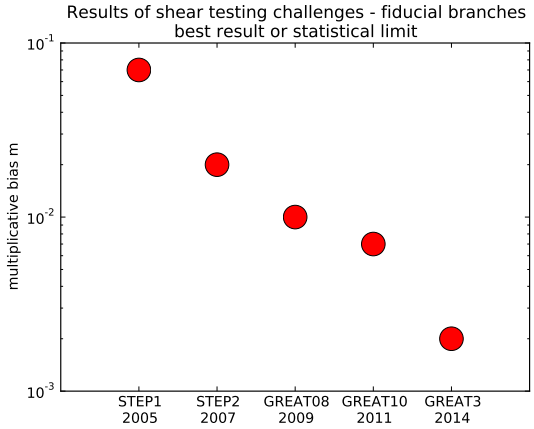

Figure 2.7. Multiplicative bias reported by shear challenges. For challenges which included branches, I list the result from fiducial branch. The concrete value represents the bias for the winning method, or the simulation's statistical limit, if it was achieved. Note that the details of the simulations varied greatly, including the parameters of the fiducial branch. Additionally, the way of reporting of results was not consistent, which is why this figure should be treated more as a guideline than actual result. Results of the GREAT3 are preliminary.

converge to a Gaussian when stacked. Such a stack should also be able to avoid noise bias, as the noise is suppressed in a linear way. However, a centroid has to be estimated very accurately for this process to work; otherwise shear might be underestimated. As for Bayesian methods, stacking methods do not use estimators, and require following stages of processing to be adjusted. In particular, it was not demonstrated how a stacking method could estimate the correlation function of galaxy shapes.

2.7 Shear testing programmes

To test and compare the performance of these methods, the weak lensing community organised a series of shear testing challenges, based on simulated data. Shear Testing Programmes 1 (STEP1) (Heymans et al. 2006) and 2 (STEP2) (Massey et al. 2006) were organised internally inside the community and provided simulated data sets with a high degree of realism, including unknown PSF, complicated galaxy morphologies, blended objects and image artefacts. These projects were followed by Gravitational Lensing Accuracy Testing Challenges: GREAT08 (Bridle et al. 2009*b*,*a*), GREAT10 (Kitching et al. 2013, 2011) and GREAT3 (Mandelbaum et al. 2014), which were opened to the Astronomy, Computer Science and Statistics communities. The approach to GREAT challenges was different than to STEP. GREAT started with a simplified version of the problem and

targeted specific questions in shear measurement, such as impact of the low SNR galaxies (GREAT08), variable shear (GREAT10) and influence of realistic galaxy morphologies (GREAT3), among others. Addressing specific questions helped to isolate the impact of different effects on shear measurement and design methods to address them.

The results of these challenges indicate a progressive improvement in the accuracy of the methods 2.7. We can notice a steady improvement in the accuracy of shear measurement methods over the last 10 years. Also, the sheer number of methods and teams entering the challenge increased greatly, reaching 23 the in recent GREAT3 challenge, which indicates increasing interest in both lensing and statistical communities.

2.8 Weak lensing with the Dark Energy Survey

Dark Energy Survey (DES) is an international project, with six participating countries and 27 institutes, which will use observations of a large fraction of the sky for a range of science goals. The main driver of DES is to learn about the properties of dark energy by observations of complementary probes such as galaxy clusters, weak lensing, baryon acoustic oscillations and supernovae. There are many other branches of science active within the DES, such as, for example, Milky Way science. The DES will additionally provide a great legacy data set.

The DES is planning to observe 5000 $\deg^2$ of southern sky to a great depth, with limiting magnitude often larger than 24. The survey also has a deep component, which will observe 6 $\deg^2$ to magnitude greater than 25. Figure 2.8 shows the survey strategy planned as for mid 2014, when the survey is entering its third year of observations. In particular, the survey overlaps fully with the area observed by the South Pole Telescope (SPT), which detected clusters using the Sunyaev-Zeldovich effect, and the Vista Hemisphere Survey (VHS). Overlap with other astronomical projects will allow us to learn more about objects observed by the DES.

DES is observing in four filters: *griz*. The detectors using charge coupled device (CCD) technology, are designed to be particularly sensitive to redshifted light, allowing probing high-redshift galaxies. Photometry in four bands is used to estimate the redshift of a galaxy, called *photometric redshift*. It is calculated by comparing the magnitudes of observed galaxies with spectra of galaxies observed using other methods. This training set, or template set, will be partially obtained using spectroscopic follow-ups to DES targets.

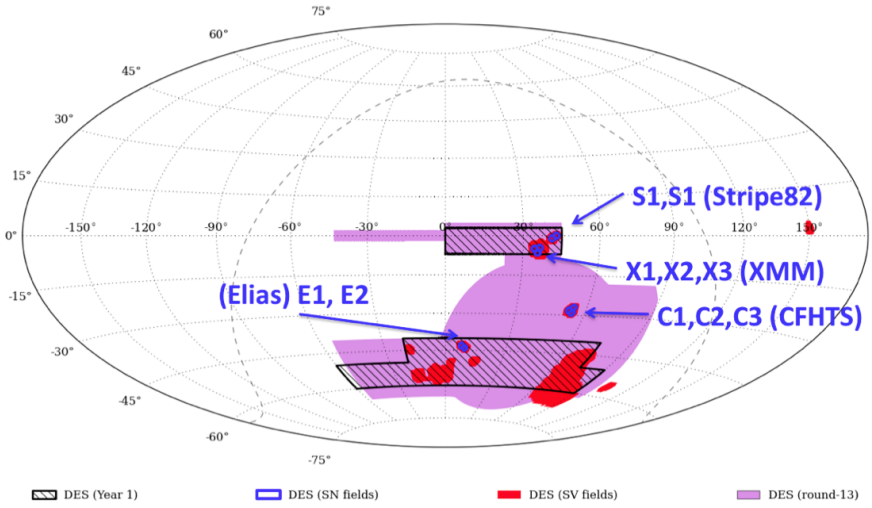

Figure 2.8. Dark Energy Survey observing area, as planned mid 2014. Area within the black shaded region was observed during first year of the survey. Area marked with red corresponds to observing season 2012-2013, during which verification of the instrument took place.

Combined with measurements of galaxy shapes, DES will provide a very powerful data set for dark energy science.

The home of the DES is the Cello Tololo Interamerican Observatory in northern Chile. A four-meter Blanco Telescope is used by the DES for a period from 2012 to 2017, observing almost exclusively for the DES in the southern summer time. The Dark Energy Camera (DECam) is a prime focus camera on the top of the Blanco, a unique wide-field camera with 62 science CCDs with 520 megapixels and images 3 square degrees with 0.27 arcsecond/pixel resolution.

First light of DECam was on 12 September 2012 and started a season of Science Verification, during which a number of technical challenges have been solved to improve the performance of the telescope. During this period observations were taken to full survey depth providing scientists with an almost-150-deg^2 deep dataset, which has already been used for a number of science publications. First observing season of the DES main survey programme started in autumn 2013. I had the great privilege to visit CTIO and perform observations for the DES from 24 to 30 January 2014.

One of the many science goals of the DES is the measurement of the dark energy equation of state. Due to its wide field nature, the DES will provide very accurate measurement of w_0. In the wCDM model, the dark energy equation of state parameter evolves with the scale factor a as $w(a) = w_0 + w_a(1-a)$. Figure 2.9 shows the Fisher-matrix-based forecast

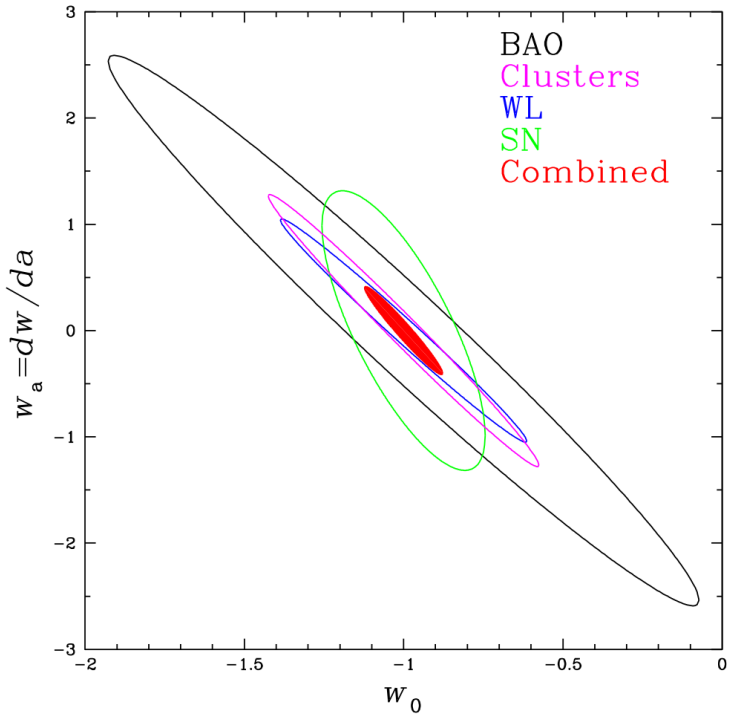

Figure 2.9. Fisher-matrix-based-forecast of the Dark Energy Survey constraints on the parameters of the dark energy equation of state. Source: Dark Energy Survey Science Program book.

of constraint on the $w_0 - w_a$ plane from The Dark Energy Survey Science Program book. Combined probes will be able to constrain the dark energy equation of state to a few percent precision. This is a very powerful test of the ΛCDM paradigm; measurement of departure from a value of $w = -1$ will indicate that Einstein's model with a cosmological constant as a driver of cosmic expansion will have to be modified.

Weak gravitational lensing is a very important, complimentary probe in this measurement. Cosmic shear is the most promising weak lensing probe, but measurements such as galaxy-galaxy lensing and three-point statistics are also being worked on. Next to cosmic shear, magnification is another observable with great potential for cosmological observations.

To achieve the full potential of the cosmic shear probe, the shear measurement pipeline has to deliver efficient and bias-free measurement of galaxy shapes. For the DES, the requirements on multiplicative bias on the shear is $m < 0.004$ and on additive bias $c < 0.0006$ (see Sec. 2.4.2). The DES shear measurement pipeline is being developed at the University of Manchester and the University of Pennsylvania, with contributions from many others. I

[0] `http://www.darkenergysurvey.org/reports/proposal-standalone.pdf`

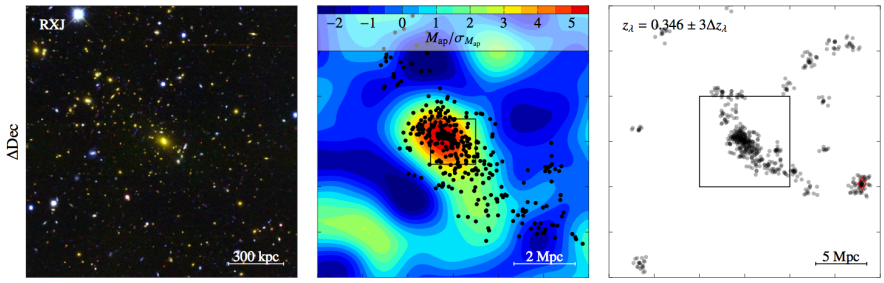

Figure 2.10. Mass reconstruction of cluster RXC J2248.7-4431 (Melchior et al. 2014). Left column is a multi-colour image the field, centred on the cluster. Middle panel is a weak-lensing aperture mass significance map, overlaid with galaxies (black dots). Cluster member galaxies were found using redMaPPer cluster finder. Right panel: same redMaPPer galaxies for larger field of view.

have been involved in the core work of the Shear Pipeline Testing working group. My main contributions were coding and testing the IM3SHAPE pipeline (see Appendix A) and running simulations to calibrate noise related effects for the Early SV clusters measurements (Appendix B) and SVA1 release.

The first publication using IM3SHAPE shear catalogues was by Melchior et al. (2014) and describes mass and light distribution in four massive clusters. Figure 2.10 shows a mass map for a RXC J2248.7-4431 cluster, the mass of which has been found to be in good agreement with the literature, and measured to be $M_{200} = 17.6^{+4.5}_{-4.0}\ 10^{14} M_\odot$. This work demonstrated the good performance of the DECam instrument and weak lensing data analysis pipeline. With the weak lensing pipeline already producing reliable shear measurements, in the next few years the DES will provide researchers with exciting opportunities for weak lensing science.

2.9 Weak lensing with the CFHTLenS survey

The Canada-France-Hawaii Telescope Lensing Survey (CFHTLenS) (Heymans et al. 2012) is a optical galaxy survey containing the largest weak lensing dataset to date. It cover 154 square degrees of deep multi-colour data obtained by the CFHT Legacy Survey, in five optical bands u^*, g', r', i', z. The 5σ limiting magnitude in the i-band was $i' \sim 25.5$. The survey had a wide and deep components, observations of which finished in early 2009. It used the 3.6m telescope and the MegaCam imager, with 36 CCDs totalling 340 megapixels. The area of the wide survey consisted of four fields: W1 (63.8 square degrees), W2 (22.6 square degrees), W3 (44.2 square degrees) and W4 (23.3 square degrees). Images were

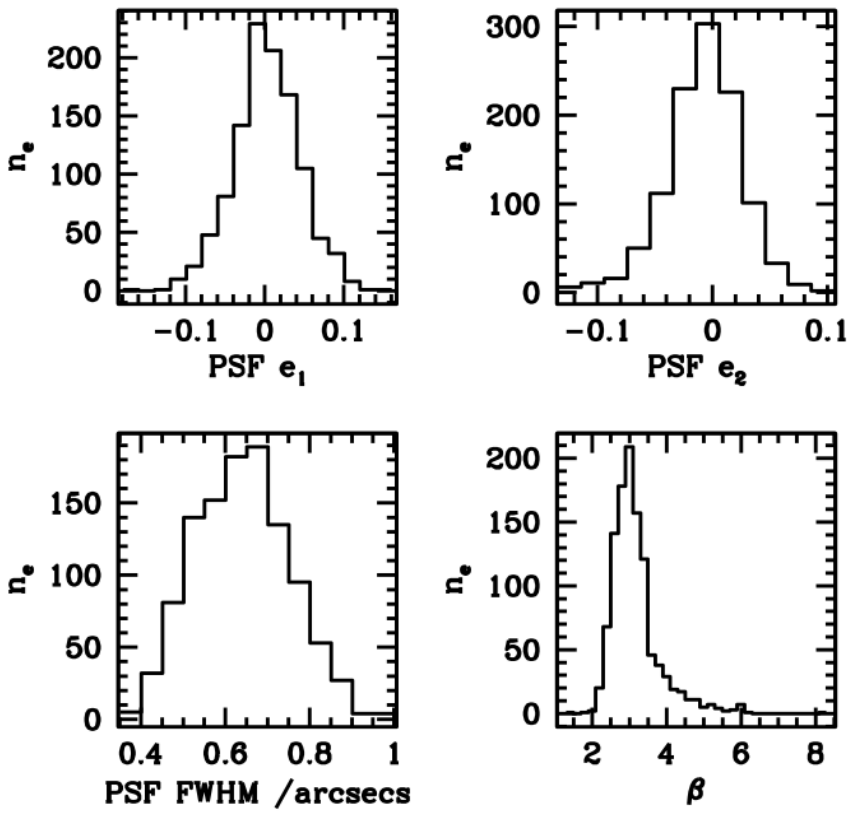

Figure 2.11. Distribution of PSF parameters in the CFHTLenS survey. Images of stars were fitted with parameter of the Moffat model $I(r) = I_0[1 + (r/r_s)^2]^{\beta}$, where $I(r)$ is the intensity at radius r, I_0 is the profile normalisation, r_s is the scale radius and β is a parameter affecting the profile of the PSF. Additionally, an elongation in e_1 and e_2 directions were fit.

reduced using the THELI pipeline (Erben et al. 2012). The survey strategy was optimised for weak lensing measurements: the lensing band i' was observed only in good seeing conditions. The distribution of PSF parameters is shown in Fig. 2.11 (Miller et al. 2012).

CFHTLenS shear was measured using the LENSFIT algorithm (Miller et al. 2012), which fits a bulge+disc galaxy model using a Bayesian framework. Only galaxies with signal-to-noise ratio of $SNR > 10$ were considered, which resulted in $\sim 1 \cdot 10^7$ galaxies in the catalog. This corresponds to galaxy density of $\sim 18/\mathrm{arcmin}^2$. Systematic errors in the shear measurement were calibrated using simulations, called GREAT-CFHTLenS and SkyMaker. Distributions of the measured ellipticity parameters, SNR and galaxy sizes are shown in Fig. 2.12.

Photometric redshifts were calculated for galaxies in the catalogue, reaching $z \sim 2$ (Hildebrandt et al. 2011). Peak of the stacked photo-z distribution was $z \sim 0.7$. Stacked

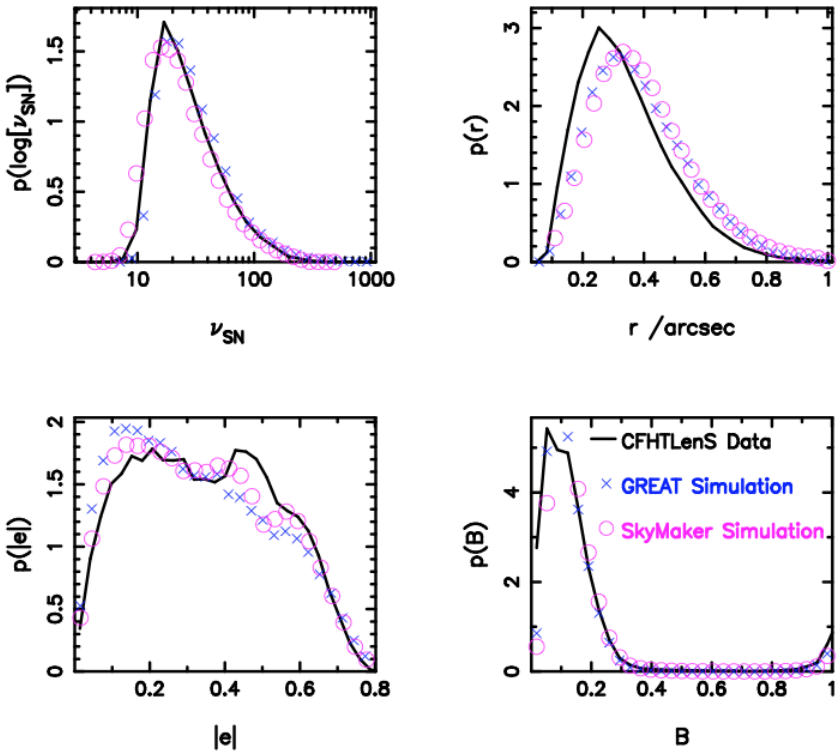

Figure 2.12. Distribution of measured parameters of galaxies in the CHFHTLenS survey. Upper panels: signal to noise ratio ν_{SN}, half light radius r/arcsec. Lower panels: ellipticity modulus $|e|$ and fraction of the flux in the bulge component B.

probability distributions of galaxy redshifts $P(z)$ is shown on Fig 2.13.

CFHTLenS has provided a dataset for many important measurements of gravitational lensing signal, used to constrain cosmological parameters. Kilbinger et al. (2013) used the 2D measurement of shear-shear correlation function to constrain cosmological parameters. Heymans et al. (2013) extended this analysis to include 5-bin tomography and mitigating the impact of intrinsic alignments. This weak lensing dataset was also used by Simpson et al. (2012) to test the laws of gravity. Higher order statistics of shear field were used to constrain cosmological parameters in Fu et al. (2014). Recently, Liu et al. (2014) used the shear peak statistics to put constraints in the Ω_m, σ_8, w_0 parameter set.

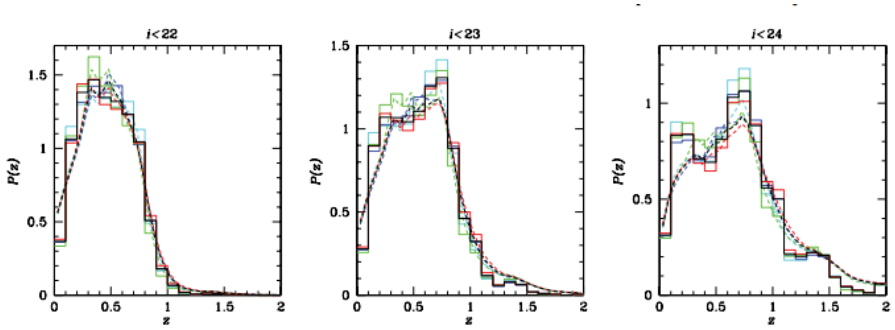

Figure 2.13. Stacked redshift probability distribution $P(z)$ for three values of limiting magnitude in the CFHTLenS dataset.

Chapter 3

Measurement and Calibration of Noise Bias in Weak Lensing Galaxy Shape Estimation

Weak gravitational lensing has the potential to constrain cosmological parameters to high precision. However, as shown by the Shear TEsting Programmes (STEP) and GRavitational lEnsing Accuracy Testing (GREAT) Challenges, measuring galaxy shears is a nontrivial task: various methods introduce different systematic biases which have to be accounted for. We investigate how pixel noise on the image affects the bias on shear estimates from a Maximum-Likelihood forward model-fitting approach using a sum of co-elliptical Sérsic profiles, in complement to the theoretical approach of an associated paper. We evaluate the bias using a simple but realistic galaxy model and find that the effects of noise alone can cause biases of order 1-10% on measured shears, which is significant for current and future lensing surveys. We evaluate a simulation-based calibration method to create a bias model as a function of galaxy properties and observing conditions. This model is then used to correct the simulated measurements. We demonstrate that, for the simple case in which the correct range of galaxy models is used in the fit, the calibration method can reduce noise bias to the level required for estimating cosmic shear in upcoming lensing surveys.

3.1 Introduction

Weak gravitational lensing is an important cosmological probe, which has the greatest potential to discover the cause of the accelerated cosmic expansion (e.g. Peacock et al. 2006; Albrecht et al. 2006, 2009). In the standard cosmological model dark energy affects both the expansion history of the universe and the rate of gravitational collapse of large scale structure. The rate of this collapse can be studied by observing the spatial distribution of dark matter at different times in the history of the universe. Gravitational lensing occurs when the path of light from distant galaxies is perturbed while passing through intervening matter. This phenomenon causes the images of galaxies to be distorted. The primary observable distortion is called gravitational shear, and typically causes the galaxy images to be stretched by a few percent. The scale of this effect is related to the amount of matter between the source and the observer, and to their relative geometry. Thus, cosmic shear can provide a valuable dataset for testing cosmology models (Kaiser 1992; Hu 1999).

Several upcoming imaging surveys plan to observe cosmic shear, including the KIlo-Degree Survey: KIDS, the Dark Energy Survey (DES)[1], the Hyper Suprime-Cam (HSC) survey[2] the Large Synoptic Survey Telescope (LSST)[3], Euclid[4] and WFIRST [5]. For these surveys, it is crucial that the systematics introduced by data analysis pipelines are understood and accounted for. The most significant systematic errors are introduced by (i) the measurements of the distance to the observed galaxies using photometric redshifts, (ii) intrinsic alignments of galaxies, (iii) modelling of the clustering of matter on the small scales in the presence of baryons, (iv) measurement of lensed galaxy shapes from imperfect images. In this chapter, I focus on the latter.

To evaluate the performance of shear measurement methods, simulated datasets have been created and released in form of blind challenges. The Shear TEsting Programme 1 (STEP1: Heymans et al. 2006), was the first in this series, followed by STEP2 (Massey et al. 2006). Both challenges aimed to test end-to-end shear pipelines and simulated galaxy images containing many physical effects including those stemming from telescope optics and atmospheric turbulence. A modified approach was taken in the GREAT08 (Bridle et al. 2009*a*,*b*) and GREAT10 (Kitching et al. 2011) challenges, which sought to isolate

[1] http://www.darkenergysurvey.org
[2] http://www.naoj.org/Projects/HSC/HSCProject.html
[3] http://www.lsst.org
[4] http://sci.esa.int/euclid
[5] http://exep.jpl.nasa.gov/programElements/wfirst/

independent parts of the data analysis process. They explored the impact of different true galaxy and image parameters on the shear measurement, by varying them one at a time among various simulation realisations. These parameters included signal to noise ratio, galaxy size, galaxy model, Point Spread Function (PSF) characteristics and others. The results showed that the shear measurement problem is intricate and complex. Existing methods proved to be sufficient for current surveys, but there is room for improvement for the future.

For a well resolved, blur-free, noise-free image, the galaxy ellipticity can be calculated by taking the moments of the image (Bonnet & Mellier 1995). However, a typical galaxy image used in weak lensing is highly affected by the observation process. The image degrading effects are (i) convolution with the PSF of the telescope, (ii) pixelisation of the image by the light buckets of the detector, (iii) pixel noise on the image due to the finite number of photons from the source and atmosphere (roughly Poisson) and detector noise (often assumed Gaussian), and (iv) galaxy colours being different from the stars used to map the PSF (Cypriano et al. 2010) and a function of position on the galaxy (Voigt et al. 2011).

Moment-based methods such as KSB (Kaiser et al. 1994), and most recently DEIMOS (Melchior et al. 2012), and FDNT (Bernstein 2010) measure the quadrupole moment of the image, using a masking function (often Gaussian) to counter the effects of noise, and then correct for the PSF. Decomposition methods, e.g. shapelets or a Gauss-Laguerre expansion, (Refregier 2001; Bernstein & Jarvis 2001; Nakajima & Bernstein 2006) use an orthogonal image basis set which can be easily convolved with the PSF. Noise is accounted for by regularisation of the coefficients matrix and truncating the basis set to a finite number of elements. Simple model fitting methods based on sums of Gaussians (Kuijken 1999; Bridle et al. 2002), Sérsic profiles (Miller et al. 2007; Peng et al. 2002), create an ellipticity estimator from a likelihood function. Stacking methods (Lewis 2009; Hosseini & Bethge 2009), which have been demonstrated for constant shear fields, average a function of the image pixels to increase the signal-to-noise ratio and then deconvolve the PSF.

All these methods introduce some level of systematic error, coming from different, method specific sources. Bias on the shear can result from inaccurate centroiding of the galaxy, for example see Lewis (2009). Another source, model bias, results from using a galaxy model which does not span the true range of galaxy shapes. Voigt & Bridle (2009) quantified the shear measurement bias from using an elliptical isophote galaxy model on

a galaxy with a a more complicated morphological structure in the presence of a PSF (see Lewis 2009 for a general proof). Melchior et al. (2009) investigated the effectiveness of shapelets at representing more realistic galaxies. Viola et al. (2010) and Bartelmann et al. (2011) quantified biases on the KSB method and investigated possibilities to correct for it.

Pixel noise bias arises from the fact that ellipticity is not a linear function of pixel intensities in the presence of noise and PSF. Hirata et al. (2004) showed its effects on second order moment measurements from convolved Gaussian galaxy images. The bias due to pixel noise on parameters fitted using Maximum Likelihood Estimators (MLEs) for elliptical shapes was demonstrated by (Refregier et al. 2012 hereafter R12), for the case when the noise is Gaussian and the correct galaxy model is known. It presented a general expression for the dependency of the bias on the signal to noise ratio. It also demonstrated the consistency of analytical and simulated results for the bias on the width for a one parameter Gaussian galaxy model. Although R12 and this chapter discuss biases that arise from ML forward fitting methods, I suspect that noise bias will play a role in every nonlinear parameter estimation method. For example, in moment - based methods, ellipticity is often defined as a ratio of moments of pixel intensities, and thus introduces nonlinearity (Melchior & Viola 2012).

In this chapter, I show the significance of this bias for weak lensing measurements using more realistic galaxy images. I find that the bias as a function of true input parameters is consistent with the theoretical framework derived in R12. Furthermore, I present a method to effectively remove this noise bias for realistic galaxy images. Using the IM3SHAPE shear measurement framework and code (Zuntz et al. 2013), I use a forward model fitting, Maximum Likelihood (ML) approach for parameter estimation. I create a model of the bias as a function of galaxy and PSF parameters by determining their bias from various simulations that sample parameter space. I apply this model to the noisy MLEs and demonstrate that this procedure successfully removes the noise bias to the accuracy required by upcoming galaxy surveys. By performing a calibration that depends on the specific statistics of every recorded galaxy, this method is independent of the overall galaxy and PSF parameter distributions.

This chapter is organised as follows. Section 3.2 summarises the equations governing the cosmic shear measurement problem and describes methods to quantify the biases on estimated parameters. I also discuss the requirements on those biases for lensing surveys,

followed by a summary of the cause of bias arising from image noise. In Section 3.3, I show the results of bias measurements. A method for correcting the noise bias based on numerical simulations is presented in Section 3.4. I conclude and briefly discuss this approach and alternatives in Section 3.5. In the Appendices, I detail the method used for measuring the multiplicative and additive bias and tabulate our results and fit parameters.

3.2 Shear measurement biases in model fitting

I first discuss the parametrisation of shear measurement biases, and present an overview of the model fitting approach. I summarise recent work on noise bias in a simple case, and then describe our shear measurement procedure and simulation parameters.

3.2.1 Quantifying systematic biases in shear estimation

In weak gravitational lensing the galaxy image is distorted by a Jacobian matrix (see Bartelmann & Schneider 1999; Bernstein & Jarvis 2001; Hoekstra & Jain 2008 for reviews)

$$\boldsymbol{M} = \begin{bmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{bmatrix}, \tag{3.1}$$

where κ is the convergence and $\gamma = \gamma_1 + i\gamma_2$ is the complex gravitational shear.

For a galaxy with elliptical isophotes I can define the complex ellipticity e as

$$e = \frac{a-b}{a+b}\, e^{2i\phi}, \tag{3.2}$$

where b/a is the galaxy minor to major axis ratio and ϕ is the orientation of the major axis anticlockwise from the positive x-axis. The post-shear lensed ellipticity e^l is related to the intrinsic ellipticity e^i by

$$e^l = \frac{e^i + g}{1 + g^* e^i} \tag{3.3}$$

for $|g| \leq 1$ (Schneider & Seitz 1994), where $g = \gamma/(1-\kappa)$ is the reduced shear. In the weak lensing regime $\kappa \ll 1$, $\gamma \ll 1$ and $g \approx \gamma$. I assume $\kappa \ll 1$ throughout this chapter.

Galaxies have intrinsic ellipticities which are typically an order of magnitude larger than the shear. If the ellipticity is defined as in Equation 3.2 then the average ellipticity is an unbiased shear estimator (e.g. Schneider & Seitz 1994). In practice e^l is averaged over a

Survey	m_i	c_i
Current	0.02	0.001
Upcoming future	0.004	0.0006
Far future	0.001	0.0003

Table 3.1. Summary of the requirements for the bias on the shear for current, upcoming and far future surveys.

finite number of galaxies and the error on the shear estimate (referred to as 'shape noise') depends on the distribution of galaxy intrinsic ellipticities and the number of galaxies analysed.

The accuracy of a shape measurement method can be tested on a finite number of images in the absence of shape noise by performing a 'ring-test' (Nakajima & Bernstein 2006). In the ring-test, the shear estimate is obtained by averaging the measured e^o estimates from a finite number of instances of a galaxy rotated through angles distributed uniformly from 0 to 180 degrees. If $\hat{e}^l$ is the measured lensed ellipticity, then the shear estimate is $\hat{\gamma} = \langle \hat{e}^l \rangle$ and the bias on the shear is

$$b[\hat{\gamma}] = \langle \hat{e}^l \rangle - \gamma^t, \tag{3.4}$$

where γ^t is the true shear. This bias on the shear is usually quantified in terms of multiplicative and additive errors m_i and c_i for both shear components $i = 1, 2$ such that

$$\hat{\gamma}_i = (1 + m_i)\gamma_i^t + c_i, \tag{3.5}$$

assuming $\hat{\gamma}_1$ does not depend on γ_2^t, and vice versa (Heymans et al. 2006). The requirements on the level of systematic errors for current and future galaxy surveys are expressed in terms of m_i, c_i in Amara & Refregier (2007) and are summarised in Table 4.1.

3.2.2 Galaxy shear from model fitting

A simple approach to measuring ellipticity is to use a parametric model. For galaxy fitting, models such as sums of Gaussians (Kuijken 1999; Bridle et al. 2002), Sérsic profiles (Miller et al. 2007), and Gauss - Laguerre polynomials (shapelets) (Refregier 2001; Bernstein & Jarvis 2001; Nakajima & Bernstein 2006) were used.

In general, model fitting methods are based on a likelihood function. Under uncorre-

lated Gaussian noise, this function is

$$\mathcal{L} = p(\theta|\mathbf{I}, \mathcal{M}) \tag{3.6}$$

$$\log \mathcal{L} = \chi^2 = \frac{1}{2}\sum_{i=1}^{N}[M_i(\theta) - I_i]^2/\sigma_i^2 \tag{3.7}$$

where θ is a set of variable model parameters, $\mathbf{I}$ is the observed galaxy image, $\mathcal{M}$ is a model function, $\mathbf{M}(\theta)$ is the model image created with parameters θ, and N the number of pixels in images $\mathbf{I}$ and $\mathbf{M}$. These equations assume a known noise level on each pixel σ_i, which is often assumed constant $\sigma_i = \sigma_{\rm noise}$. Sometimes a prior on the parameters is used to create a posterior function.

Usually an ellipticity estimator is derived from this likelihood function; so far maximum likelihood estimators (MLE; e.g. IM3SHAPE, Shapelets), mean likelihood (IM2SHAPE) and mean posterior (e.g. LENSFIT) have been used. I use the MLE in this chapter.

Parametric models based on elliptical profiles typically use the following galaxy parameters: centroid, ellipticity, size, flux and a galaxy light profile parameter. Often a combination of two Sérsic profiles (Sérsic 1963) is used to represent the galaxy bulge and disc components, with identical centroids and ellipticities.

The model also contains information about other effects influencing the creation of the image. These image parameters are not often a subject of optimisation: noise level $\sigma_{\rm noise}$, PSF kernel and the pixel integration kernel. SNR is often defined as SNR $= \sqrt{\sum_{i=1}^{N} I_i^2}/\sigma_{\rm noise}$ and this definition will be used throughout this chapter. This definition of SNR is the same as in GREAT08, but different to GREAT10: SNR=20 here corresponds to SNR=10 in GREAT10.

3.2.3 Noise bias

The bias of parameter estimation for MLEs in the context of galaxy fitting was first studied by R12. The authors derived general expressions for the covariance and bias of the MLE of a 2D Gaussian galaxy model convolved with a Gaussian PSF. For a nonlinear model, in the Taylor expansion of χ^2 (in equation 3.7) the terms in even power of the noise standard deviation are found to contribute to the estimator bias. The analytical results were confirmed by simulations using a single parameter toy model. It was also noted that the bias is sensitive to the chosen parametrisation, especially if the parameter space is

bounded.

Expanding on work presented in R12, I introduce a generalisation of noise bias equations in R12 to include the case when the galaxies measured have unknown morphologies. I use an expansion around the *parameters of the best fit model to the noiseless data*. Note that in this formalism there is no *true* value of the parameters, as the real galaxy can have morphology which is not captured by the model. This in fact gives rise to model bias, which has to be evaluated empirically from low noise calibration data.

Let us define the following set of variables

$$g_p \quad \text{- true noiseless image} \tag{3.8}$$

$$\mathbf{a}^t \quad \text{- vector of best fitting parameters for } g_p \tag{3.9}$$

$$f_p(\mathbf{a}) \quad \text{- model image with parameters } \mathbf{a} \text{ at pixel } p \tag{3.10}$$

$$n_p \quad \text{- noise at pixel } p \tag{3.11}$$

$$\sigma_n \quad \text{- noise standard deviation} \tag{3.12}$$

$$\rho = \frac{f_0}{\sigma_n} \quad \text{- signal to noise, assume total flux } f_0 = 1 \tag{3.13}$$

Log - likelihood of image data given a set of parameters is

$$-2\log\mathcal{L} = \chi^2(\mathbf{a}) = \frac{1}{\sigma^2}\sum_p \left[g_p + n_p - f_p(\mathbf{a})\right]^2 \tag{3.14}$$

Maximum likelihood point $\hat{\mathbf{a}}$ is defined as

$$\hat{\mathbf{a}} = \arg\min_{\mathbf{a}} \chi^2(\mathbf{a}) \quad \Leftrightarrow \quad \frac{\partial\chi^2(\mathbf{a})}{\partial a_k} = 0 \quad \forall \quad k \tag{3.15}$$

where

$$\frac{\partial\chi^2(\mathbf{a})}{\partial a_k} = \frac{2}{\sigma^2}\sum_p \left[\left[g_p + n_p - f_p(\mathbf{a})\right](-1)\frac{\partial f_p(\mathbf{a})}{\partial a_k}\right] \tag{3.16}$$

An expansion of $\hat{\mathbf{a}}$ around the true parameters $\mathbf{a}^t$ gives

$$\hat{a}_k = a_k^t + \underbrace{\alpha a_k^{(1)} + \alpha^2 a_k^{(2)}}_{\delta a_k} \tag{3.17}$$

Expanding $f_p(\hat{\mathbf{a}})$ about $\mathbf{a}^t$ gives

$$f_p(\hat{\mathbf{a}}) = f_p(\mathbf{a}^t) + \sum_i \delta a_i \frac{\partial f_p(\mathbf{a}^t)}{\partial a_i} + \frac{1}{2}\sum_{ij} \delta a_i \delta a_j \frac{\partial^2 f_p(\mathbf{a}^t)}{\partial a_i \partial a_j} \tag{3.18}$$

And finally, expanding $\frac{\partial f_p(\hat{\mathbf{a}})}{\partial a_i}$ gives

$$\frac{\partial f_p(\hat{\mathbf{a}})}{\partial a_k} = \frac{\partial f_p(\mathbf{a}^t)}{\partial a_k} + \sum_i \delta a_i \frac{\partial^2 f_p(\mathbf{a}^t)}{\partial a_k a_i} + \frac{1}{2}\sum_{ij} \delta a_i \delta a_j \frac{\partial^3 f_p(\mathbf{a}^t)}{\partial a_k a_i a_j} \tag{3.19}$$

Now we substitute (3.17), (3.18), (3.19) into (3.16), ignore terms $O(\alpha > 2)$, and use the residual $r_p := g_p - f_p(\mathbf{a}^t)$. For convenience let's use $\frac{\partial f_p(\mathbf{a}^t)}{\partial a_k} = \frac{\partial f_p^t}{\partial a_k}$, and then we can write

$$\begin{aligned}
&\frac{\partial \chi^2(\mathbf{a})}{\partial a_k} = \\
&\frac{-2}{\sigma^2}\sum_p \left(\alpha \left[n_p \frac{\partial f_p^t}{\partial a_k} - \sum_i a_i^{(1)} \frac{\partial f_p^t}{\partial a_i}\frac{\partial f_p^t}{\partial a_k} + r_p \sum_i a_i^{(1)} \frac{\partial^2 f_p^t}{\partial a_i \partial a_k} \right] \right. \\
&\qquad +\alpha^2 \left[n_p \sum_i a_i^{(1)} \frac{\partial^2 f_p^t}{\partial a_i \partial a_k} - \sum_{ij} a_i^{(1)} a_j^{(1)} \frac{\partial f_p^t}{\partial a_i}\frac{\partial^2 f_p^t}{\partial a_j \partial a_k} \right. \\
&\qquad\quad - \sum_i a_i^{(2)} \frac{\partial f_p^t}{\partial a_i}\frac{\partial f_p^t}{\partial a_k} - \frac{1}{2}\sum_{ij} a_i^{(1)} a_j^{(1)} \frac{\partial^2 f_p^t}{\partial a_i \partial a_j}\frac{\partial f_p^t}{\partial a_k} \\
&\qquad\quad \left. \left. + r_p \sum_i a_i^{(2)} \frac{\partial^2 f_p^t}{\partial a_i \partial a_k} + \frac{1}{2} r_p \sum_{ij} a_i^{(1)} a_j^{(1)} \frac{\partial^3 f_p^t}{\partial a_i \partial a_j \partial a_k} \right] \right)
\end{aligned} \tag{3.20}$$

Using definitions in the following Eqn. 3.21

$$\begin{aligned}
D_{ip}^{(1)} &:= \frac{\partial f_p^t}{\partial a_i} \\
D_{ijp}^{(2)} &:= \frac{\partial^2 f_p^t}{\partial a_i \partial a_j} \\
D_{ijkp}^{(3)} &:= \frac{\partial^3 f_p^t}{\partial a_i \partial a_j \partial a_k} \\
F_{ij} &:= \frac{\partial f_p^t}{\partial a_i}\frac{\partial f_p^t}{\partial a_j} = D_{ip}^{(1)} D_{jp}^{(1)}
\end{aligned} \tag{3.21}$$

and summation over repeated indices, we can collect the first order terms

$$O(\alpha) = 0 \quad \Leftrightarrow \quad D_{kp}^{(1)} n_p - F_{ik} a_i^{(1)} + r_p D_{ikp}^{(2)} a_i^{(1)} = 0 \quad \forall \quad a_k \tag{3.22}$$

Solving for first order terms

$$a_i^{(1)} = (F_{ik} - r_p D_{ikp}^{(2)})^{-1} D_{kp}^{(1)} n_p \tag{3.23}$$

$$\langle a_i^{(1)} \rangle = 0 \tag{3.24}$$

As expected from the result of (Refregier et al. 2012), the first order terms average to zero. For convenience, we can use a Fisher matrix modified by the residual r_p

$$\tilde{F}_{ik} := (F_{ik} - r_p D_{ikp}^{(2)}) \tag{3.25}$$

Covariance between two parameters is

$$\langle a_i^{(1)} a_j^{(1)} \rangle = \tilde{F}_{ij}^{-1} D_{jp}^{(1)} \underbrace{\langle n_p n_p \rangle}_{\sigma^2} D_{jp}^{(1)} \tilde{F}_{ij}^{-1} \tag{3.26}$$

$$= \sigma^2 \tilde{F}_{ij}^{-1} F_{ij} \tilde{F}_{ij}^{-1} \tag{3.27}$$

Covariance between a parameter and a noise pixel, using (3.23), is

$$\langle a_i^{(1)} n_m \rangle = \langle \tilde{F}_{ij}^{-1} D_{jp}^{(1)} n_p n_m \rangle = \tilde{F}_{ij}^{-1} D_{jp}^{(1)} \delta(m = p) \sigma_n^2 \tag{3.28}$$

Solving for second order terms in (3.20) gives

$$O(\alpha^2) = 0 \quad \Leftrightarrow \tag{3.29}$$

$$\begin{aligned} & a_i^{(1)} D_{ikp}^{(2)} n_p \\ & + a_i^{(1)} a_j^{(1)} \left(\frac{1}{2} r_p D_{ijkp}^{(3)} - D_{jp}^{(1)} D_{ikp}^{(2)} - \frac{1}{2} D_{kp}^{(1)} D_{ijp}^{(2)} \right) \\ & - a_i^{(2)} \tilde{F}_{ik} \qquad = 0 \quad \forall \quad a_k \end{aligned} \tag{3.30}$$

Averaging second order terms, rearranging and substituting (3.28) and (3.27) we arrive at

$$\begin{aligned}\langle a_i^{(2)}\rangle = \sigma_n^2 \tilde{F}_{ik}^{-1}\big[\\ &+\tilde{F}_{lj}^{-1} D_{jp}^{(1)} D_{lkp}^{(2)} \\ &+\frac{1}{2}\tilde{F}_{lj}^{-1} F_{lj} \tilde{F}_{lj}^{-1} D_{ljkp}^{(3)} r_p \\ &-\tilde{F}_{lj}^{-1} F_{lj} \tilde{F}_{lj}^{-1} D_{jp}^{(1)} D_{lkp}^{(2)} \\ &-\frac{1}{2}\tilde{F}_{lj}^{-1} F_{lj} \tilde{F}_{lj}^{-1} D_{kp}^{(1)} D_{ljp}^{(2)}\big]\end{aligned} \tag{3.31}$$

This is our main result, showing how the residual between the best fit and the noiseless real galaxy image is modifying the noise bias equations. It introduces the third derivative of the model function scaled by the residual $D_{ljkp}^{(3)} r_p$. It also modifies the Fisher matrix, and may prevent the cancellation of two terms in $D_{jp}^{(1)} D_{lkp}^{(2)}$. If there is no model bias, and the residual $r_p = 0$, and the expression reduces to the result from (Refregier et al. 2012) shown in Eqn. 3.32.

$$\begin{aligned}\langle a_i^{(2)}\rangle &= \sigma_n^2 F_{ik}^{-1}\big[\\ &+ F_{lj}^{-1} D_{jp}^{(1)} D_{klp}^{(2)} \\ &- F_{lj}^{-1} D_{jp}^{(1)} D_{klp}^{(2)} \\ &- \frac{1}{2} F_{lj}^{-1} D_{kp}^{(1)} D_{ljp}^{(2)}\big] \\ &= -\frac{1}{2}\sigma_n^2 F_{ik}^{-1} F_{lj}^{-1} D_{kp}^{(1)} D_{ljp}^{(2)}.\end{aligned} \tag{3.32}$$

The presence of the terms scaling with σ_n^2 gives rise to the noise bias effect. The reason for this effect can be understood intuitively, as follows: the non-linear nature of the model fitting problem is causing the final maximum likelihood solution to be a non-linear function of normally distributed pixel intensities. When averaged over many pixel noise realisations, these estimators are not normally distributed. In fact, this distribution has a non-zero skew and which causes its mean to be biased.

3.2.4 Im3shape pipeline

The analyses in this chapter were performed using the IM3SHAPE shear measurement framework and code. Here I outline the system, which will be described in more detail in Zuntz et al. (2013).

Each simulated galaxy is fitted with a model containing two co-centric, co-elliptical Sérsic components, one de Vaucouleurs bulge (Sérsic index=4) and one exponential disc (Sérsic index=1). The amplitudes of the bulge and disc were free but the ratio of the half light radii was fixed to 1.0. They are convolved with the true Moffat PSF model to produce a model image. Since there is high resolution structure in de Vaucouleurs bulges I made the models at a higher resolution than the final images. I use a resolution three times higher in the outer regions and 45 times higher in the central 3×3 pixels of the final image. Since very highly elliptical images are hard to simulate accurately I restrict the allowed space of models to those with $|e| < 0.95$.

I find the peak of the likelihood using the Levenberg-Marquadt method (Lourakis Jul. 2004) using numerical gradients of each image pixel in the likelihood. I tested the performance of the optimiser for variety of input galaxy and image parameters to ensure that the optimiser always converges to a local minimum by evaluating the likelihood in the neighbourhood of the found best fit point for multiple test noise realisations. In this nonlinear optimisation problem multiple likelihood modes are possible. However, for our simple model, I found that usually there was only one local minimum (i.e. the bias results did not depend on the starting parameters given to the minimiser). I will discuss this further in (Zuntz et al. 2013).

3.2.5 Simulation parameters

The galaxies used for this study were created using a two component model: a Sérsic profile of index 4 for the bulge and a Sérsic profile of index 1 for the disc. Both components have the same centroid, ellipticity and scale radius. The galaxy model used for fitting encompassed the one used to create the true galaxy image; therefore I am isolating the noise bias effect from the model bias effect in this study. The PSF was modelled as a Moffat profile with a FWHM of 2.85 pixels and Moffat β parameter of 3 (see, e.g., Bridle et al. 2009*a* for a definition of the Moffat and the notation adopted here). I use the same PSF in the fit as in the simulated images to prevent any bias effects caused by incorrect modelling of the PSF. I fit a total of 7 parameters: galaxy centroid x, y; galaxy ellipticity e_1, e_2; galaxy size r; bulge flux F_b; and disc flux F_d.

I expect variation in the following physical parameters to have the most significant influence on the noise bias, and therefore the bias will be evaluated as a function of:

- Signal-to-noise ratio (SNR),
- Intrinsic galaxy ellipticity,
- PSF ellipticity,
- Size of the galaxy compared to the size of the PSF, expressed as R_{gp}/R_p, which is the ratio of the FWHM of the convolved observed object and the FWHM of the PSF. Note that this is not the same as the parameter I fit. This is because the noise bias strongly depends on the PSF parameters, and the galaxy radius parameter alone would not fully capture this dependence.
- Light profile of the galaxy, described by the flux ratio $F_b/(F_b + F_d)$, which is the flux of the bulge component divided by total flux of the galaxy. For a purely bulge galaxy, $F_b/(F_b + F_d) = 1$ and for a disc galaxy $F_b/(F_b + F_d) = 0$. In our model, I allow the amplitudes of the components to be negative, so the flux ratio can take both values $F_b/(F_b+F_d) > 1$ and $F_b/(F_b+F_d) < 0$. Therefore, for $F_b/(F_b+F_d) > 1$, the galaxy has a negative disc component, which results in the galaxy being less 'peaky' than a galaxy with $F_b/(F_b+F_d) = 1$, and the galaxy model image may even be more similar to a galaxy with $F_b/(F_b + F_d) < 1$. An alternative might be to use a more flexible radial profile, for example a larger number of Sersic components, or allowing the Sersic indices to be free parameters in the fit.

These parameters will be used to create a model for the noise bias. I expect these physical parameters to best encapsulate the main dependencies of the bias, although I am aware that there may exist other statistics that better capture bias variation.

I do not show the effect of the galaxy centroid on the bias, as no significant dependence on this parameter was found in our experiments. I measured the noise bias for a simulated galaxy image with identical model parameters, once located in the middle of a pixel and once on the edge of a pixel. I found no difference in ellipticity bias to our desired precision.

I note that centroiding errors in the case of model fitting may impact ellipticity estimates differently when compared with moment based and stacking methods (e.g. Melchior et al. 2012). For model fitting approaches, the centroid is just another parameter in the fitted model. In the simulations, the galaxy centroid is randomised. Should there be any dependence of the ellipticity bias on the galaxy centroid, uniform randomisation of the

	Parameter	Fiducial	Deviations
D1	SNR	20	40, 200
D2	R_{gp}/R_p	1.62	1.41, 1.82
D3	$F_b/(F_b+F_d)$	0.5	0 , 1
D4	e_1^{PSF}	0.05	0, 0.1

Table 3.2. Summary of parameters used for simulations. For D3 two different parametrisations are shown for clarity.

true galaxy positions within one pixel allows this potential source of bias to be included fairly in our simulations.

The values for the simulation parameters are summarised in Table 3.2. Their choice is based on galaxies used in GREAT08. I define a fiducial parameter set and make departures D1 to D4 in one parameter at a time using the values given in the Table. I restrict our analysis to SNR values of 20 and greater because I find convergence of the minimiser does not pass our quality tests at lower values.

However, the SNR values of most interest for upcoming surveys are low, and therefore I use the lowest SNR I can use with confidence by default for all simulations. I investigate a SNR value of 200 which matches that of the GREAT08 LowNoise simulation set, plus an intermediate value of 40 which is also used in GREAT08. By default, I use a galaxy with half the flux in a bulge and half in a disc. The two perturbations I consider are to pure bulge and pure disc. Finally I explore the dependence of noise bias on the PSF ellipticity, spanning the range from zero to 10%.

For the minimisation parameters used in this chapter, IM3SHAPE takes around one second per galaxy, which is typical for model fitting methods. To obtain our desired accuracy on noise bias I needed to simulate 2.5 million galaxies for each set of simulation parameters shown in Table 3.2. Therefore the computations shown in this chapter took of order 1 year of CPU time. This computational burden limited the number of points I could show on the figures to 3 per varied parameter.

3.3 Evaluation of the noise bias effect

In this section I evaluate the noise bias as a function of galaxy and image parameters. I define the noise bias on an ellipticity measurement as

$$b[\hat{e}] = \langle \hat{e} \rangle - e^{\text{true}}. \tag{3.33}$$

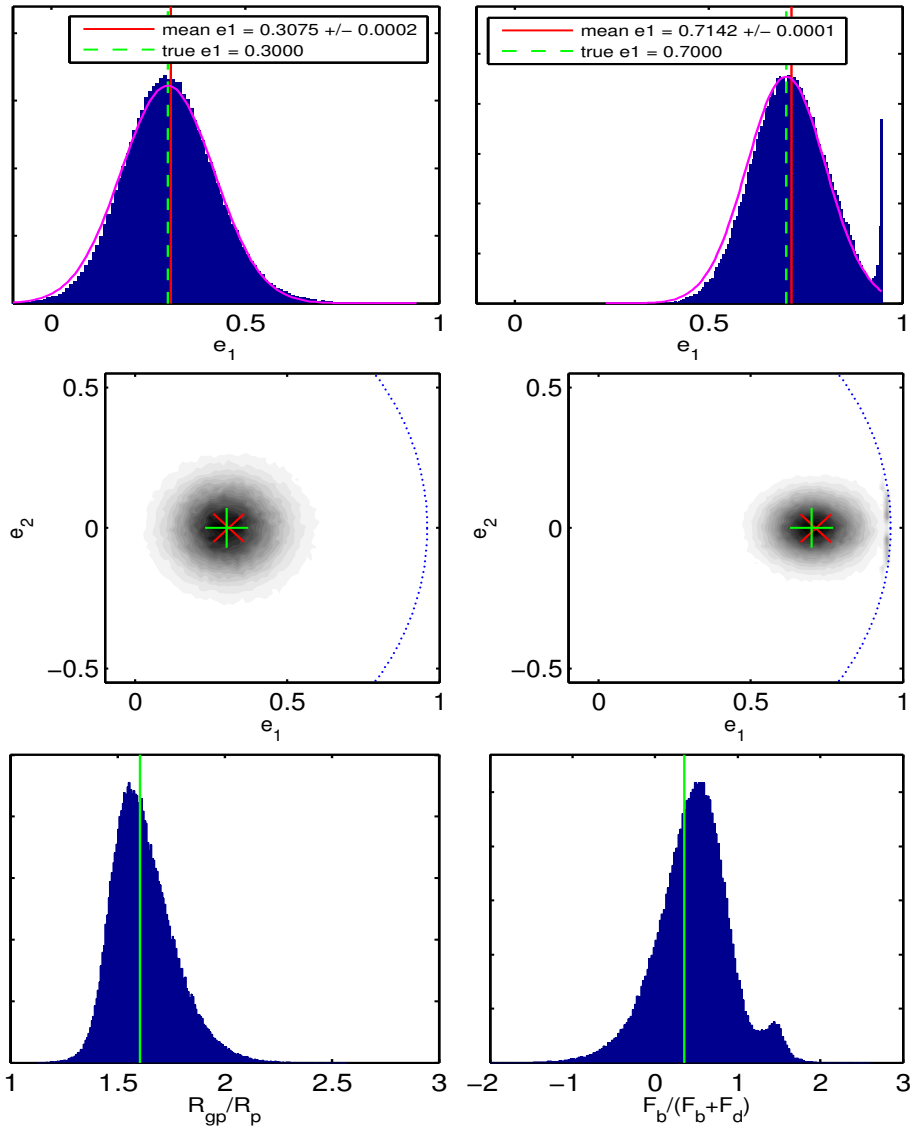

Figure 3.1. Histograms of ML parameter estimates for the fiducial galaxy model. Top panels show the distribution of measured ellipticity ($\hat{e}_1$) parameters for true intrinsic galaxy ellipticity of 0.3 (left) and 0.7 (right), marked with green dashed line. The empirical mean of these distributions is marked with a red solid line. The magenta line shows the Gaussian probability distribution centered on the true ellipticity and with the same variance as the distribution of $\hat{e}_1$. The middle panels shows the distribution of the ML estimates for both ellipticity components – for true intrinsic ellipticity of [0.3,0.0] (left) and [0.7,0.0] (right), marked with the plus sign. The mean of this distribution is marked with a cross sign. The effective boundary on the ellipticity parameter space ($|e| = 0.95$) is marked with black dotted line. The bottom panels show histograms of measured size (R_{gp}/R_p) and light profile ($F_b/(F_b + F_d)$) parameters. True values for these parameters are marked with red solid line – true $R_{gp}/R_p = 1.6$ and true $F_b/(F_b + F_d)$.

I calculate the bias using the following procedure: I create a galaxy image with some true ellipticity, add a noise map and measure the MLE of the ellipticity. Then, I repeat this procedure with different noise realisations which results in a distribution of noisy MLE ellipticities. The difference between the mean of this distribution and the true galaxy ellipticity is the bias on ellipticity.

The histograms of ML estimates for 300 thousand noise realisations are plotted in Figure 3.1 to illustrate the nature of the noise bias. The galaxy and image had default parameters described in Section 3.2.5 and intrinsic ellipticities of $e_1 = 0.3$, $e_2 = 0$ and $e_1 = 0.7$, $e_2 = 0$ in the left and right upper and middle panels, respectively. The spread of values comes from the Gaussian noise added to the images to approximate the finite number of photons arriving on the detector. As discussed in Section 3.2.5, I assume a default SNR value of 20.

Two effects contribute significantly to the bias on ellipticity for the left hand panels in which the true ellipticity is $e_1 = 0.3$, $e_2 = 0$. The ellipticity distribution is slightly skewed away from being a Gaussian. There is a larger tail to high ellipticity values than to negative ellipticity values. The peak is shifted to lower ellipticities, which is also visible in the two-dimensional histogram in the middle-left panel of Figure 3.1. Overall there is a net positive bias to larger ellipticity values, as shown by the vertical solid line which is to be compared with the vertical dashed line placed at the true value. Although this net positive bias is hard to see by eye, it is significant at the level of shear measurement accuracy required from future observations. This is discussed in more detail in the following sections.

Furthermore, the ellipticity parameter space is theoretically bounded at an ellipticity modulus of unity. This is exacerbated by any realistic measurement method which will break down just short of unity. The consequence of this effect is visible for a galaxy with true intrinsic ellipticity of $|e| = 0.7$, shown in the upper-right and middle-right panels. For this example, it counteracts the noise bias effect by reducing the amount of overestimation. For more noisy or smaller galaxies, which will have larger variance in the ellipticity MLEs, this effect will be stronger and may even cause the ellipticity to be underestimated, see 3.2 for an illustration of this.

No skew of the $\hat{e}_2$ distribution is visible on the middle panel of Figure 3.1. In fact, when $e_2^{galaxy} = e_2^{PSF} = 0$, then $b[\hat{e}_2]$ is consistent with zero to our accuracy.

Distributions of other fitted parameters are also biased and skewed, as discussed in R12. I show histograms of fitted galaxy size and galaxy light profile in the two bottom panels

of Figure 3.1. The convolved galaxy to PSF size ratio peaks at lower values than the ones that are used in the input simulation but there is a tail to larger values. Overall the mean is biased low by around 2%. The flux ratio is skewed to larger values and overestimated by around 8%. Moreover, this distribution has two modes; one close to the truth, and one close to $F_b/(F_b + F_d) = 1.5$. This kind of bimodality is not unexpected in nonlinear problems. Here it may be related to the characteristics of the $F_b/(F_b + F_d)$ parameter; one of components in models with $F_b/(F_b + F_d) = 1.5$ can have a negative flux and the corresponding image can in fact be more similar to the one with $F_b/(F_b + F_d) = 0.5$.

The shear measurement biases thus depend on the galaxy intrinsic ellipticity in a non-trivial way. However, this can be converted into the shear measurement bias for a population of galaxies at different orientations using the ring test.

3.3.1 Measurement of the bias on the shear

The multiplicative and additive bias was measured using the following procedure.

1. Evaluate the bias on a grid in observed ellipticity: A grid in observed ellipticity parameter was created for each test galaxy in Table 3.2. This grid consisted of 8 angles on a ring. At each angle, 15 ellipticity magnitudes were used in range $\{0, 0.05, \ldots, 0.7\}$. This grid is presented in Figure 3.2. For each point on this grid, I evaluate 20000 noise realisations, and average them to obtain the bias. The number of noise realisations is chosen so that the uncertainty on the mean was smaller than $\sigma_e < 10^{-3}$.

2. Create a model of the bias as a function of observed ellipticity: A third order 2D polynomial was fit to the surface of the bias. Not all terms in the 2D expansion were used to avoid overfitting of the data. In particular, I used $\{1, e_1, e_1^2, e_2^2, e_1^3\}$ for fitting the bias on e_1, analogously for e_2. This expansion takes into account the inherent rotational symmetry of the problem: rotating galaxy ellipticity and PSF ellipticity vectors results in the rotation of the bias vector.

3. Perform a ring-test to calculate m and c: The parametric model of the bias surface allows us to perform a ring test at any desired intrinsic ellipticity.

The upper panels of Figures 3.2 present the grid (dots) and interpolated surface (colour scale) of the magnitude of bias as a function of true e_1 and e_2 for a circular and elliptical

PSF. I note that for circular PSF within the modelled range, the bias surface has a circular symmetry which demonstrates that the problem is symmetric and that the effect of the pixel orientation with respect to the galaxy is not strong. The lower panels of Figure 3.2 present cross sections of the above grid and surface for each angle. The bias on ellipticity changes sign for large intrinsic ellipticities. This is due to the edge effect of the ellipticity parameter space, described in 3.3.

3.3.2 Characteristics of the shear bias

For the default galaxy and image parameters I find a multiplicative shear measurement bias of a few per cent. For an intrinsic galaxy ellipticity of 0.3 I find $m = 0.02$ which is an order of magnitude larger than the requirement for upcoming surveys. The additive shear measurement bias is around $c = 2 \times 10^{-3}$ which is larger than the requirement for upcoming surveys, and around an order of magnitude larger than the requirement for far-future surveys.

The multiplicative and additive shear measurement bias is shown as a function of galaxy and image parameters in Figure 3.3. Data points for those plots are listed in Table 3.3, and the functions I fitted are given in equations in Table 3.4, both in the following section 3.3.3.

The upper panels show the dependence on the image SNR. This demonstrates clearly that the bias I observe is truly a noise bias, since the biases tend to zero at high SNR. Indeed for a SNR of 200 the biases are well below the requirement even for far-future surveys. The dependence on SNR is well described by a quadratic function, shown as a fitted line, as discussed anecdotally (Bernstein, priv. com.) and as expected from the derivations in Hirata & Seljak (2003) and R12.

The upper middle panels of Figure 3.3 show the dependence on the ratio of convolved galaxy to PSF size, as defined in Section 3.2.5. The derivations in R12 showed that for Gaussian functions, the bias on the size parameter increases with the size of the PSF (Eq. 17). In our simulations the bias on the shear has a similar trend, as I observe an increased bias with decreased galaxy size relative to the PSF. The bias is reduced by a factor of almost three when the convolved galaxy to PSF size increases from 1.41 to the default value of 1.62. I modelled this dependence by using inverse power expansion with terms in $(R_{gp}/Rp - 1)^{-2}$ and $(R_{gp}/Rp - 1)^{-3}$.

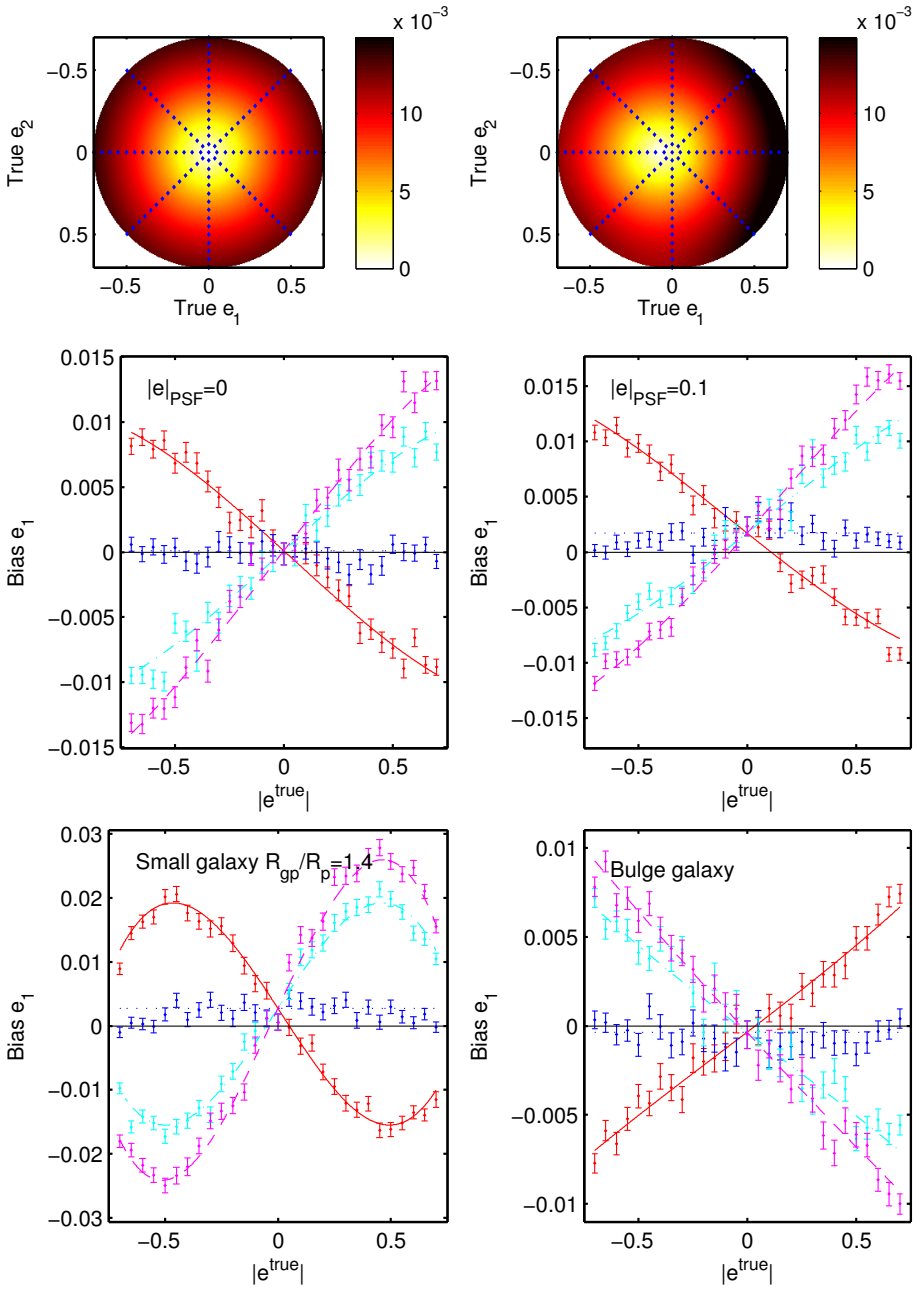

Figure 3.2. Top panels: The colourscale presents the model of the magnitude of the bias on the estimated galaxy ellipticity, $|b[\hat{e}_1] + ib[\hat{e}_2]|$ as a function of true galaxy ellipticity e_1 and e_2 for PSF ellipticity $e_{\rm PSF} = \{0.0, 0.0\}$ (left) and $e_{\rm PSF} = \{0.1, 0.0\}$ (right). The model was created using biases measured from simulations on a grid of true ellipticity values shown by the diamond points. Middle and bottom panels: bias on $\hat{e}_1$ as a function of true absolute ellipticity $|e|$ for $e_{\rm PSF} = \{0.0, 0.0\}$ (middle left) and $e_{\rm PSF} = \{0.1, 0.0\}$ (middle right), $R_{gp}/R_p = 1.4$ (bottom left) and pure bulge (bottom right). All other parameters are held at the fiducial values (see Section 3.2). Lines (dashed magenta , dash - dotted cyan, dotted blue, solid red) correspond to true ellipticity angles $\{0, \pi/8, \pi/4, 3\pi/8\}$ joined with $\{\pi/2, 5\pi/8, 3\pi/4, 7\pi/8\}$. Lines are third order polynomial fits to the points. The middle left and right panels correspond to the fiducial galaxy model with circular and highly elliptical PSF, respectively, as in the top panels.

The lower middle panels of Figure 3.3 show the bias as a function of the flux ratio. Both multiplicative and additive bias change signs when the galaxy light profile changes from bulge to disc. Bulges are underestimated and discs are overestimated. This peculiar behaviour of the bias demonstrates the complexity of this problem. I use a straight line to fit the points, and this works reasonably well.

The dependence on PSF ellipticity is shown in the bottom panels of Figure 3.3. As expected, e.g. from Paulin-Henriksson et al. (2008), the dependence of the additive shear measurement bias is much greater than that of the multiplicative bias. The additive shear bias dependence is very close to linear (shown by the fitted lines). Rotational symmetries in the problem, also visible on Figure 3.2 indicate that there is very little dependence on the pixel orientation with respect to the PSF and galaxy. This essentially means that I can use results for the PSF aligned with the x - axis for any other PSF angle, by rotating the coordinate system. Moreover, this indicates that the size of a pixel with respect to the size of the convolved object does not play a significant role for the sizes used here (size of the postage stamp was 39 pixels and of convolved object FWHM was 4.5 pixels).

For the fiducial galaxy model I find that the multiplicative bias on the shear is positive. In contrast, many methods tested on the STEP and GREAT08/10 simulation sets underestimate the shear (i.e. report a negative m). I note, however, that the characteristics of the noise bias will depend not only on the shear measurement method, but also on the distribution of true galaxy surface brightness profiles (as shown by the $F_b/(F_b + F_d)$ dependence in 3.3). Moreover, the results from the STEP and GREAT08/10 challenges are affected by other types of biases, such as underfitting and centroiding, which are not included here.

3.3.3 Parameters and functions used to create models of the bias on ellipticity and shear

In this section I list the parameters for the fitting functions which characterise the bias on the shear as a function of galaxy size, signal to noise ratio, PSF ellipticity and galaxy morphology.

Table 3.3 contains the multiplicative and additive bias measurements for all galaxies used in this work. See Appendix 3.3.1 for details of how these values were calculated. Fiducial galaxy parameters were: SNR = 20, $R_{gp}/R_p = 1.6$, $\mathrm{FWHM_{PSF}} = 2.85$, $e_{\mathrm{PSF}} =$

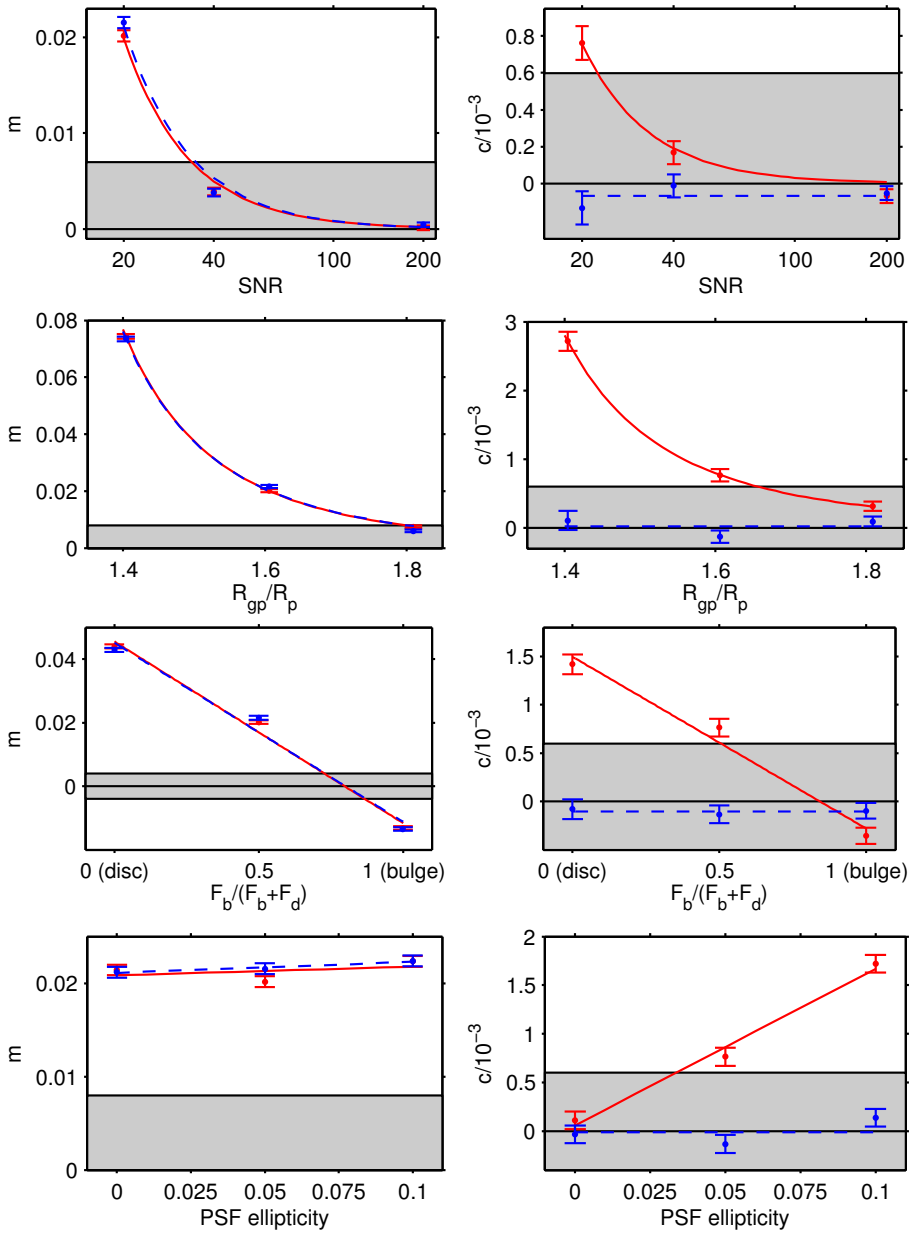

Figure 3.3. Multiplicative (left column) and additive (right column) bias as a function of galaxy and image parameters at intrinsic ellipticity of $|e| = 0.3$. First and second ellipticity components are marked with red and blue dashed, respectively. Note that on some of the plots the errorbars are too small to be visible. Typical standard error on the multiplicative bias was of order (5 - 10) $\cdot 10^{-4}$ and on additive bias of order (5 - 10) $\cdot 10^{-5}$. Lines are fits to the measured points, not the theoretical prediction. m_1, m_2 and c_1 as a function of SNR were fitted with SNR^{-2} function, c_2 with a constant. For m_1, m_2 and c_1 vs R_{gp}/R_p the basis expansion for the fit was $\{(R_{gp}/R_p)^{-2}, (R_{gp}/R_p)^{-3}\}$, for c_2 - a constant. For the other parameters a linear fit was used. Section 3.3.3 contains the data points (Table 3.3) and equations for fitted functions (Table 3.4). The grey shaded area corresponds to requirements for upcoming surveys.

	m_1	m_1	c_1	c_2
fiducial	$+0.01962 \pm 0.00040$	$+0.02094 \pm 0.00044$	$+0.00084 \pm 0.00040$	-0.00006 ± 0.00008
SNR $= 200$	$+0.00010 \pm 0.00017$	$+0.00021 \pm 0.00018$	-0.00007 ± 0.00017	-0.00003 ± 0.00004
SNR $= 40$	$+0.00383 \pm 0.00028$	$+0.00416 \pm 0.00031$	$+0.00016 \pm 0.00028$	-0.00009 ± 0.00006
$R_{gp}/R_p = 1.4$	$+0.05809 \pm 0.00060$	$+0.05263 \pm 0.00067$	$+0.00274 \pm 0.00060$	-0.00003 ± 0.00013
$R_{gp}/R_p = 1.8$	$+0.00809 \pm 0.00031$	$+0.00688 \pm 0.00034$	$+0.00037 \pm 0.00031$	$+0.00009 \pm 0.00007$
disc	$+0.03992 \pm 0.00045$	$+0.03929 \pm 0.00050$	$+0.00166 \pm 0.00045$	-0.00002 ± 0.00010
bulge	-0.01325 ± 0.00036	-0.01171 ± 0.00040	-0.00036 ± 0.00036	-0.00006 ± 0.00008
$e_{\rm PSF} = \{0.0, 0.0\}$	$+0.02050 \pm 0.00040$	$+0.02067 \pm 0.00044$	$+0.00010 \pm 0.00040$	-0.00005 ± 0.00008
$e_{\rm PSF} = \{0.1, 0.0\}$	$+0.02176 \pm 0.00040$	$+0.02119 \pm 0.00044$	$+0.00195 \pm 0.00040$	$+0.00006 \pm 0.00009$
$R_{gp}/R_p = 1.8$ disc	$+0.01740 \pm 0.00035$	$+0.01785 \pm 0.00039$	$+0.00089 \pm 0.00035$	-0.00006 ± 0.00007
$R_{gp}/R_p = 1.8$ bulge	-0.02694 ± 0.00031	-0.02101 ± 0.00034	-0.00066 ± 0.00031	-0.00012 ± 0.00007
$R_{gp}/R_p = 1.4$ disc	$+0.05899 \pm 0.00063$	$+0.05634 \pm 0.00070$	$+0.00289 \pm 0.00063$	-0.00003 ± 0.00013
$R_{gp}/R_p = 1.4$ bulge	$+0.03450 \pm 0.00054$	$+0.03459 \pm 0.00060$	$+0.00161 \pm 0.00054$	$+0.00016 \pm 0.00012$

Table 3.3. Measured multiplicative and additive biases for all simulated galaxies. Biases here are shown for a ring test using intrinsic ellipticity of 0.3. All parameters of the galaxies were the same as the fiducial model, except the ones indicated in the first column.

	m_1 m_2	c_1 c_2
$D_1 := \mathrm{SNR}$	$+7.956 \cdot 10^{+00} \cdot D_1^{-2}$ $+8.470 \cdot 10^{+00} \cdot D_1^{-2}$	$+3.026 \cdot 10^{-01} \cdot D_1^{-2}$ $-6.685 \cdot 10^{-05}$
$D_2 := R_{gp}/R_p - 1$	$-2.190 \cdot 10^{-03} \cdot D_2^{-2} + 5.791 \cdot 10^{-03} \cdot D_2^{-3}$ $-1.923 \cdot 10^{-03} \cdot D_2^{-2} + 5.639 \cdot 10^{-03} \cdot D_2^{-3}$	$-4.002 \cdot 10^{-05} \cdot D_2^{-2} + 1.953 \cdot 10^{-04} \cdot D_2^{-3}$ $+2.089 \cdot 10^{-05}$
$D_3 := \frac{F_b}{F_b+F_d}$	$-5.716 \cdot 10^{-02} + 4.557 \cdot 10^{-02} \cdot D_3$ $-5.641 \cdot 10^{-02} + 4.518 \cdot 10^{-02} \cdot D_3$	$-1.775 \cdot 10^{-03} + 1.496 \cdot 10^{-03} \cdot D_3$ $-1.034 \cdot 10^{-04}$
$D_4 := e_{PSF}$	$+2.084 \cdot 10^{-02} + 9.193 \cdot 10^{-03} \cdot D_4$ $+2.111 \cdot 10^{-02} + 1.185 \cdot 10^{-02} \cdot D_4$	$+5.697 \cdot 10^{-05} + 1.612 \cdot 10^{-02} \cdot D_4$ $-1.107 \cdot 10^{-05}$

Table 3.4. Equations for noise bias model function. These are the equations fitted to the data points in Figure 3.3.

	$a_1^{(0)}$	$a_1^{(1)}$	$a_1^{(2)}$	$a_1^{(3)}$	$a_1^{(4)}$	$a_2^{(0)}$	$a_2^{(1)}$	$a_2^{(2)}$	$a_2^{(3)}$	$a_2^{(4)}$
fiducial	+0.0008	+0.0201	+0.0003	+0.0004	−0.0013	−0.0001	+0.0216	+0.0006	−0.0047	−0.0070
SNR = 200	−0.0001	+0.0001	+0.0004	−0.0009	+0.0004	−0.0001	+0.0004	+0.0002	−0.0009	−0.0007
SNR = 40	+0.0002	+0.0039	+0.0004	−0.0013	+0.0013	−0.0000	+0.0037	+0.0001	+0.0047	+0.0021
$R_{gp}/R_p = 1.4$	+0.0027	+0.0767	−0.0076	−0.1154	−0.1075	+0.0001	+0.0754	−0.0031	−0.1048	−0.1125
$R_{gp}/R_p = 1.8$	+0.0003	+0.0073	+0.0008	+0.0055	+0.0054	+0.0001	+0.0059	−0.0003	+0.0073	+0.0066
disc	+0.0014	+0.0443	+0.0004	−0.0282	−0.0256	−0.0001	+0.0432	+0.0004	−0.0247	−0.0254
bulge	−0.0004	−0.0132	+0.0007	−0.0024	−0.0000	−0.0001	−0.0137	+0.0002	−0.0023	+0.0133
$e_{PSF} = \{0.0, 0.0\}$	+0.0001	+0.0216	−0.0008	−0.0072	−0.0042	−0.0000	+0.0213	−0.0009	−0.0046	−0.0048
$e_{PSF} = \{0.1, 0.0\}$	+0.0017	+0.0223	+0.0014	−0.0056	−0.0039	+0.0001	+0.0227	−0.0015	−0.0006	−0.0107
$R_{gp}/R_p = 1.8$, disc	+0.0008	+0.0172	+0.0010	−0.0008	+0.0049	−0.0000	+0.0173	−0.0002	+0.0052	+0.0016
$R_{gp}/R_p = 1.8$, bulge	−0.0006	−0.0290	+0.0003	+0.0030	+0.0140	−0.0001	−0.0230	−0.0004	+0.0020	+0.0108
$R_{gp}/R_p = 1.4$, disc	+0.0030	+0.0843	−0.0110	−0.1420	−0.1485	+0.0001	+0.0844	−0.0010	−0.1540	−0.1425
$R_{gp}/R_p = 1.4$, bulge	+0.0014	+0.0425	−0.0012	−0.0500	−0.0468	+0.0001	+0.0450	−0.0004	−0.0593	−0.0529

Table 3.5. Parameters of equations for the bias on ellipticity. These are the parameters used with Eq. 3.34.

$\{0.05, 0\}$, $\beta^{\text{Moffat}} = 3$, $\text{flux}_{\text{bulge}}/\text{flux}_{\text{total}} = 0.5$, $r_{\text{bulge}}/r_{\text{disc}} = 1.0$. Table 3.4 contains equations of the functions in Figure 3.3. Table 3.5 contains the parameters of polynomial function fitted to the bias on ellipticity, for example in Figure 3.2. The equation used with these parameters is

$$b[\hat{e}_1] = a_1^{(0)} + a_1^{(1)}\hat{e}_1 + a_1^{(2)}\hat{e}_1^2 + a_1^{(3)}\hat{e}_1^2\hat{e}_2 + a_1^{(4)}\hat{e}_1^3, \tag{3.34}$$

accordingly for $b[\hat{e}_2]$ with parameters a_2.

3.4 Noise bias calibration

In this section I investigate how the bias measurements can be used to calibrate out the noise bias effect. First, I create a model of the bias on the ellipticity measurement as a function of four measured parameters: $\hat{e}_1$, $\hat{e}_2$, $\widehat{Rgp/Rp}$, $F_b/\widehat{(F_b + F_d)}$, similar to Figure 3.2 (note that I do not directly use the functions presented on Figure 3.3, as they show a bias on shear in the form of m and c, instead of the bias on the ellipticity). I apply an additive correction predicted by our model directly to the measured ellipticity values. Finally I verify the accuracy of this procedure by testing it using a ring test consisting of 10 million noisy fiducial galaxies.

This approach will not provide a perfect calibration, as our model of biases is calculated for a set of galaxies with particular true galaxy and image properties. In practice I will only know the measured galaxy parameters, which are noisy, as illustrated in Figure 3.1. Therefore, if I read off the bias values from the measurements of the noisy measured galaxy parameters they will not be exactly the correct bias values for that galaxy. In this section, I investigate the scale of this effect.

The estimator of the ellipticity $\hat{e}$ is biased, so that $\hat{e} = \tilde{e} + b[\hat{e}]$, where $\tilde{e}$ is the unbiased estimator. By definition $\tilde{e}$ averaged over noise realisations is equal to the true ellipticity, so that $\langle\tilde{e}\rangle = e^{\text{true}}$.

I estimate the true shear g with an estimator $\hat{g}$ in a ring test. I write the following equations to show mathematically what is happening when I do the correction on the

individual galaxy ellipticities.

$$\hat{g} = \langle\langle\widehat{e^l}\rangle_N\rangle_R = \langle\langle\widetilde{e^l}\rangle_N + b[\widehat{e^l}]\rangle_R \tag{3.35}$$

$$= g + \langle b[\widehat{e^l}]\rangle_R \tag{3.36}$$

where $e^l = e + g$ is the lensed ellipticity, and subscripts N and R denote averages over noise realisations and around the ring respectively. Eq. 3.36 shows that the bias of the shear estimator will be equal to the bias on the lensed ellipticity $e+g$, averaged over noise realisations and the ring. This is the bias I aim to calibrate.

I create a correction model which describes $b[\hat{e}]$ as a function of four galaxy parameters, i.e.

$$b[\hat{e}] = \beta(\theta) = \beta(e_1, e_2, R_{gp}/R_p, F_b/(F_b + F_d)) \tag{3.37}$$

Then I apply this correction to the noisy estimates $\hat{\theta}$, creating an estimator of the correction $\beta(\hat{\theta})$ and I update our ellipticity estimate to be

$$\hat{e}^\beta \leftarrow \hat{e} - \beta(\hat{\theta}). \tag{3.38}$$

Using this correction in the ring test implies

$$\hat{g}^\beta = g + \langle b[\widehat{e+g}] - \langle\beta(\hat{\theta})\rangle_N\rangle_R. \tag{3.39}$$

Because I am applying the correction to the noisy maximum likelihood estimates, the correction itself can be biased under noise, so that $b[\beta(\hat{e}, \dots)] = \langle\beta(\hat{\theta})\rangle - b[\hat{e}]$. Including this 'bias on the correction', I expect the the final bias on the shear after applying our calibration procedure to be

$$b[\hat{g}^\beta] = \langle\langle b[\beta(\widehat{e+g})]\rangle_N\rangle_R \tag{3.40}$$

$$c^\beta = \langle\langle b[\beta(\hat{e})]\rangle_N\rangle_R \tag{3.41}$$

$$m^\beta = \frac{\langle\langle b[\beta(\widehat{e+g})]\rangle_N\rangle_R - \langle\langle b[\beta(\hat{e})]\rangle_N\rangle_R}{g}. \tag{3.42}$$

Testing this procedure will include finding out how big the term in Eq. 3.41 is.

In practice I create the model of the bias $\beta(\theta)$ (Eq. 3.37) using a learning algorithm

based on Radial Basis Functions (RBF) Interpolation[6], trained on all our simulated results. This learning algorithm performed a regression task, and interpolated the measured points in the θ space predicting the value of β for any new parameter vector. This algorithm used a nonparametric kernel method. Then I use Eq. 3.38 to correct the ellipticity estimates.

The calibration procedure was tested by generating nearly ten million galaxy images using the default galaxy parameters. The ring test was performed as follows: a set of galaxies was simulated with the galaxy intrinsic ellipticity angles equally spaced at 16 values from 0 to π, (i) with no shear applied (ii) with a shear of $g_1 = 0.1$ applied. In total 300,000 galaxies were simulated at each angle in the ring, for each shear value. To compute the uncalibrated shear measurement bias, the measured ellipticity was averaged over all galaxies with a given shear to obtain a shear estimate for that population. Then a straight line was fitted to the resulting shear estimates as a function of input shear to obtain the usual m and c. To compute the calibrated shear measurement bias, the measured ellipticities were corrected using Eq. 3.38 before averaging to obtain the shear estimate.

The uncalibrated and calibrated shear measurement biases are presented in Figure 3.4. I see that the uncalibrated shear measurement biases are well outside the requirement for upcoming surveys, as discussed earlier. The calibration reduces the additive bias by a factor of around three, and the multiplicative bias by a factor of around ten. I find that the bias term in Eq. 3.41 is insignificantly small to the accuracy afforded by our simulations. Therefore the calibrated biases are now within the requirement for upcoming surveys for both additive and multiplicative shear biases.

3.5 Conclusions

In this chapter I have investigated the effect of noise on shear measurement from galaxy images. I have found that this can significantly bias shear measurement from realistic images, even though the bias goes away completely for images with lower noise levels. This was previously studied in (Hirata et al. 2004) and R12, who demonstrated the existence of this noise bias effect. I quantified noise bias using images simulated from more realistic galaxy models and I used a forward fitting shear measurement method which fitted a matching set of galaxy models to the simulations (IM3SHAPE, Zuntz et al. (2013)). These

[6] http://www.mathworks.com/matlabcentral/fileexchange/10056

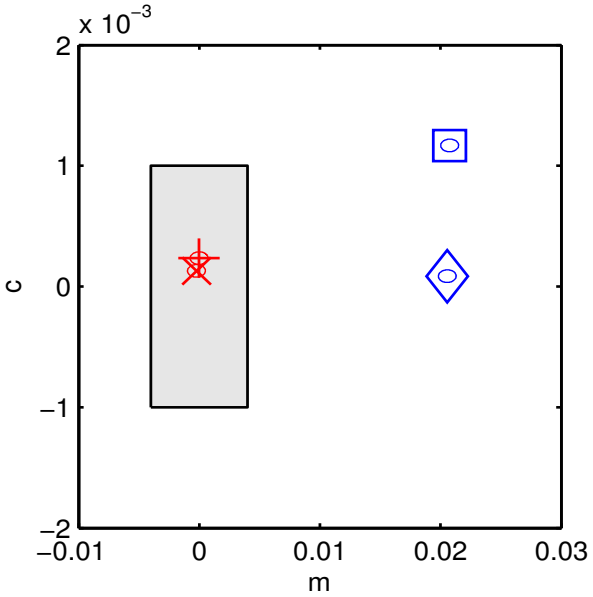

Figure 3.4. Values of multiplicative (m) and additive (c) bias for uncalibrated (blue) and calibrated (red) shear estimates. Ellipses indicate one sigma error bars.

models are based on observationally-motivated combinations of exponential disk and de Vaucouleurs bulge models and are broadly representative of the light profiles of realistic galaxies. They have also formed the basis of previous weak lensing simulation programmes (Heymans et al. 2006; Bridle et al. 2009*a*; Kitching et al. 2013). I use a maximum likelihood estimator (MLE) to obtain galaxy ellipticity estimates from the images, and use these ellipticity estimates as our noisy shear estimates. I find that the shear measurement biases often exceed ∼1% and even approach ∼10% for the smallest galaxies and highest noise values I consider in this chapter.

One feature of the simulations presented is that they are deliberately internal: test galaxies are generated using the same models and routines used later for fitting them, the only difference being the addition of noise. In this way I are able to explore the effects of noise biases in isolation from the contribution of underfitting or model bias (e.g. Melchior et al. 2009; Voigt & Bridle 2009; Bernstein 2010). The fact that the biases I detect are considerable, even when fitting with perfect knowledge of the parametric galaxy model, is striking. I conclude that, for many methods, bias from unavoidable noise in galaxy images must be considered an important potential source of systematic error when seeking shear inference at sub-percent level accuracy. The existence of noise bias is likely to be a common feature to many shape measurement methods (Hirata et al. 2004; R12). Unless shape measurement methods are theoretically constructed to avoid noise bias, empirical

calibration with simulations is necessary.

I quantified the noise bias as a function of image and galaxy parameters and found a strong dependence. I found that the dependence on image signal-to-noise ratio is inverse square, as expected from symmetry arguments (e.g. see R12). The dependence on galaxy size is quite non-linear and rises steeply as the galaxy size decreases relative to the PSF size. The bias depends on the galaxy profile in a complicated way. I find that for our fiducial parameters shears are overestimated for exponential disc galaxies and underestimated for de Vaucouleurs bulge galaxies. The dependence on bulge to total flux ratio is reasonably consistent with a linear relation. There is a good linear relation between the additive shear measurement noise bias and the PSF ellipticity.

Many shape measurement methods are potentially subject to noise bias, and for these methods this sort of calibration will be an important step in order to reduce systematic errors below the level required for upcoming survey datasets. I illustrate a correction scheme based on a model of the measured biases, as a function of observed galaxy properties. Note that this is not expected to remove the bias completely because the observed galaxy properties are not the true galaxy properties and therefore I will be using slightly the wrong bias correction. This correction was able to reduce ellipticity estimator biases to lower levels than those required for the upcoming lensing surveys, for a fiducial galaxy with SNR=20 and a typical intrinsic ellipticity of magnitude 0.3.

There is a small residual bias remaining after this first level of correction. This is due to the scatter and bias in measured galaxy parameters about their true values. This scatter and bias is an output of the simulations and could therefore be propagated into a second level of bias correction which would reduce the residual bias yet further, into the realm of far-future surveys.

The calibration scheme I proposed can only be applied to a method which, in addition to ellipticity, also produces estimates of other parameters; it will probably be difficult to use it with a method such as KSB, which primarily aims to estimate only the ellipticity parameters.

This calibration approach is extremely computationally expensive and would ideally be carried out for a large range and sampling of image and galaxy parameters. The resolution of our results was limited by the available computing time. The final results shown in this chapter took over 1 year of CPU time.

These results are obtained using the same two-component co-elliptical galaxy model

in the simulations and in the fits. In practice it will be necessary to investigate more complicated galaxy morphologies, which may not be precisely modelled in the fits. See (Zuntz et al. 2013) for noise bias calibration applied to GREAT08 data.

For future surveys the simulated data must be carefully constructed in order to recreate realistic observing conditions, and the realistic properties of the underlying galaxies (the latter requirement poses greater difficulties than the former). The deep imaging of the real sky is potentially an expensive overhead for future surveys, but may prove necessary for confidence in the final results. Accurate estimates of gravitational shear from methods affected by noise bias will rely on consistent strategies for measuring and correcting these systematic effects.

The presented calibration scheme does not use the information about the galaxy parameters distribution in the universe. I found that the measured galaxy parameters were a sufficiently good proxy for the true galaxy parameters that the noise bias could be corrected well enough for upcoming surveys. If this result were generally true then this places less stringent requirements on the simulations because the galaxy population demographics would not need to match exactly with reality, and the simulations would only have to span a realistic range of galaxy parameters. However, different calibration schemes could be created based on the distributions of galaxy parameters. The simplest solution would be to calculate one m and c for the whole population of galaxies, randomly drawing not only noise maps but also galaxy and image parameters from histograms of measured parameters from galaxies in the survey. Using this method is not limited to maximum likelihood fitting; potentially all shear measurements methods could be calibrated that way.

I have used a white Gaussian noise model. In general it should be possible to repeat this procedure for a case of correlated noise. It should also be possible to repeat the procedure for Poisson noise. Our bias results will also depend on the number of parameters used in the fitting. I have used seven free parameters and fixed the ratio of radii of the bulge and disc galaxy components to unity. I also assumed no constant background in the image, whereas this could also be included as a free parameter in the fit. An uncertain variable background level would complicate the analysis further.

Another approach would be to use a fully Bayesian analysis: use the full likelihood distribution (or samples) of ellipticity given the noisy images and propagate this uncertainty to the cosmological parameters. In this case the calibration would not be necessary.

Chapter 4

Sérsic galaxy models in weak lensing shape measurement: model bias, noise bias and their interaction

Cosmic shear is a powerful probe of cosmological parameters, but its potential can be fully utilised only if galaxy shapes are measured with great accuracy. Two major effects have been identified which are likely to account for most of the bias seen for maximum likelihood methods in recent shear measurement challenges. Model bias occurs when the true galaxy shape is not well represented by the fitted model. Noise bias occurs due to the non-linear relationship between image pixels and galaxy shape. In this chapter I investigate the potential interplay between these two effects when an imperfect model is used in the presence of high noise. I present analytical expressions for this bias, which depends on the residual difference between the model and real data. They can lead to biases not accounted for in previous calibration schemes.

By measuring the model bias, noise bias and their interaction, I provide a complete statistical framework for measuring galaxy shapes with model fitting methods from GRavitational lEnsing Accuracy Testing (GREAT)-like images. I demonstrate the noise and model interaction bias using a simple toy model, which indicates that this effect can potentially be

significant. Using real galaxy images from the Cosmological Evolution Survey (COSMOS) I quantify the strength of the model bias, noise bias and their interaction. I find that the interaction term is often a similar size to the model bias term, and is smaller than the requirements of current and near future galaxy surveys.

4.1 Introduction

Weak gravitational lensing is a very important and promising probe of cosmology (see Schneider 1995; Bartelmann & Schneider 1999; Hoekstra & Jain 2008 for reviews). Matter between a distant galaxy and an observer causes the image of the galaxy to be distorted. This distortion is called gravitational lensing. Almost all distant galaxies I observe are lensed, mostly only very slightly, so that the observed image is sheared by just a few percent. Weak gravitational lensing shear is particularly effective in constraining cosmological model parameters (Albrecht et al. 2006; Peacock et al. 2006; Fu et al. 2007; Kilbinger et al. 2013). Measuring the spatial correlations of those shear maps in the tomographic bins of redshift can shed light on the evolution of dark energy in time (Hu 2002; Takada & Jain 2005; Huff et al. 2011; Heymans et al. 2013; Benjamin et al. 2012) and modified gravity (Simpson et al. 2012; Kirk et al. 2013).

Several projects are planning to measure cosmic shear using optical imaging. The KIlo-Degree Survey (KIDS)[1], the Dark Energy Survey (DES)[2], the Hyper Suprime-Cam (HSC) survey[3], the Large Synoptic Survey Telescope (LSST)[4], Euclid[5] and Wide Field Infrared Survey Telescope (WFIRST)[6].

However, accurate measurement of cosmic shear has proved to be a challenging task (Heymans et al. 2006; Massey et al. 2006; Bridle et al. 2009*a*; Kitching et al. 2013). There is a range of systematic effects that can mimic a shear signal. In this chapter, I focus on biases in the measurement of shear. Other important systematics are: intrinsic alignments of galaxy ellipticities, photometric redshift estimates and modelling of the clustering of matter on small scales in the presence of baryons.

Prior to shearing by large scale structure, galaxies are already intrinsically elliptical.

[1] http://kids.strw.leidenuniv.nl/
[2] http://www.darkenergysurvey.org
[3] http://www.naoj.org/Projects/HSC/HSCProject.html
[4] http://www.lsst.org
[5] http://sci.esa.int/euclid
[6] http://exep.jpl.nasa.gov/programElements/wfirst/

This ellipticity is oriented randomly on the sky in the absence of intrinsic alignments. Then a few percent change in this ellipticity is induced as the light travels from the galaxy to the observer through intervening matter. During the observation process, the images are further distorted by a telescope Point Spread Function (PSF) and, in case of ground-based observations, also by atmosphere turbulence. Additionally the image is pixelised by the detectors. Due to the finite number of photons arriving on the detector during an exposure, the galaxy images are noisy. Additional noise is induced by the CCD readout process in the detector hardware.

The complexity of this forward process makes the unbiased measurement of the shear signal very challenging. Bonnet & Mellier (1995) showed that for very good quality data, simple quadrupole moments of the images can be used as unbiased shear estimators. Example methods utilising this approach are Kaiser et al. (1994); Kaiser (1999); Hirata & Seljak (2003); Okura & Futamase (2010). For PSF convolved galaxy images, one can use DEconvolution In MOment Space (DEIMOS Melchior et al. 2012) to remove the effects of the PSF from the quadrupole.

Model fitting methods use a parametric model for the galaxy image to create a likelihood function (often multiplied by the prior function) in all parameters, from which ellipticity estimators are then extracted. Galaxy images are often modelled by Sérsic functions: IM3SHAPE (Zuntz et al. 2013) uses a 7-parameter bulge + disc model and infers the parameter values by maximum likelihood estimation. Miller et al. (2007, 2012) uses a similar model, and mean posterior for the estimator. Another frequently used galaxy model is a decomposition into a Gauss-Laguerre orthogonal set (Refregier 2001; Bernstein & Jarvis 2001; Nakajima & Bernstein 2006), 'shapelets'. The complexity of this model can be controlled by changing the number of coefficients in the shapelet expansion.

In the context of model fitting methods, if the model is not able to represent realistic galaxy morphologies well enough, then the shape estimator will be biased (Bernstein 2010; Voigt & Bridle 2009). This bias is often called the *model bias* or *underfitting bias*.

In the presence of pixel noise on the image, the unweighted quadrupole moment will give an unbiased estimate of the quadrupole, with a very large variance. However, the shear is defined as a ratio of quadrupole moments, and Hirata & Seljak (2003); Melchior & Viola (2012); Okura & Futamase (2012) showed that this induced non-linearity leads to a bias. This bias is often called the *noise bias*.

For model fitting methods, the noise related bias was studied in Refregier et al. (2012

hereafter R12), where analytical expressions were given for the estimator bias. It presented the noise bias as a simple statistical problem, assuming that the true galaxy model was perfectly known. The noise was added to images which were created by the same model function which then later was used to fit it. In (Kacprzak et al. 2012 hereafter K12) I have used the Sérsic galaxy models to quantify the magnitude of noise bias and presented a calibration scheme to correct for it. Again, here the true galaxy model was perfectly known. In Zuntz et al. (2013) I have further developed the calibration scheme. It was then applied to the GREAT08 simulation set (Bridle et al. 2009*b*,*a*), with satisfactory results on most challenge branches.

In this chapter I investigate the next piece in the puzzle: what if model biases and noise biases occur at the same time? This will certainly be the case for real galaxy surveys: the galaxies have realistic morphology and the images are noisy. In that situation there may be some interaction between noise and model bias.

Figure 4.1 demonstrates this concept. Is shows two effects studied so far:

- *model bias* - when the galaxy image is noiseless, and the fitted model does not represent the complicated galaxy features well
- *noise bias* - when galaxy image is noisy, and the fitted model is perfectly representing the galaxy

and the effect I investigate more in this work:

- *model bias, noise bias and their interaction* - when the true galaxy has realistic morphology, the fitted model does not represent the galaxy well, and the observed image is noisy.

I present analytic equations for the noise bias when the true galaxy model is not perfectly known. Then I use a toy model to show that the interaction terms have the potential to be significant. To evaluate the significance of this interaction terms I use 26113 real galaxy images from the COSMOS survey (Mandelbaum et al. 2011), available in the GalSim[7] toolkit (RoI et al., in prep.).

Therefore, by evaluating the model bias, noise bias and their interaction, I attempt to answer the question: *can I use Sérsic profiles to represent galaxies in fitting to realistic noisy data?* However, our answer is limited to a simple bulge + disc galaxy model, which

[7] `github.com/GalSim-developers/GalSim`

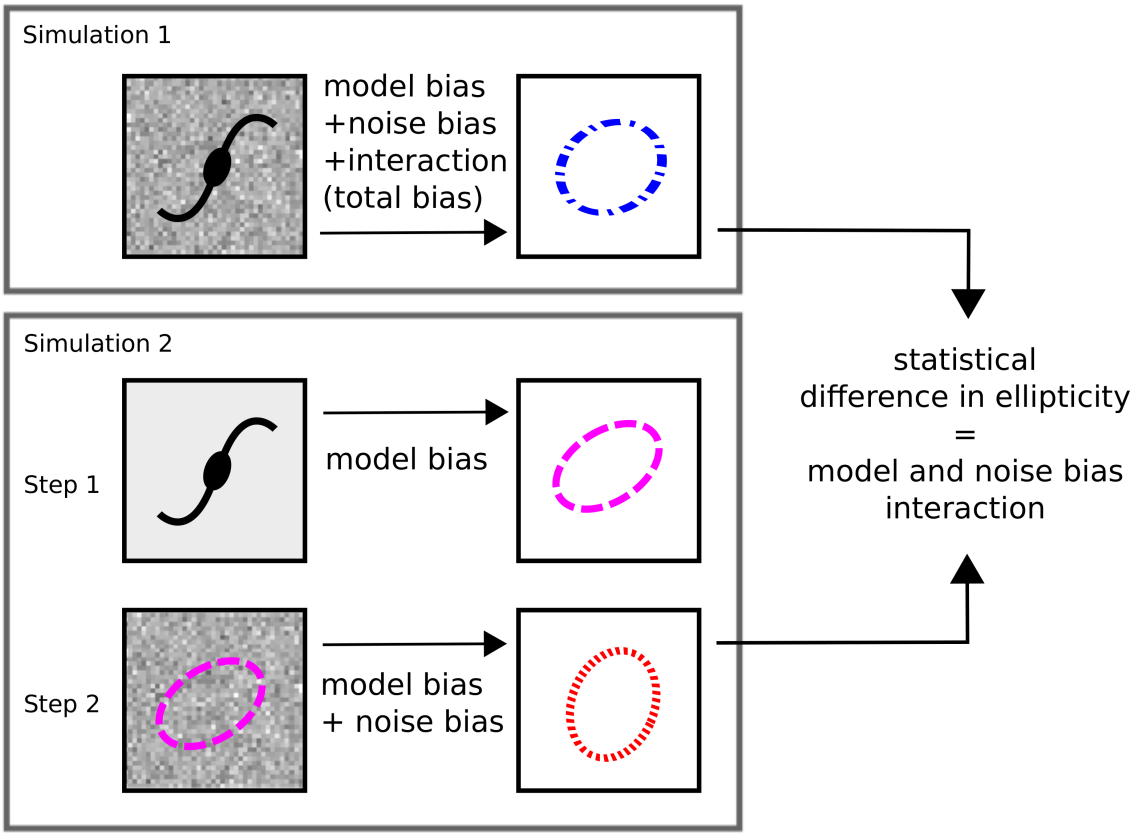

Figure 4.1. A demonstration of the concept of noise and model bias interaction. Left-hand side postage stamps show the images which are going to be fit with a parametric galaxy model. Right-hand side postage stamps represent the images of best fitting models. In this work I compare the results of these two simulations. Simulation 1 uses the real galaxy image directly to measure noise bias, model bias and their interaction jointly in a single step (blue dot-dashed ellipse). Simulation 2 first finds the best fitting parametric model to a real galaxy image (black solid galaxy cartoon), and create its model image (dashed magenta ellipse). This process introduces the model bias. The next step is to measure the noise bias using this best fitting image as the true image. The fit to the noisy image is represented by red dotted ellipse. In this simulation noise and model bias interaction terms are absent. The difference of the results of Simulations 1 and 2 measures the strength of noise and model bias interaction terms.

is commonly used in model fitting shape measurement methods for weak lensing. Also, I use the PSF and pixel size that corresponds to the parameters of a conservative upcoming ground-based Stage III survey. It is possible to perform similar analysis for different galaxy models and telescope parameters, as well as for different types of estimator (maximum posterior, mean posterior).

Calibration of biases in shear measurement is becoming an increasingly important part of shear pipelines. Miller et al. (2012) used a calibration for noise bias as a function of galaxy size and signal to noise ratio. Systematic effects may depend on many different galaxy and image quality properties, and increasing the accuracy of the bias measurement will require including more of these effects into the pipeline. This will increase the complexity of the modelling and the computational cost of the simulation. Moreover, it will require knowledge of the true underlying parameters for a galaxy sample, either in the form of a representative calibration sample of images, or an inferred galaxy parameters distribution. Refregier & Amara (2014) introduced a procedure to infer these parameters via a Monte Carlo Control Loops approach. So far, in various simulations, galaxy images were often created using a parametric Sérsic model function. An important question is: are the realistic galaxy morphologies important for shear bias measurement? The upcoming GREAT3 challenge (Mandelbaum et al. 2013, in prep) is planning to answer this question by testing using images of galaxies in the COSMOS survey. In this chapter I provide an initial answer to this question. A more detailed study of galaxy model selection is going to be presented in (Voigt et al. 2013, in prep.) and calibration sample requirements in (Hirsch et al. 2013 in prep.).

It is worth noting that some shear measurement methods apply a fully Bayesian formalism and use a full posterior shear probability (Bernstein & Armstrong 2013) in subsequent analyses. These methods are considered to be free of noise and model related biases and present a promising alternative approach to problems studied in this chapter.

This chapter is organised as follows. Section 4.2 contains the principles of cosmic shear analysis. Section 3.2.3 presented the analytic formulae for noise and model bias interaction, as a generalisation of the noise bias equations derived in R12. I present a toy model for the problem in section 4.3. In Section 4.4 I use the COSMOS sample to evaluate the noise and model biases, and show the significance of the interaction terms. I conclude in Section 4.5.

4.2 Systematic errors in model fitting

In this section, first I present the basics of the model fitting approach to shear measurement and then discuss the biases it introduces. I discuss the requirements on this bias in the context of current and future surveys. Then I introduce analytical expressions for the noise bias, including interaction terms with the model bias.

4.2.1 Shear and ellipticity

Cosmic shear as a cosmological observable can be related to the gravitational potential between the distant source galaxy and an observer (for a review see Bernstein & Jarvis 2001). Ellipticity is defined as a complex number

$$e = \frac{a-b}{a+b}\, e^{2i\phi}, \tag{4.1}$$

where a and b are the semi-major and semi-minor axes, respectively, and ϕ is the angle (measured anticlockwise) between the x-axis and the major axis of the ellipse. The observed (lensed) galaxy ellipticity is modified by the complex shear $g = g_1 + ig_2$ in the following way

$$e^l = \frac{e^i + g}{1 + g^* e^i}. \tag{4.2}$$

In the absence of intrinsic alignments the lensed galaxy ellipticity is an unbiased shear estimator (Seitz & Schneider 1996).

To recover this ellipticity, the shear measurement methods correct for the PSF and the pixel noise effects. This procedure can introduce a bias. Usually shear bias is parametrised with a multiplicative m and additive c component

$$\hat{\gamma}_j = (1 + m_j)\gamma_j^t + c_j, \tag{4.3}$$

where $\hat{\gamma}$ is the estimated shear and γ_j^t is the true shear. Additive shear bias is usually highly dependent on the PSF ellipticity and can be probed using the star-galaxy correlation function (e.g. Miller et al. 2012), which can be measured from real data. Efforts are made to calibrate it using simulations (Miller et al. 2012 K12). Requirements for multiplicative and additive bias are summarised in Table 4.1, derived from Amara & Refregier (2007).

Survey	area (sq deg)	m_i	c_i
Current	200	0.02	0.001
Upcoming future	5000	0.004	0.0006
Far future	20000	0.001	0.0003

Table 4.1. Requirements for the multiplicative and additive bias on the shear for current, upcoming and far future surveys, based on Amara & Refregier (2007)

.

When fitting a galaxy model with co-elliptical isophotes, this ellipticity is often amongst the model parameters. An estimator is created using a likelihood function, which in the presence of white Gaussian noise has the form

$$-2\log\mathcal{L} = \frac{1}{\sigma^2}\sum_p [g_p + n_p - f_p(\mathbf{a})]^2 \tag{4.4}$$

where p is the pixel index running from 1 to N, where N is the number of pixels, $\mathbf{a}$ is a set of variable model parameters, g_p is the noiseless galaxy image, n_p is additive noise with standard deviation σ_n and f_p is a model function.

In this approach, the two main systematic effects are *noise bias* and *model bias*. Model bias has so far been studied in the context of low noise (or noiseless) images. Lewis (2009); Voigt & Bridle (2009); Bernstein (2010) demonstrated that fitting a galaxy image with a model which consists of a basis set (or degrees of freedom) smaller than in the true galaxy image, then the model fitting method can be biased. This bias can even be larger than the current survey requirements. Voigt et al. (2013, in prep) quantify the model bias using real galaxy images from the COSMOS survey, for different galaxy models used in the fit.

Noise bias arises when the observed galaxy images contain pixel noise. K12 showed that for a galaxy with noise level of $S/N > 200$ noise bias is negligible. However, for galaxies with $S/N \approx 20$ it can introduce a multiplicative bias up to $m = 0.08$, which is large compared to our requirements for future surveys (see Table 4.1). This is a very important systematic effect that will have to be accounted for to utilise the full statistical power of a survey.

R12 showed analytic expressions for the noise bias in the case when the true model is perfectly known. The key idea was that the value of the parameter estimator obtained from the noisy data lies near the true value for this parameter, so an expansion can be used to calculate it. It turns out that the terms which depend on the first power of noise variance disappear when averaged over pixel noise realisations. Second order terms,

however, give rise to the noise bias. The bias was quantified for a simple Gaussian galaxy model using both analytical expressions and simulation, which were found to be in good agreement.

In section 3.2.3 I expanded the derivation in R12 to the case when the true galaxy model is not known.

4.3 Toy model for the problem

To demonstrate the noise and model bias interaction I create a simple toy example. Galaxy models used in this example will be single Sérsic profiles (Sérsic 1963), where surface brightness at image position $\mathbf{x}$:

$$I(\mathbf{x}) = A \exp(-k[(\mathbf{x} - \mathbf{x_0})^\intercal \mathbf{C}^{-1} (\mathbf{x} - \mathbf{x_0})]^{1/(2n)}) \tag{4.5}$$

where $\mathbf{x_0}$ is galaxy centeriod, $\mathbf{C}$ is galaxy covariance matrix (see e.g. Voigt & Bridle (2009) for relation to ellipticity), $k = 1.9992n - 0.3271$ and n is the Sérsic index. Therefore, the model has six parameters which are being fit: two centroid parameters, two ellipticity parameters, a scale radius and an amplitude parameter.

Instead of a real galaxy, I use a single Sérsic profile with a fixed index. Hereafter I will refer to that galaxy as the *real* galaxy, as it will serve as a real galaxy equivalent in our toy example, even though it is created using a Sérsic function. I fit it with another single Sérsic profile, but with a different index. In this section I will refer to it as the *galaxy model*. This way I introduce model bias - the real galaxy is created using a Sérsic profile with a different index than the model I fit.

The impact of the noise and model interaction bias terms introduced in Eqn. 3.31 can be measured in the following way (Fig. 4.1). I use the real galaxy, add noise to it, and measure the ellipticity by fitting the galaxy model. This process will include the noise bias, model bias and their interaction terms, which are dependent on the pixel residual between the best fit to the noiseless image and the real galaxy image (Eqn. **??**). This process is described in Simulation 1 in Fig 4.1. These residuals will be zero if instead of the real galaxy image I use the best-fitting galaxy model (Simulation 2 in Fig 4.1). This is the pure noise bias scenario, as described in Eqn. 3.32. To measure the model bias and noise bias jointly, I use a two step procedure. In step 1 I obtain the best fit image to the noise-free real galaxy. Thus I obtain the model bias measurement. In step 2 I use

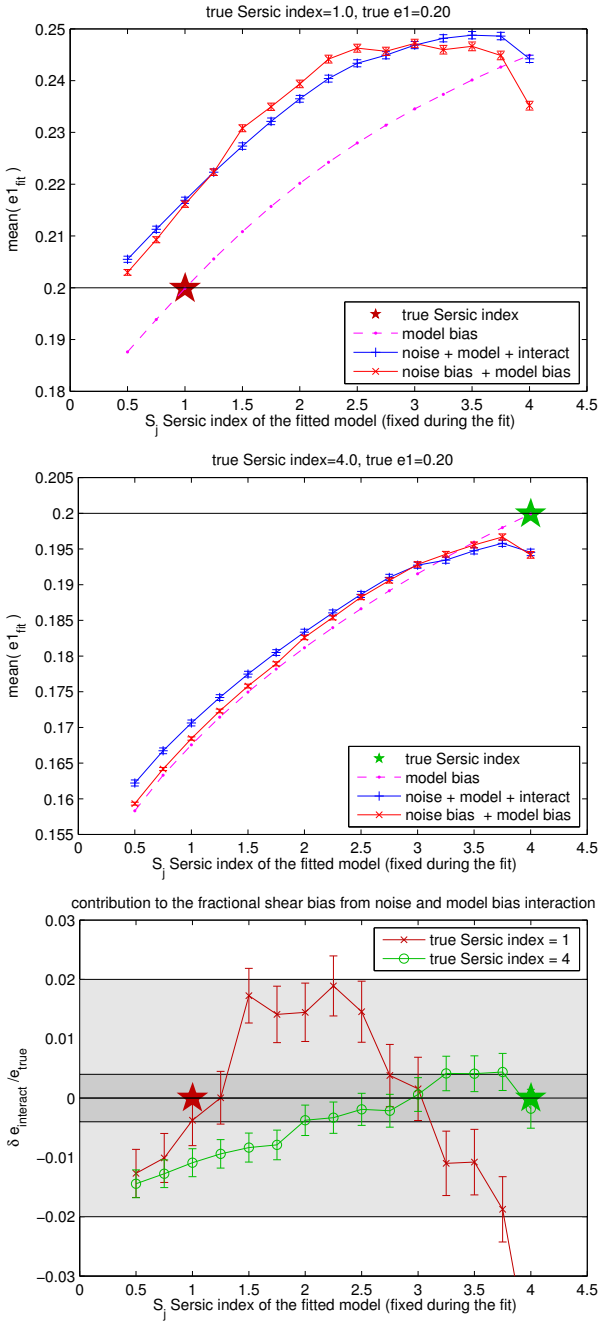

Figure 4.2. Noise and model bias toy model. The upper and middle panels show the galaxy ellipticity estimated when a wrong Sérsic index S_j is used. True ellipticity was $e_1^{true} = 0.2$. True Sérsic index was 1 and 4 for the top and middle panels respectively. Additionally, a star marks the true Sérsic index, to guide the eye. Magenta dash-dotted line with cross dot markers shows the model bias; ellipticity estimated by a fit with Sérsic S_j in the absence of noise. Blue and red lines with + and × markers, respectively, show the mean ellipticity measured from galaxy image with added noise, when: (1) galaxy image has the true Sérsic index, (2) galaxy image is the best fit with Sérsic index S_j. Bottom panel shows the fractional difference between $[(1) - (2)]/e_1^{true}$. Grey bands mark the requirement on multiplicative bias for current and upcoming surveys.

this model image to measure the mean of noise realisations by repeatedly adding noise to it. The strength of the noise and model bias interaction terms scaled by the residual r_p in Eqn. 3.32 can be measured by taking the difference between the mean of the noise realisations from the cases when the true image is a *real* galaxy and when the true image is the *best-fitting* galaxy, obtained earlier.

Fig. 4.2, upper and middle panels 2 show these cases when the real galaxy is a Sérsic with an index of 1 and 4, respectively. I used a Moffat PSF with FWHM = 2.85 pixels, and kept the half light radius of these galaxies so that the FWHM of the convolved object divided by the FWHM of the PSF is 1.4. These settings were used in the GREAT08 challenge and in previous noise bias work. The true ellipticity of the real galaxy was set to $e_1^{\rm true} = 0.2$. For the noisy cases, I used S/N = 20, where S/N = $\sqrt{\sum_{p=1}^{N} g_p^2}/\sigma_{\rm noise}$ (see Bridle et al. 2009*a*). This is an optimal, matched-filter S/N: it uses the profile which is assumed to be known in advance. In practice, it is not possible to window aperture or model-fitting S/N using the precisely known profile, and so real data S/N are always lower. They can commonly be lower by as much as a factor of 2: the SExtractor `FLUX_AUTO / FLUXERR_AUTO` often gives ~ 0.5 the reported S/N of the (Bridle et al. 2009*a*) definition (for typical SExtractor settings in simulations of deep optical astronomical data). Therefore the S/N minimum explored therefore corresponds approximately to what would be designated as ~ 10 in real data, such as CFHTLenS.

Fig. 4.2 shows the mean measured ellipticity $e_1^{\rm fit}$ as a function of the Sérsic index of the fitted model. The magenta dashed line shows the pure model bias measurement. There was no noise on the images in this case: the mean is taken from the average of simulated images, which had a centroid randomly sampled within a central pixel in the image. Standard errors for this measurement are so small that they are not visible on this plot. I see that if the correct model is used, then the measured ellipticity is equal to the true ellipticity $e_1^{\rm true} = 0.2$. This point is additionally marked with a star. For the real galaxy case with a Sérsic index of 4 (middle panel), ellipticity is underestimated when the index of the fit is smaller than 4. The same applies to the case when the Sérsic index of the real galaxy is 1. For that case, ellipticity is overestimated if the fitted index is greater than 1.

The mean of the noisy realisations for the noise bias + model bias + their interaction is marked by a blue solid line with plus + markers. For this case I were adding noise to the real galaxy images (Sérsic profiles with indices 1 for the top panel and 4 for the middle

true image	m_1	m_2	$std(m_1)$	$std(m_2)$	c_1	c_2	$std(c_1)$	$std(c_2)$
real galaxies	0.0238	0.0227	0.0006	0.0006	-0.0019	-0.0018	0.0001	0.0001
best-fitting galaxies	0.0230	0.0218	0.0007	0.0006	-0.0017	-0.0017	0.0001	0.0001
difference	0.0008	0.0009	0.0009	0.0009	-0.0001	-0.0001	0.0001	0.0001

Table 4.2. Multiplicative and additive biases measured from the noise realisations with $S/N = 20$. This is the bias when all galaxies are included, without splitting into size or redshift bins. The difference is the strength of the model and noise bias interaction terms.

panel). Noise bias effects cause overestimation of ellipticity, relative to the underlying model bias, for fitted Sérsic indices of $0.5 - 3$. For indices $3 - 4$ the ellipticity becomes underestimated.

The red solid line with cross $\times$ markers shows the model bias and noise bias without the interaction terms. To obtain it, I use the best fit model image to the noiseless real galaxy with a given Sérsic index (x-axis). Then I add noise realisations with the same noise variance as in the case above. The mean of noise realisations is marked with the red line and cross points. The interaction terms cause the blue and red lines to differ. When the true model is used in the fit (Sérsic index 1 and 4 for the left and right panels respectively), then noise and model interaction bias terms are zero, and mean biases are consistent.

The bottom panel shows the difference between the noise + model + interaction measurement bias and the noise + model only bias, which probes the strength of the interaction terms. In this case I plotted this difference as the fractional bias, $\delta e_{\rm interact}/e_{\rm true}$. This fractional bias will be similar to the multiplicative bias in Eqn. 4.3 measured from a ring test (Nakajima & Bernstein 2006). Red crosses and green circles correspond to cases where the true Sérsic index was 1 and 4, respectively. The current and upcoming survey requirements are shown as light and dark grey bands, respectively. In the toy example, noise and model interaction bias terms in Eqn. 3.31 can be as significant as $\delta e_{\rm interact}/e_{\rm true} = 0.02$. This demonstrates the potential significant impact of effect: this contribution can exceed the upcoming survey requirements.

The difference is smaller than the effects of the noise bias or model bias alone. In fractional terms, the noise bias on ellipticity was of order $\delta e_{\rm noise} = (\langle e_{\rm fit}\rangle - e_{\rm true})/e_{\rm true} = 0.1$. The maximum model bias measured in this example was $\delta e_{\rm model} = (\langle e_{\rm fit}\rangle - e_{\rm true})/e_{\rm true} = 0.2$ when the real galaxy was created using Sérsic index $S_j = 1$ and fitted model galaxy had Sérsic index of $S_j = 4$.

I have demonstrated that using the wrong galaxy model can give rise to significant noise and model interaction bias terms. However, this bias will strongly depend on the model I use and the real galaxy morphologies. In the next section I quantify the strength of the model bias, noise bias and their interaction using real galaxies from the COSMOS survey, and use a more realistic bulge + disc galaxy model to perform the fit.

4.4 Biases for real galaxies in the COSMOS survey

In this section I evaluate the strength of the noise and model bias interaction terms using real galaxy postage stamp images from the COSMOS survey (Mandelbaum et al. 2011), available for download alongside the GalSim software[8] (RoI et al., in prep.). This training dataset consists of 261113 galaxy images, selected with magnitudes $i_{\rm F814W} < 22.5$ from the full COSMOS sample. This sample comes from May 2011 release, see SHERA web page[9] for details.

The comparison is done in a very similar way to that in Section 4.3: I compare the mean of the estimators from noisy images for the case when the true image is a real galaxy, relative to the case when the true image is a best fit of our model to the real galaxy without added noise. This time, I use more a realistic fitted galaxy model, consisting of two components: bulge and disc. The bulge component is a de Vaucouleurs profile with Sérsic index $n = 4$, disc is an Exponential profile with Sérsic index $n = 1$, ratio of the half light radii of the components was set to $r_B/r_D = 1$. Both components were fixed to have the same centroids and ellpticities.This model was previously used in Zuntz et al. (2013) and K12.

This time, however, I can not afford to repeat this procedure for all 26113 galaxies available in the COSMOS sample, due to computational reasons; constraining the multiplicative bias to $\sigma_m < 0.004$ for a single galaxy image with $S/N = 20$ requires of order 4 million noise realisations. Instead, I group those galaxies in bins of size, redshift, morphological classification, and also model bias. For each galaxy in a bin I use a number of ring tests, with different shears.

I create a set of 8 shears by using all pairs of $g_1, g_2 \in \{-0.1, 0.0, 0.1\}$, except for $g_1 = g_2 = 0.0$. The reason for missing out this middle point is that it brings very little statistical power for constraining the multiplicative bias, whereas additive bias is already

[8] https://github.com/GalSim-developers/GalSim/wiki/RealGalaxy%20Data
[9] http://www.astro.princeton.edu/ rmandelb/shera/shera.html

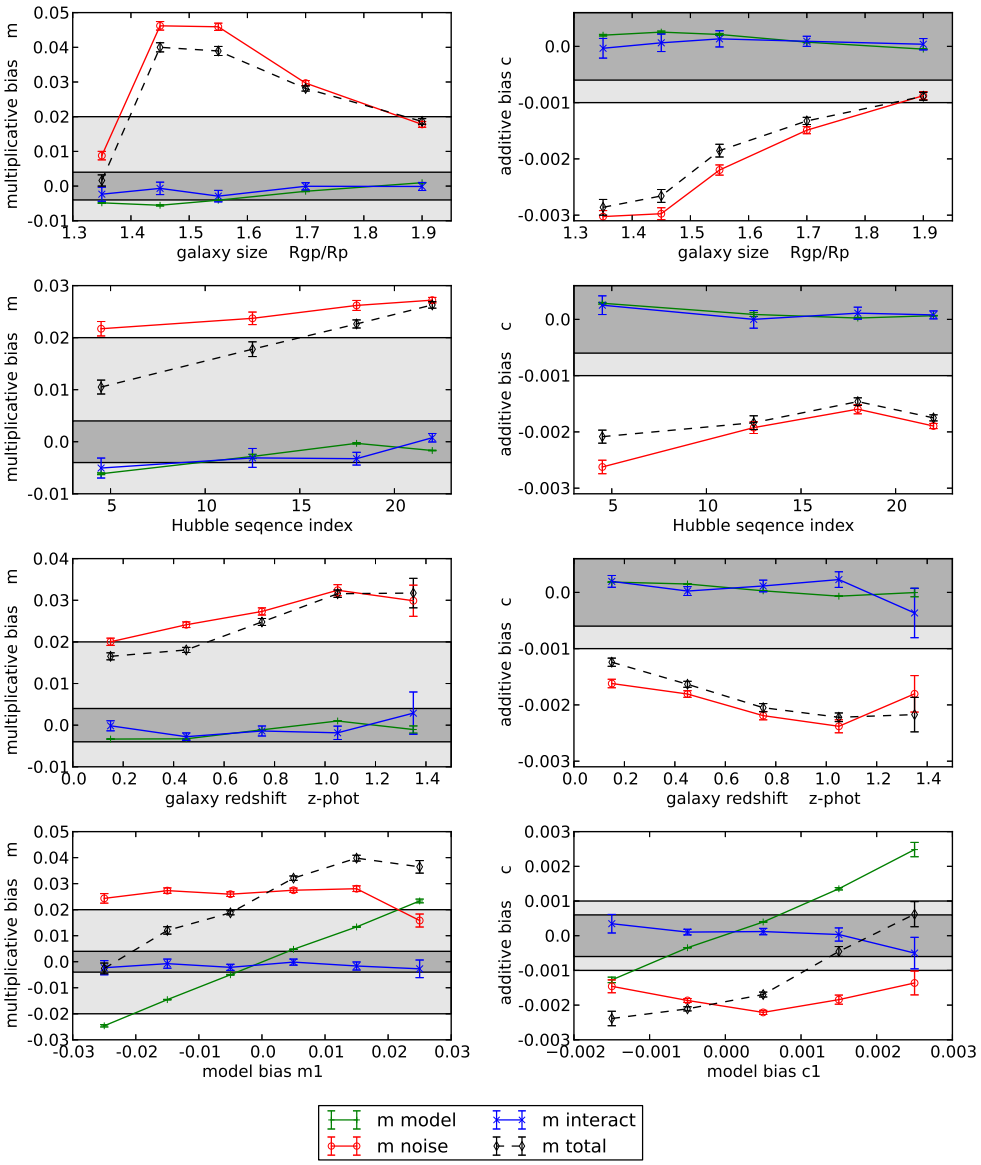

Figure 4.3. Model bias, noise bias and their interaction as a function of various galaxy properties. Left and right panels show multiplicative and additive biases, respectively. The biases shown here are the average of two components: $m = (m_1 + m_2)/2$, ditto for additive bias. The series on these plots presents: (i) the model bias, obtained from noiseless real galaxy images from the COSMOS catalogue, available in GalSim (green plus), (ii) the noise bias, calculated from noise realisations, when the true galaxy was a bulge + disc Sérsic model (red circles); the parameters of that model were that of a best-fit to the noiseless real galaxy images, (iii) the noise bias and model bias interaction terms (blue crosses), obtained by subtracting results of Simulation 1 and Simulation 2, described on Fig 4.2, (iii) model bias + noise bias + their interaction (black diamonds), calculated from noise realisations, when the true galaxy was the real COSMOS galaxy image. The upper left panel shows the biases as a function of galaxy size, measured with respect to the PSF size (FWHM of the convolved galaxy divided by the FWHM of the PSF). The upper right panel shows the biases as a function of the Hubble Sequence Index, where the index corresponds to types: 1 . . . 8 Ell-S0, 9 . . . 15 Sa-Sc, 16 . . . 19 Sd-Sdm, $\geq$ 20 starburst. The lower left panel shows the biases as a function of galaxy photometric redshift. The lower right panel shows the biases as a function of the value of the galaxy model bias.

constrained quite well. For each shear I create a ring test with 8 equally spaced angles.

Our simulation data set consisted of two sets of almost 10 million galaxy images: the first set was created using real galaxy images, the second using their best-fitting Sérsic representations. The number of noise realisations per COSMOS galaxy image was variable and dependent on the number of galaxies in bins of redshift and morphological classification. There are very few galaxies with high redshift or low Hubble Sequence index available in the COSMOS sample, so the number of noise realisations for each of these galaxies had to be larger than for others to reach the desired statistical uncertainty for these bins. However, when showing results for the entire sample, like in Table 4.2 described in following paragraphs in this section, I re-used every galaxy the same number of times.

The image pixel size was 0.27 arcsec, and I fitted to postage stamps of size 39×39 pixels. I use a Moffat profile for the target reconvolving PSF, with FWHM of 0.7695 arcsec= 2.85 pixels, and $\beta = 3$. The ellipticity of the PSF was $g_1^{\rm PSF} = g_2^{\rm PSF} = 0.05$. When I added noise to the simulated images, the signal to noise ratio was again $S/N = 20$.

I measure the ratio of the FWHM of the convolved galaxy (without noise) to the FWHM of the PSF ($\rm R_{g*p}/R_p$). FWHM of the PSF-convolved galaxy is measured numerically from an image of the best fitting bulge plus disc model drawed on a fine grid, with both galaxy and PSF ellipticities set to zero. For small galaxies the shear is strongly underestimated due to noise bias. Multiplicative bias for galaxies in size bin of $\rm R_{g*p}/R_p \in (1.2, 1.3)$ is $m_i \sim -0.2$, whereas model bias is still on sub-percent level. Therefore I remove all galaxies with $\rm R_{g*p}/R_p < 1.3$ for the purpose of this work.

If I perform a reconvolution and shearing operation on a galaxy image containing noise, the final image will contain correlated noise which may align with the shear direction. To assure that this effect is not dominant in our results, I only consider COSMOS galaxies for which S/N > 200. Details of sample selection and the properties of the remaining galaxies are shown in section 4.4.1, and also discussed in Mandelbaum et al. (2011). The majority of the galaxies contained in this sample are of redshift $z < 1.5$ and the convolved size of FWHM < 3 arcsec and are blue (Hubble sequence index > 9). There is a dependence between the galaxy size and Hubble sequence index: most of the elliptical galaxies are small, with convolved FWHM < 1 arcsec, while spiral galaxies span a range of sizes. Sizes of galaxies also depends on redshift: high reshift galaxies tend to be small, while bigger galaxies can have bigger sizes. Dependence of Hubble index on redshift is not prominent.

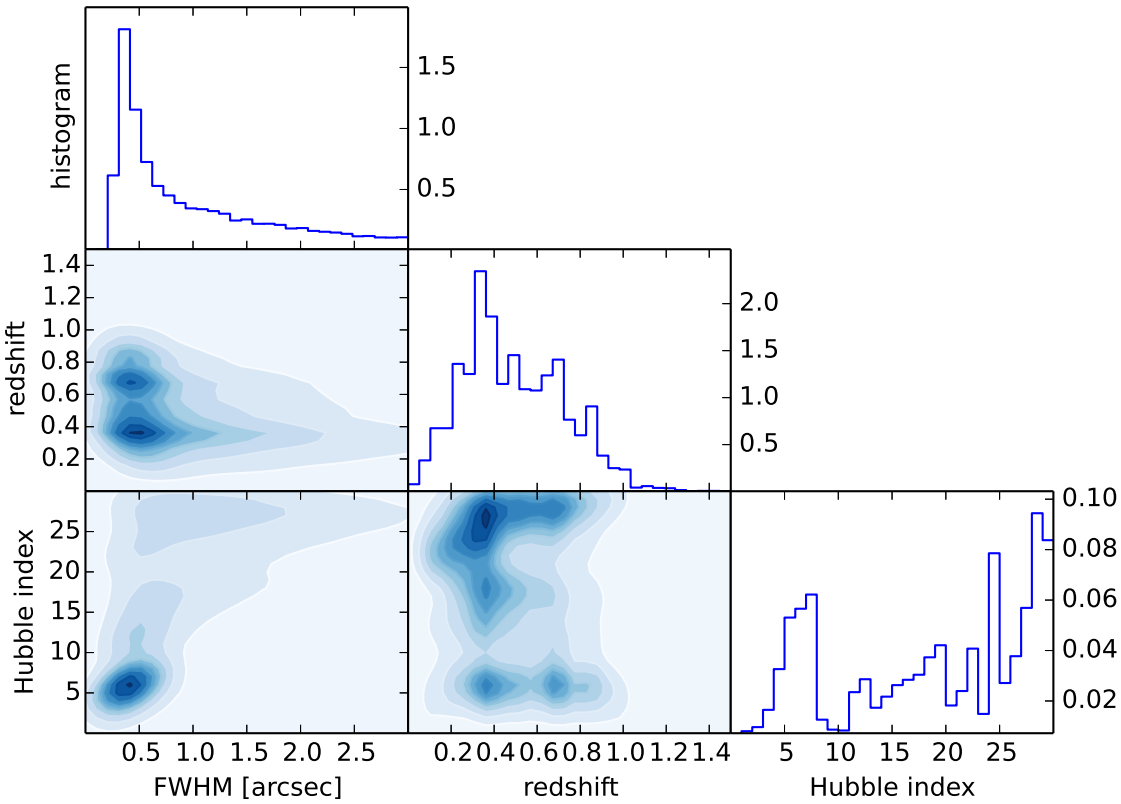

Figure 4.4. Distribution of convolved galaxy size, redshift and Hubble Sequence index. For discussion of the figure, see Section 4.4. The two parameter distributions are estimated using a Kernel Density Estimator for clarity. One parameter histograms are normalised.

4.4.1 Properties of the galaxy sample

This appendix describes a distribution of convolved size, redshift and Hubble Sequence index in the sample of galaxies used in this work. Distribution of these parameters is shown in Fig 4.4, and discussed in Section 4.4.

From the full sample of 26113 galaxies from the COSMOS survey (available via the GalSim software package), we selected 18311 galaxies with HST noise limit of S/N > 200 and valid redshift estimates (we removed all galaxies where `ZPHOT` $= -99$). The noise considered here comes from the HST ACS camera, and does not relate to the noise bias effects studied in this paper. After reconvolving the high resolution HST images with ground PSF, we add more noise at S/N > 20 level. We limit the sample to low HST noise images to avoid influencing our main results by the noise on original HST images.

This noise is further sheared and convolved by the reconvolution procedure, and this can influence the model bias and noise bias results, if S/N is low. However, at later stage of analysis we perform another cut, on the $R_{gp}/R_p < 1.3$. This kind of cut will be applied to real survey data too, as small galaxies are usually removed from analysis. We find that this size cut removes almost all galaxies with S/N > 200 anyway, as more noisy galaxies tend to be small.

Nevertheless, selection of images used for calibration is a very important task; the images have to represent the survey sample well for the calibration to be accurate. Various cuts made on different stages of analysis (including the HST COSMOS survey selection) can create a difference between the calibration and survey samples. This difference may be hard to understand. More discussion on the selection issues for calibration samples can be found in Mandelbaum et al. (2011) and Rowe et al. (2014).

4.4.2 Results

I compared the bias obtained from real galaxy images (including the interaction terms) and best-fitting images (excluding interaction terms) for the whole sample available in GalSim. Table 4.2 shows the differences between those. The results indicate that for the entire selected sample the difference is positive and on a very small level of $m \approx 0.001$, and only 1σ significant.

The mean model bias measured by Im3shape with a bulge + disc model was of order $m = 0.005$ and $c = 0.0003$. Additionally, I measured the model bias for each individual galaxy in the sample and obtained standard deviation on multiplicative bias $std(m) = 0.02$. Model bias can vary significantly, but happens to average out to a rather small value.

Fig. 4.3 shows the noise bias, model bias and their interaction as a function of different galaxy and image properties. The series are as following: (i) model bias only (green plus), measured from noiseless images, (ii) noise bias only (red circles), measured by subtracting two bias measurements: noisy measurement from best fitting Sérsic models, and noiseless real images, (iii) noise bias and model bias interaction (blue cross), obtained by subtracting measurement from noisy images of real galaxies, and bias obtained from noisy images of their best Sérsic representations (subtracting results of Simulation 1 and Simulation 2, described in Fig 4.1), (iv) total bias, containing all these effects (black diamonds with dashed line), measured from noisy images of real galaxies.

The upper left panel shows those biases as a function of the true size of the real

galaxy, measured as R_{g*p}/R_p. For more discussion and investigation of the model bias results alone see Voigt et al. (2013, in prep.). Estimates from galaxies with $R_{g*p}/R_p > 1.3$ seem to be biased positive in the presence of noise. The noise bias decreases as the size of the galaxy increases. This dependence is similar to the one presented in K12. For a galaxy with $R_{g*p}/R_p = 1.6$, in K12 the noise bias was $m = 0.039$ for disc galaxies, $m = -0.013$ for a bulge galaxies and $m = 0.02$ for galaxies with bulge-to-disc ratio of 1. Here, galaxies with this size are biased on the level of $m = 0.035$.

The difference between noise bias obtained from the real galaxies and their best-fitting Sérsic models is shown using green line with plus markers. This difference shows directly the strength of the noise and model bias interaction terms. The difference is very small, consistent with zero to the accuracy limit imposed by the finite number of noise realisations I simulated. This is the case for all the size bins. This means that these terms are either small enough not to make a difference, or their contribution averages out, which should therefore also make no difference to shear statistics, to first order.

The upper middle panels show the multiplicative and additive bias as a function of Hubble Sequence index. The index corresponds to the following types: $1 \ldots 8$ Ell-S0, $9 \ldots 15$ Sa-Sc, $16 \ldots 19$ Sd-Sdm, ≥ 20 starburst from (Bruzual & Charlot 2003). The majority of galaxies in our COSMOS sample have Hubble index > 9. For those galaxies, I can not see a significant difference between noise bias from real galaxy images and their best Sérsic representations. For the elliptical galaxies, however, see a $1-2\sigma$ deviation from zero. This difference is still small within the requirements for upcoming galaxy surveys, although they would dominate the error budget.

The noise and model bias interaction does not significantly deviate from zero as a function of galaxy redshift, for either additive or multiplicative bias.

The model bias should be dependent on the residuals of the best fit image to the real galaxy image, and so should the noise and model bias interaction. I wanted to investigate if there is a dependence between those two effects. Therefore the bottom panels show the biases as a function of model bias. It seems that the noise and model interaction bias terms do not depend strongly on the model bias, which may indicate that their contribution averages out to a very small value, for this galaxy sample.

4.5 Discussion and conclusions

I investigated the problem of shape estimation using model fitting methods, studying in detail the impact of using a model which does not represent the galaxy morphology completely, with noise included. For noiseless galaxies, Bernstein (2010); Voigt & Bridle (2009) introduced and investigated the model bias, which will depend on the complexity of the fitted model. Noise bias, studied in R12 and K12 was derived for the case when the true noiseless galaxy *can* be represented perfectly by our fitted model - there was no bias in the absence of noise. This bias depends on the signal to noise of the image, as well as galaxy size and other properties.

In this work I generalised the noise bias derivations to the case when an imperfect model is used. The interaction between the noise and model bias depends on the residual of the best fit model and the real galaxy image. I isolated the terms dependent on this residual, by comparing the noise bias arising when the true underlying image is the *real* galaxy (residual present), and when it is a *best-fitted* Sérsic image (residuals absent). I call this difference a *noise and model bias interaction.* Thus, model bias, noise bias and their interaction make a complete picture of biases I can expect for fitting parametric models to isolated real galaxy images.

A simple toy model was shown to demonstrate the potential influence of these terms. In our simple toy model, I encountered model biases on a level of 20%, noise biases of order 10%, and the interaction terms of order $1-2\%$. For the toy model, the interaction terms are small compared to model and noise biases, but significant enough that I can detect it using a simulation with finite number of noise realisations.

I then investigated the shear bias induced by model fitting for the real galaxy images available in a COSMOS sample (Mandelbaum et al. 2011). The mean model bias measured by IM3SHAPE with a bulge + disc model was small; of order $m = 0.005$ and $c = 0.0003$. The standard deviation of model bias for individual galaxies is higher; on the level of $std(m) = 0.02$. This indicates that the model bias can be quite variable, but it averages to a very small value. This fact suggests that to properly calibrate a shape measurement method for future high precision lensing experiments knowledge of the spatial distribution of properties of lensing sources may be needed, including evolution with redshift. This could come, for example, from galaxy clustering and galaxy evolution studies. However, I leave quantification of the requirements on the knowledge of spatial and redshift properties

of galaxies to future work (see Voigt et al. in prep).

For the signal to noise level of $S/N = 20$, the maximum noise bias I obtained was on the level of $m \sim 0.04$ and $c \sim -0.003$, for galaxies falling into a size bin of $\mathrm{R_{g*p}/R_p} \sim (1.4, 1.6)$. As a function of galaxy size, the noise bias dependence is very compatible with K12; it decreases as the galaxy size increases. The magnitude of this bias corresponds to a galaxy sample with approximately 2/3 disc-like and 1/3 bulge-like galaxies. The intuitive explanation in the dependence can be based on the derivation in Section 4.2. In general, the bias scales as $\mathrm{S/N^{-2}}$. But it also depends on the inverse of the Fisher matrix. This matrix is dependent on the derivatives of pixel image with respect to the model parameters. If those derivatives are large, then the image contains more information about the galaxy. If the galaxy is small compared to the PSF, those derivatives will be small. If the galaxy is close to a delta function, those derivatives will be close to zero and this will result in large bias. Therefore for smaller galaxies, the bias is larger. The decrease of magnitude of bias for galaxies with size $R_{gp}/R_p < 1.4$ is difficult to ascribe to a single effect, and can be attributed to a few factors, optimizer inefficiency, inclusion of higher order terms in $\mathrm{S/N^{-1}}$, effects of the limited ellipticity boundary.

Noise bias as a function of other galaxy properties, such as redshift and morphological classification, is also variable. However this is almost certainly due to the correlation of the property with galaxy size, which influences the noise bias the most.

The noise and model bias interaction terms prove to be very small, usually almost consistent with zero to the accuracy I simulated, which varied from $\sigma_m \sim 0.001$ to $\sigma_m \sim 0.006$ across bins in different parameters. The only deviation from zero was found for galaxies with Hubble Sequence Index < 19, which corresponds to ellipticals. The noise and model bias interaction terms were estimated as $m = 0.004 \pm 0.002$, which is a 2σ deviation from zero and on the borderline of near future surveys requirements.

The most important shear bias dependence is on galaxy redshift, as it impacts the shear tomography the most. I split our galaxy sample into bins of redshift and measured the model and noise biases for each bin. Noise and model bias interaction does not seem to deviate more than $1 - 2\sigma$ from zero for each bin.

This indicates that the bulge and disc Sérsic model should be a good basis for noise bias calibrations for current surveys. Furthermore, the model bias and noise bias calibrations can be performed in two separate steps. Using Sérsic profile parameters to evaluate the noise bias can reduce the simulation volume significantly, by using a procedure described

in K12. Similarly, if noise bias equations can be evaluated analytically or semi-analytically with sufficient precision, the fact that the noise and model bias interaction terms can be neglected to the accuracy of roughly $\sigma_m \sim 0.003$, means that the noise bias prediction task can be simplified even further. Also, if a galaxy model used proves to be robust enough and introduce negligible model bias, simulations based on Sérsic models should prove sufficient for the shear calibration.

For the results presented in this chapter I used a Sérsic, co-elliptical, co-centric bulge + disc model, with fixed radii ratio of the two components. This model choice is well motivated, but arbitrary to some extent. Selecting a good model for a galaxy is not a trivial task. Increasing the model complexity by adding more parameters (for example allowing for different, variable ellipticities for bulge and disc components) can result in lower model biases (Voigt & Bridle 2009). Our intuition is that for any model-fitting method applied to ground based observations should show negligible noise and model bias interaction, if the complexity of models used has similar number of degrees of freedom.

For upcoming and far future surveys, the noise and model bias interaction terms might become important. Minimizing this contribution may involve finding a better galaxy model (simultaneously decreasing the model bias), and/or using representative calibration images and reconvolution technique. When a good quality, representative image sample is available, or model bias calibration is necessary, it will always be more accurate to use the images themselves.

This page was intentionally left blank

Chapter 5

Detection of weak lensing by dark matter filaments in CFHTLenS

I search for the weak lensing signal from dark matter filaments between LRGs residing in massive clusters in CFHTLenS fields. I use a model for two halos and a filament and measure the likelihood of its parameters. The lensing signal from filaments is very weak, making them difficult to detect individually. Thus I combine measurements from multiple pairs of halos, taking a product of likelihood estimates from filament model fits. To partially account for intrinsic variability in the properties of each halo I parametrize its density relative to the densities of connected halos. Using 19 pairs of halos, selected to have a detectable weak lensing signal in each halo and having at least one halo with mass $M_{200} > 2.5 \times 10^{14}\ h^{-1}\mathrm{M}_{\odot}$, I detect the filament signal at $\simeq 3.9\sigma$ significance. The peak filament density relative to the mean density of paired halos at their R_{200} is measured as $\Delta\Sigma^{\mathrm{filament}}_{\mathrm{peak}}/\Delta\Sigma^{\mathrm{halos}}_{R200} = 0.50 \pm 0.13$. For two halos of mass $M_{200} = 3 \times 10^{14}\ h^{-1}\mathrm{M}_{\odot}$, this corresponds to a peak projected filament density of $\Delta\Sigma^{\mathrm{filament}}_{\mathrm{peak}} \approx 10^{13}\ h\mathrm{M}_{\odot}\mathrm{Mpc}^{-2}$. The filament radius is more weakly constrained as $R_{\mathrm{scale}} = 1.66^{+0.81}_{-0.54}\ h^{-1}\mathrm{Mpc}$. Overall, the model containing a filament is preferred over a model without filaments with a p-value of 0.0003. To further validate the detection, I investigate the potential influence of effects that can mimic a filament signal, including the lensing signal from the connected halos and lensing from other large scale structures at different redshifts. Null tests using simulated halo pairs at random points in CFTHLenS indicate that these effects do not

have a significant impact on the detection.

5.1 Introduction

A prominent feature of large-scale matter structure in standard cosmological models is the existence of intercluster filaments. These filaments are believed to lie between most massive clusters and consist of dark matter, hot gas and galaxies. In the future, these objects can be useful for studying cosmology and gravity. In particular, comparison of properties of filaments at different redshifts and in different density environments can tell us more about the collapse of matter in gravitational instability scenarios, including with modified gravity.

Filaments have recently been detected using gravitational lensing (Jauzac et al. 2012; Dietrich et al. 2012; Clampitt et al. 2014), as well as in the distribution of galaxies (Zhang et al. 2013; Tempel et al. 2013) and the thermal Sunyaev-Zel'dovich effect using X-ray and CMB data (Planck Collaboration 2012). Many theoretical studies using N-body simulations (e.g. Cautun et al. 2014; Higuchi et al. 2014; Sousbie 2013; Bond et al. 2010) have quantified the properties of the cosmic web. In particular, early work by Colberg et al. (2005) showed that filaments are abundant: for massive halos separated by up to $12h^{-1}$Mpc the probability of finding a filament approaches 80%.

In this chapter I look for weak lensing signal from dark matter filaments in CFHTLenS shear data. I use a probabilistic approach, fitting a model consisting of two dark matter halos and a connecting filament. Measuring this signal is challenging. Individual filaments are difficult to detect with current weak lensing data as their signal is faint. Mead et al. (2009) use simulations to estimate the signal to noise ratio (SNR) of filament detections as a function of number of background galaxies used for lensing, using multiple approaches. They report that individual filaments can be detected with $\sim 2\sigma$ confidence using lensing data with $\sim$30 galaxies per arcmin^{-2}.

Recently, Dietrich et al. (2012 D12 hereafter) detected a single filament with a significance of 4.1σ, using SuprimeCam data with high source galaxy number density. This filament is serendipitously oriented almost parallel to the line of sight. Such a configuration increases the projected density contrast and the lensing signal, but such objects are of course rare. To increase the SNR of detection, one can combine measurements from multiple halo pairs. Clampitt et al. (2014) adopted a stacking technique to combine

measurements from SDSS lensing data. In this work, I select 19 pairs of lensing-detected halos and combine estimates of filament parameter likelihoods derived from model fits to CFHTLenS data.

A second difficulty is that the filament signal is degenerate with the lensing introduced by the connected halos themselves. Any method aiming to detect filaments must account for this. Maturi & Merten (2013) derive a set of optimal filters to achieve this, while Clampitt et al. (2014) use a nulling technique which exploits symmetries in the lensing signal around individual halos. Dietrich et al. (2012) and Mead et al. (2009) employ multiple methods including parametric model fitting, using NFW (Navarro, Frenk & White 1997) profiles for the halos. I also present multiple methods to remove the halo signal: halo parameter marginalisation with both derived and flat priors, and the subtraction of maximum likelihood and expectation halo contributions (see Sec. 5.4).

A challenge when combining measurements from multiple halo pairs is the selection of pairs likely to have filaments, since not all close halos are predicted to be connected by filaments (e.g. Colberg et al. 2005). Filaments are more likely to connect close halos. As including unconnected halos in a sample of filament candidates will dilute the signal and make detection more difficult, it is important to select the pairs most likely to contain a filament. To address this problem I use a simple algorithm to compare probable connections between halos in CFHTLenS.

Not all filaments are predicted to be equally dense, and each filament's density will depend on many factors including halo masses, environment and history. When combining estimates from multiple halo pairs this variation in density would ideally be accounted for, as would variation in other parameters, such as the radius of a filament, which may be similarly affected. To partially address density variations due to the immediate environment I use a model in which the filament density is dependent on the mass of halos it connects.

The chapter is organised as follows. In Sec. 5.2 I describe the details of our model. Sec. 5.3 presents our algorithm for selecting halo pairs which probable to have a dark matter filament. Details of our analysis are shown in Sec. 5.4. In Sec. 5.5 I present results, and conclude in Sec. 5.6.

5.2 Filament model

I use a simple parametric mass model for the shear signal due to two halos and a connecting filament. For the halos I use an NFW profile (Navarro et al. 1997). The filament is a cylindrical profile with two free parameters: a projected peak mass density and a scale radius governing thickness.

I expect the filament density to be related to the density of connected halos. Although this relationship is probably very complicated, we do not expect the peak surface density of filaments to exceed the peak density of halos. Therefore some dependence of filament density relative to the density of halos is needed.

I scale the peak projected filament density to the mean of the paired halos' projected density at their respective virial radii R_{200}. The free scaling parameter is

$$D_{\rm f} = \Delta\Sigma^{\rm filament}_{\rm peak} / \Delta\Sigma^{\rm halos}_{R200} \tag{5.1}$$

where $\Delta\Sigma^{\rm halos}_{R200} = (\Delta\Sigma^{\rm halo1}_{R200} + \Delta\Sigma^{\rm halo2}_{R200})/2$, and where $\Delta\Sigma$ denotes sky-projected mass density. The scaling relation (5.1) is oversimple, as filaments will depend non-trivially on their environment and formation history. However, current wide field data do not support greater model complexity, and this filament model has the advantage of taking some account of connected halo properties without adding parameters. I do not use the same scaling for the radius parameter, which is measured in h^{-1}Mpc.

Another property of filaments is that the projected density contrast is dependent on the line of sight distance between halos. Observed density $\Delta\Sigma$ will be related to density observed if I were looking at this filament 'side-on', $\Delta\Sigma_{\rm side-on}$, in the following way:

$$\Delta\Sigma \approx \Delta\Sigma_{\rm side-on} \cdot \frac{D_{\rm tot}}{D_{\rm sky}} \tag{5.2}$$

where $D_{\rm tot}$ is the total length of the filament and $D_{\rm sky}$ is its length on the sky plane. This means that filaments with the same density distribution in 3D will have different projected densities in 2D, depending on orientation with respect to an observer. For a method which combines measurements from multiple pairs, this effect will introduce a scatter in measured $\Delta\Sigma$. That is why, in our parametrisation, the $\Delta\Sigma$ is always measured as though the filament was seen side-on. I include that effect when the shear model is drawn, effectively fitting the $\Delta\Sigma_{\rm side-on}$ parameter. Throughout the rest of the chapter I

will use only $\Delta\Sigma_{\rm side-on}$, and refer to it with just $\Delta\Sigma$.

I use the same model filament convergence profile as D12:

$$\kappa(R,\theta) = \frac{\kappa_0}{1+\exp\left(\frac{|\theta|-\theta_l}{\sigma}\right)+\left(\frac{R}{R_{\rm s}}\right)^2} \tag{5.3}$$

where the coordinate R is angular distance perpendicular to the filament axis, θ is a coordinate along the filament axis with an origin at the mid-point, $R_{\rm s}$ is a scale radius, $\theta_{\rm l}$ is the point where the filament begins, and σ regulates the length of the smooth increase of density at the beginning of the filament region. Density contrast is very close to constant along the line joining the halos. This is also a simplification, as N-body simulations Colberg et al. (2005) indicate that a filament's density vary along its line with a smallest density in the mid-point.

I fix θ_l such that filaments begin at 1 h^{-1}Mpc from each halo center, motivated by the typical virial radii of $M_{200} \simeq 1-3\times10^{14}\ h^{-1}{\rm M}_\odot$. The transition scale was fixed at $\sigma = 0.25h^{-1}$ Mpc, close to the value used by D12, who found that this parameter was very poorly constrained by the data. I verified that changing this parameter to 0.5 did not noticeably change our results. Following Mead et al. (2009) and Maturi & Merten (2013), I use reduced shear $g_1(R,\theta) = -\kappa(R,\theta)$, $g_2(R,\theta) = 0$. I notice that this equation is true for a infinite filament bar. Symmetry arguments can be used to derive this dependence. For the infinite bar model, the derivatives of deflection potential with respect to the coordinate varying along the axis connecting the halos. This means that all derivatives in Eqn. 2.9 can be set to zero. For a case of a finite filament, as used in our model, some non-zero g_2 is expected. However, I do not model this part, as most of the signal is indeed in the g_1 part.

For the halos I use spherical NFW profiles despite the predicted, and detected, asphericity of real halos (e.g. Evans & Bridle 2009). However, this work is focused on filament detection, and uses halos pairs of moderate projected separation (4 – 12 Mpc; see Sect. 5.3) for which extreme halo ellipticity would be required to mimic a filament signal. With future data it will be interesting to explore how halos align with connecting filaments. Additionally, it is a matter of debate whether a halo elongation towards the axis of the filament should be considered a part of the filament itself.

I fit a single parameter for each halo profile, M^i_{200} for $i = 1, 2$, relating each halo's concentration c_i to mass via the scaling of Duffy et al. (2011). Halo centroids are fixed

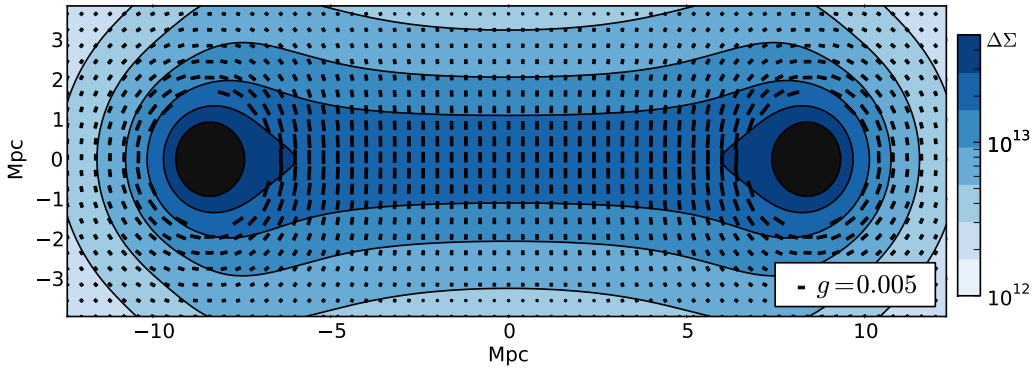

Figure 5.1. Example model for lensing signal used in this work. For clarity, no shear whiskers were drawn within the radius of 1 Mpc/h from the centres of halos. In this region the shear signal is much higher than in the filament region. This mask was not present in the fitting process. Peak density contrast of this filament is $\Delta\Sigma = 10^{13}$ $M_\odot$Mpc^{-2}h, and radius is $R_s = 1.5$ Mpc/h. Masses of both halos are $3 \cdot 10^{14}$ $M_\odot$/h. As the mean $\Delta\Sigma^{\rm halos}_{R200} = 2 \cdot 10^{13}$ $M_\odot$/h, this corresponds to $D_f = 0.5$.

to the position of the Luminous Red Galaxy (LRG) used to select the halo. The shear contribution from the halos is calculated using the equations in Wright & Brainerd (1999).

Thus, our model has four degrees of freedom: $D_{\rm f}$, $R_{\rm s}$, M^1_{200} and M^2_{200}. An example density map and corresponding shear in the surrounding region is shown in Fig. 5.1. Regions within 1 h^{-1}Mpc from halo centres were masked for clarity on the plot, as the shear is significantly larger there.

5.3 Pair selection

In this section I present the procedure I followed to create a sample of halo pairs, between which I look for filament signal. I use LRGs from SDSS DR10 (SDSS Collaboration 2014) as halo tracers, finding 2187 inside the CFHTLenS wide fields. These LRGs have spectroscopic redshifts, giving precise distance estimates.

I wish to select pairs of LRGs which are likely to trace haloes with a connecting filament. This is important, as non-connected halo pairs will bias our result low and possibly prevent detection. As a pre-processing step, I fit a spherical NFW profile (with a single free parameter M_{200} as in Sec. 5.2) to the shear around selected LRGs in CFHTLenS (see Sec. 5.4 for details of the fitting process), and measure the lensing halo detection

significance. Equipped with a sample of LRGs with known positions and lensing mass, I proceed to select pairs of halos for filament detection.

From all possible LRG halo pairs, a filament sample is selected according to the following considerations:

1. A filament is likely to connect a halo only to its closest neighbour in a particular direction. I use the following algorithm to create a graph of possible connections. Starting with most massive halo, I find all its neighbours within 30 Mpc/h. Then, starting with the closest neighbour, I begin to add to the connections list. Before I add a new connection, I check the angle that this new connection would form with all connections which have already been added. I add a connection only if this angle is always greater than 90 degrees. I repeat this procedure for all halos, making sure I do not duplicate connections. This procedure creates 474 pairs.

2. After creating a graph of possible connections, I perform another selection. I require the distance between halos to be between 4-12 Mpc/h. According to Colberg et al. (2005), these pairs have a good chance of being linked a dark matter filament. The halos can not be too close to each other on the sky plane, as the lensing signal, although enhanced by large projection angle, would be very hard to distinguish from halo signal. Therefore I set the minimum of 4 Mpc/h for the separation in the sky plane.

3. Connected halos have to be massive. I select only those pairs, for which at least one halo has measured mass $M_{200} > 2.5 \cdot 10^{14}$ using its lensing signal. Significance level of minimum 1σ is also required. This selection, combined with above cut on distance, produces 20 filament pairs.

4. As I do not want to re-use the shears in a filament region, data stamps can not overlap greatly. I remove one pair, which overlapped with another one. Here I choose a shorter connection, leaving the sample size at 19. Within the sample, 4 pairs remain which share a cluster, but the filament regions do not overlap. To avoid re-using information assumed to be independent, I should remove one of pairs in such system. However, I do not do this for our main sample, as the halos signal is removed from by marginalisation, so I expect that effect influences the filament parameters very little. Indeed, when I compare main result with a result with those 4 connections removed, it changes only very slightly (see Sec. 5.5).

The final set of connections is shown on Fig 5.2. Colour of the halo signifies its redshift. There are no pairs in CFTHLenS W2 field, as it was not sufficiently covered by BOSS spectra. Further details of each pair are described in Table 5.2. Grey lines indicate all 476 connections found by first step of our procedure. I consider that the distance calculations can be affected by peculiar velocities of LRGs. Clampitt et al. (2014) estimated that the error on line of sight distance due to peculiar velocities of LRGs can be of order of 6 Mpc/h. In worse case scenario, this effect would increase the connection length to about 20 Mpc/h. This can be tolerated, as it is still probable for clusters separated by this distance to be connected by a filament (Colberg et al. 2005).

In all calculations I assumed flat ΛCDM cosmology, with $\Omega_m = 0.27$. I verified that the results change insignificantly when $\Omega_m = 0.3$ is used instead.

5.4 Analysis

I begin our analysis with fitting isotropic NFW halos to the position of BOSS LRGs. The fit is done using a grid search in M_{200} parameter, with concentration taken from the mass-concentration relation. I prepare shear postage stamps, within 4x4 Mpc region around a LRG. Shear catalogues are taken from the official CFHTLenS release (Heymans et al. 2012; Miller et al. 2012; Erben et al. 2012). I bin shears with pixel size of 0.1 MPc, using the following formula:

$$g_p = \frac{\sum_{i \in p} e_i w_i}{\sum_{e_i \in p} (1 + m_i) w_i} \tag{5.4}$$

where g_p is the shear at pixel p, e_i is a galaxy shear estimator in CFHTLenS, w_i is the weight corresponding to that galaxy, and m_i is the multiplicative bias correction. I create maps for both g_1 and g_2 components, but I additionally subtract the additive correction from the g_2 component before averaging. There may be other systematics present in the data, and some CFHTLenS fields had to be excluded from cosmic shear calculations. However, these fields were found not to impact galaxy-galaxy lensing science (Hudson et al. 2013), and I also do not attempt to remove them. Additionally, as the pair orientation angle should be random, I expect any additive systematic to be uncorrelated with it and not bias our measurement. As the weights in CFHTLenS catalogs are close to, but not exactly corresponding to inverse variance errors, I follow Hudson et al. (2013) and use a fixed error on the shear: $\sigma(g_p) = 0.28/\sqrt{n_{\rm eff}}$, where $n_{\rm eff} = (\sum_{i \in p} w)^2 / (\sum_{i \in p} w^2)$ is

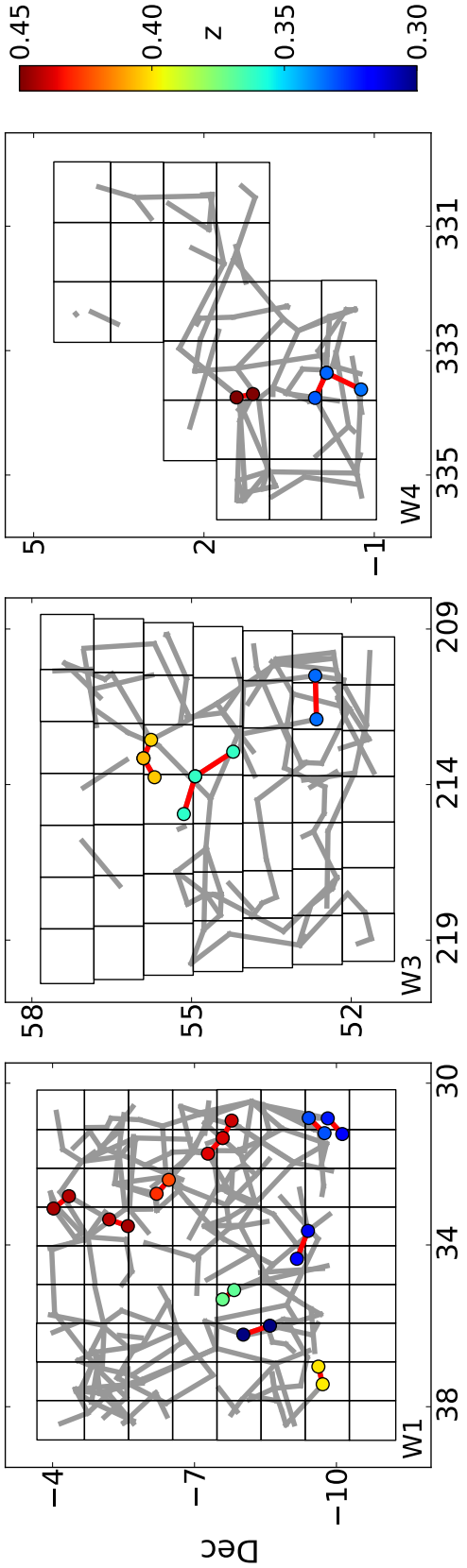

Figure 5.2. Locations of filaments in CFHTLenS fields. Colour of the markers corresponds to redshift of the LRGs. Grey connections show the network of 476 probable filaments obtained using first step of pair finding procedure. For main analysis, I did not use any pairs from the W2 field, as there are only few BOSS LRGs in that area. More details about these pairs can be found in Table 5.1.

the effective number of galaxies in pixel. I have verified that treating weights as inverse variance errors by using equation $\sigma(g_p) = [\sum_{i \in p} w_i]^{-1/2}$ changes our results only very slightly. I require CFHTLenS galaxies to have parameter `fitclass`=0. The selection of background source galaxies is performed by using a strict cut on `Z_B` column; I use only those galaxies for which redshift is greater than the mean redshift of halos plus 0.2. For creating a shear signal, I use $p(z)$ created using average probability of normalised `PZ_full`. I measure the likelihoods of M_{200} and significance intervals of individual detections. This task is performed using grid search with 2000 points distributed evenly between 10^{12} and 10^{15} $M_\odot/h$. I use measured masses to select halo pairs for further filament analysis, as explained in section 5.3.

For each of the 19 selected pairs, I draw a rectangular box around the pair region, as in Fig. 5.1, with 4 Mpc/h in y-direction and leaving additional 4 Mpc/h on each side of halos. I create shear postage stamps as in Eqn. 5.4. I proceed to fit a model described in Sec. 5.2. I use the following ranges for parameters: $M_{200} \in [10^{13}, 10^{15}]$ $M_\odot$/h for both halos, $D_f \in [0, 1.2]$ and $R_s \in [0.001, 4]$. The default grid sizes were (15, 15, 30, 30) for four model parameters, respectively. Likelihood is measured for each combination of parameters on a grid, which gives a total of 202500 grid points. I verified that increasing the density of the grid changes our results at a negligible level.

With these likelihood cubes I proceed to analysing the filament signal. I aim to plot the constraints on D_f and R_s parameters. To combine the measurements from 19 pairs, I multiply the two parameter likelihoods of individual pairs. Before that, however, I have to remove the contribution from halo lensing. I compare four schemes of removing this signal: marginalisation with constructed prior, marginalisation with a flat prior, subtraction of halo signal inferred via maximum likelihood, and subtraction using the expectation of M_{200}. The aim of this comparison is to demonstrate that the detection is robust to halo signal removal methods.

First method, constituting our main result, is a full probabilistic treatment: I marginalise the halo parameters with a prior on halo M_{200}. I construct this prior by taking a mean of likelihoods of M_{200} calculated in the pre-processing step (from 4x4 Mpc/h postage stamps), for all halos in the sample of 19 pairs. Note that this prior was inferred using stamps which could consist a part of a filament, but I do not expect this to influence the results significantly, as halo shear signal is much greater than the filament signal. Fig. 5.3 presents the prior calculated this way. Marginalisation is performed as following

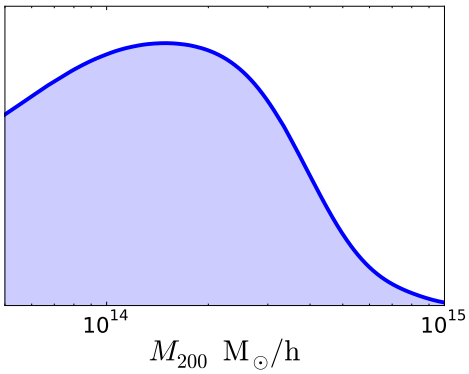

Figure 5.3. Prior on M_{200} created for 38 halos. This prior was created by taking the average likelihood of M_{200} parameter, which was fit to 4x4 Mpc/h cutouts surrounding the LRG.

$$p(D_f, D_f|\mathcal{S}) = \prod_i \int \mathcal{L}(\mathcal{S}_i|D_f, R_s, M_1, M_2)\pi(M_1)\pi(M_2)dM_1dM_2 \tag{5.5}$$

where $\mathcal{L}$ is the likelihood function and π are priors, and $\mathcal{S}_i$ is the data vector, with index i iterating over halo pairs.

In the second method a flat prior on M_{200} of halos is used. The third approach multiplies those D_f – R_s distributions, for which halos M_{200} correspond to maximum likelihood fits to halo signal measured earlier from single parameter fits to 4x4 Mpc/h postage stamps. The fourth approach also removes the halo signal using a point estimate, and this estimate is an expectation of M_{200}. Note that the last approach is the most conservative, as M_{200} values are always higher than for the maximum likelihood method.

I plot the constraints on D_f and R_s parameters, improving the resolution of the grid with a spline interpolation.

5.5 Results

With a set of 19 pairs, I calculate the posterior probability on filament parameters using the methods described in Sec. 5.4. As described in section 5.2, these parameters are: filament's peak density relative to mean density of halos at their R_{200} (D_f) and filament's scale radius (R_s). Coordinates and measured parameters for these pairs are presented in Table 5.2.

5.5.1 Significance

I report a detection of mean filament lensing, with maximum posterior probability at $D_{\rm f} = 0.50 \pm 0.13$ and $R_{\rm s} = 1.66^{+0.81}_{-0.54}$. Significance for $D_{\rm f}$ is 3.9σ for best fit radius and 3.3σ for marginalised $R_{\rm s}$. The radius is poorly constrained, especially its upper limit has very large error. $R_{\rm s}$ is significant at 3.1σ level for the best fit $D_{\rm f}$ and 2.7σ for marginalised $D_{\rm f}$. Figure 5.4 shows the measured constraint. Middle and lower panels show likelihood for single parameters, with the other either fixed to maximum likelihood or marginalised. When I compare the model with a filament against the null model using a maximum likelihood ratio test, I obtain a p - value of 0.0003. That suggests strong evidence for the extended model.

Alternative methods of accounting for the halo signal give similar detection significance, as shown in Table 5.1. The lowest threshold of 3.1σ is reached for halo removal with maximum likelihood M_{200}. The fact that the detection holds for all the proposed halo signal removal methods indicates that the results is not strongly dependent on the choice of the method. In particular, the main result is not driven by the use of a custom prior on mass. The last line shows the result if I did not allow for pairs in the sample to share one connecting halo, reducing our sample to 15 pairs. This result is shown for first halo removal method. There is no noticeable difference in central values of parameters, and the significance decreases as expected.

The uncertainty on filament density is large, but it is possible to try to relate filament density contrast to mass of the connected halos. For two halos with mean mass $M_{200}^{\rm halos} = 3 \cdot 10^{14} {\rm M}_{\odot}/{\rm h}$, the filament density contrast is between $\Delta\Sigma = 0.9\text{–}2.2 \cdot 10^{13}$ ${\rm M}_{\odot}{\rm Mpc}^{-2}{\rm h}$. Assuming the length of a filament to be 10 Mpc/h, I estimate the filament mass between $0.7 \cdot 10^{13}$ and $1.9 \cdot 10^{14} {\rm M}_{\odot}/{\rm h}$ for best-fit radius, and can reach $3 \cdot 10^{15}$ for large radii, given our poor constraint on this parameter. Although these estimates do overlap, I point out that our measurement considered an ensemble of halo pairs, each of which had a filament (or accidentally selected unconnected region) of different mass. Moreover, our best fit filament radius was more than twice as large as in D12, where the radius was not well constrained and set to 0.5 Mpc. These large errors do not allow us to draw conclusions about a typical mass of a filament in the Universe.

The filament signal is much higher for the most massive pairs of halos. I found five pairs for which the mass of both halos to exceeds $M_{200}^{\rm halos} > 2 \cdot 10^{14} {\rm M}_{\odot}/{\rm h}$. For these pairs,

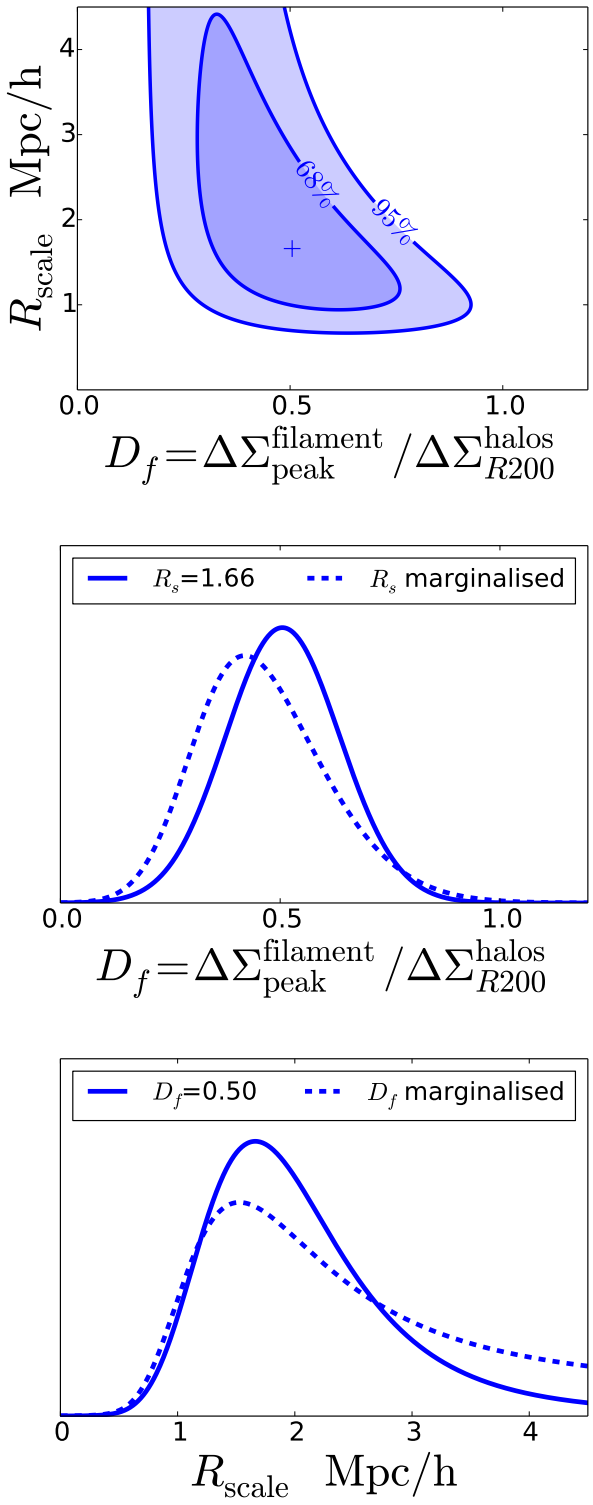

Figure 5.4. Posterior probability of mean filaments parameters, from a sample of 19 halo pairs. Top panel shows the constraint on $D_{\rm f}$ and $R_{\rm s}$ parameters. The maximum of this distribution is at $D_{\rm f} = 0.50 \pm 0.13$ and $R_{\rm s} = 1.66^{+0.81}_{-0.54}$, which corresponds to 3.9σ detection for density and 3.1σ for the radius. Lines correspond to 68% and 95% confidence intervals. Lower panels contain single parameter distributions, for other parameter set to either best fit (solid) or marginalised (dashed) with a flat prior in a volume, which was the same as shown in top panel. Best marginalised parameters were $D_{\rm f} = 0.42^{+0.16}_{-0.13}$ (3.3σ) and $R_{\rm s} = 1.53^{+1.27}_{-0.56}$ (2.7σ).

	# pairs	$D_\mathrm{f} = \Delta\Sigma^\mathrm{filament}_\mathrm{peak}/\Delta\Sigma^\mathrm{halos}_{R200}$	R_scale
halo mass marginalisation with a derived prior	19	0.50 ± 0.13 (3.9σ)	$1.66^{+0.81}_{-0.54}$ (3.1σ)
halo mass marginalisation with a flat prior	19	0.43 ± 0.12 (3.6σ)	$1.79^{+1.01}_{-0.62}$ (2.9σ)
point estimate halo removal with maximum likelihood	19	0.41 ± 0.12 (3.5σ)	$1.76^{+1.05}_{-1.05}$ (1.7σ)
point estimate halo removal with expected M_{200}	19	0.42 ± 0.12 (3.6σ)	$1.77^{+1.02}_{-0.67}$ (2.8σ)
avoid connections sharing a halo	15	0.50 ± 0.15 (3.4σ)	$1.48^{+0.72}_{-0.50}$ (2.9σ)
most massive pairs, $M^1_{200}, M^2_{200} > 2 \cdot 10^{14}\mathrm{M}_\odot/\mathrm{h}$	5	1.17 ± 0.23 (4.3σ)	$0.90^{+0.27}_{-0.23}$ (3.7σ)
longer filaments, 10 – 22 Mpc/h	36	0.19 ± 0.06 (3.2σ)	$1.54^{+0.94}_{-0.66}$ (2.4σ)

Table 5.1. Summary of results obtained with several halo removal approaches. The main results used a full probabilistic method. The detection significance varies slightly between these methods, but the detection holds for all of them. The impact of our derived prior is not dominating the significance for our main result. For last two rows, halo signal was removed via marginalisation with a prior. Fifth row shows the results for a reduced sample of 15 pairs, in which I did not allow connections to share a halo. Second to last row shows the constraint for five most massive pairs. Last row shows the result for longer filaments, from 10 to 20 Mpc/h, allowing 12 of them to share a halo. For that case halo mass requirement was lowered, requiring at least one of the halos to have $M_{200} > 2 \cdot 10^{14}\mathrm{M}_\odot/\mathrm{h}$.

I measured $D_\mathrm{f} = 1.17 \pm 0.23$ (4.3σ) and $R_\mathrm{s} = 0.90^{+0.27}_{-0.23}$ (3.7σ). This is more than twice the signal from 19 pairs, however, due to small number of pairs, it's quite uncertain. Constraints from this sample and the main sample of 19 pairs overlap on approximately 2σ level. It is not unexpected for more massive systems to have more massive filaments. Moreover, the linear dependence of mass of filament on connected halos I assumed in our model is likely to be flawed. Another possible reason for this discrepancy can be the fact that connections between halos of smaller mass can be less abundant; our 19 pair sample may contain accidentally selected disconnected halos. I do not attempt to quantify the influence of this selection effect in this work. Selection effects can be studied further by testing the algorithm on simulations.

Allowing for longer filaments also gives a good detection significance, however, the constraint on the radius is much worse. I found 40 pairs with maximum separation of 20 Mpc/h, and after removing 5 pairs which overlapped in shear region, I obtained a constraint on $D_\mathrm{f} = 0.31 \pm 0.07$ and $R_\mathrm{s} = 1.9^{+0.8}_{-0.6}$. Density parameter D_f is smaller than for shorter filaments, but the fact that I am using additional pairs gives comparable significance. The requirements on halo mass was kept the same as before. For this sample there are many filaments which share a halo, if I were to remove connections so that no

halos are re-used, sample size would decrease to 23. Detection significance would also fall accordingly to between 2 – 3 σ depending on the random choice of removed connections.

5.5.2 Null test: random points

Measurement of filament parameters may also be affected by other signals not included in our model. Large scale structures (LSS) on different redshifts can create a lensing signal, which can be mimic a filament. Note that to make a detection, fI am only concerned about effects which can bias the measurement high. If the parameters are biased and the detection still holds, I would not attempt to correct for such a systematic. One way to test the effects of LSS is create a sample of random points in CFHTLenS and measure filament parameters centred at these points, with a random orientation. In this *random points* null test, I use the same parameters for pairs as for our main sample, including their redshifts and separations. I add signal from two halos with M_{200} parameter taken from best fits to halos in our main result sample. I make new shear cut-outs for 32 sets of 19 pairs. I go through the same fitting procedure with a four parameter model. None of the 32 random points sets produced a filament lensing detection with significance greater than 2.5σ in both $D_{\rm f}$ and $R_{\rm s}$. The right panel of Fig. 5.5 shows the product of likelihood for this null test. This likelihood is a mean of 32 sets; within each set likelihoods of 19 pairs were multiplied together. Notice the smaller range on the x-axis compared to Fig. 5.4. These contours are quite broad, but notice that in the region of our best fit, the mean probability is low. This suggests that it is very unlikely that the mean likelihood measured from CFHTLenS filaments (Fig. 5.4) is a result of a noise or LSS fluctuation. Blue dashed lines shows the result of multiplying $32 \cdot 19$ likelihood surfaces, instead of taking a mean of 32. The fact that this product is consistent with zero indicates that there is no bias due to cosmic shear or other large scale structures. This allows us to conclude that our measurement was not affected by LSS to a level which could affect the detection.

5.6 Conclusions and discussion

I have detected weak lensing signal between massive clusters in CFHLenS centred on BOSS LRGs. The detection significance was close to 3.9σ, depending slightly on halo signal removal method. The central parameters measured for filament peak density relative to mean halos density at their R_{200} was $D_{\rm f} = 0.50 \pm 0.13$. It is interesting to compare our

	RA_1	Dec_1	RA_2	Dec_2	z_1	z_2	D_{sky}	D_{los}	$M^1_{200}/10^{14}$	$\sigma(M^1_{200})$	$M^2_{200}/10^{14}$	$\sigma(M^2_{200})$
1	30.870	−9.817	31.258	−10.124	0.3217	0.3174	5.786	8.304	8.3	3.7	2.1	2.3
2	35.996	−8.595	36.218	−8.031	0.2705	0.2721	6.362	3.437	5.2	3.8	0.6	1.5
3	34.343	−9.165	33.653	−9.397	0.3179	0.3148	8.433	6.206	3.7	2.6	0.7	1.5
4	32.794	−4.345	33.099	−4.020	0.4433	0.4446	6.193	2.190	3.8	2.0	1.9	1.8
5	333.357	−0.160	333.761	0.044	0.3337	0.3318	5.199	3.603	2.9	2.3	0.5	1.2
6	37.445	−9.710	37.009	−9.618	0.3987	0.3998	5.993	1.961	3.3	2.4	1.2	1.7
7	333.698	1.137	333.753	1.426	0.4662	0.4690	4.397	4.394	2.6	2.2	2.0	1.6
8	213.155	55.904	213.765	55.695	0.4068	0.4044	5.531	4.171	2.5	2.3	2.3	1.9
9	213.155	55.904	212.563	55.761	0.4068	0.4037	4.977	5.520	2.5	2.3	0.7	1.2
10	212.944	54.218	213.735	54.932	0.3620	0.3619	10.875	0.097	2.4	2.3	2.9	2.0
11	213.735	54.932	214.919	55.145	0.3619	0.3606	9.300	2.194	2.9	2.0	1.0	1.7
12	31.354	−7.595	30.924	−7.785	0.4386	0.4398	6.738	1.982	4.4	2.7	1.4	1.9
13	31.354	−7.595	31.742	−7.285	0.4386	0.4373	7.124	2.186	4.4	2.7	0.7	1.3
14	211.903	52.652	210.499	52.676	0.3349	0.3345	10.363	0.763	2.6	1.6	1.8	2.0
15	35.319	−7.593	35.116	−7.837	0.3717	0.3701	4.355	2.823	4.1	2.8	1.5	2.3
16	33.368	−5.196	33.534	−5.592	0.4420	0.4453	6.226	5.517	2.7	2.6	1.9	1.7
17	32.735	−6.198	32.391	−6.456	0.4293	0.4253	6.092	6.694	3.6	2.3	1.2	1.7
18	333.624	−0.767	333.357	−0.160	0.3352	0.3337	8.064	2.919	1.3	1.3	2.9	2.3
19	31.235	−9.749	30.862	−9.414	0.3308	0.3297	6.002	2.126	0.6	1.4	3.1	2.8

Table 5.2. List of halos pairs used in our analysis. M_{200} estimates are best fit parameters using isotropic NFW model. $\sigma(M_{200})$ is the significance level of halo lensing detection.

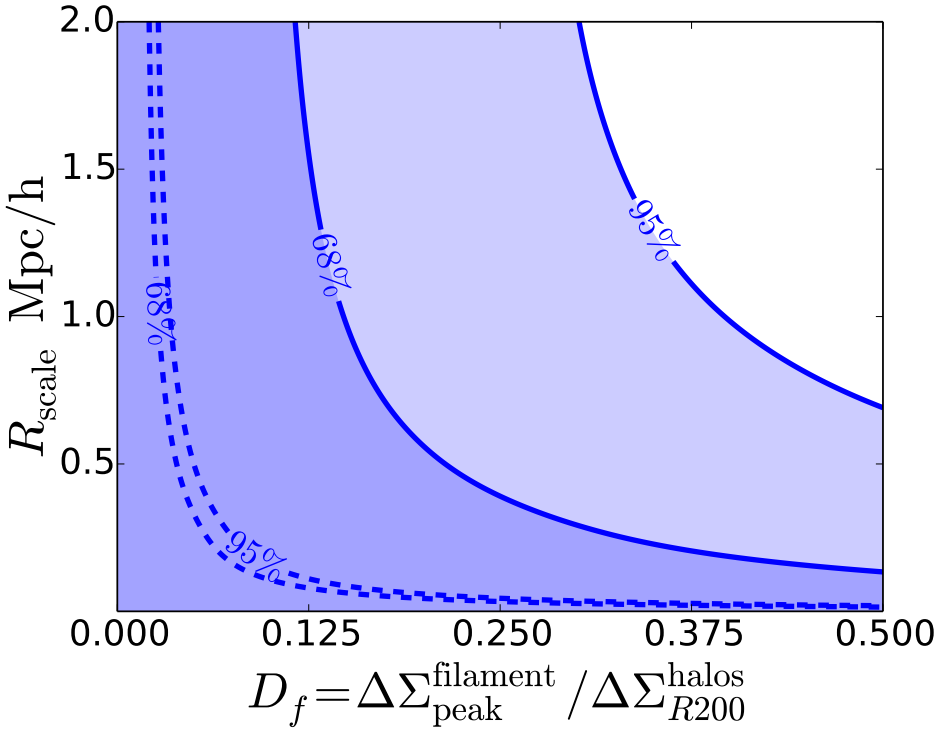

Figure 5.5. Average likelihood of filament parameters for random points null test. Shaded regions show the mean likelihood for 32 products of 19 pairs. This distribution is broad, as expected, due to no filament signal present. The fact that our best fit point lies outside the mean contours suggests that our measurement significance is not dominated by noise. Confidence intervals for a product of likelihoods of $19 \cdot 32$ measurements is shown in dashed lines. This measurement is consistent with zero for the $D_{\rm f}$ parameter, which indicates that there is not bias on this parameter due to noise and large scale structure.

result to previous measurements in the literature. For example, the A222/A223 system studied by D12, had two clusters separated by $\sim$ 11 Mpc/h, and the mean mass was $\sim 3 \cdot 10^{14}$. A simple calculation shows that, if it was measured face-on, its peak $\Delta\Sigma$ would be approximately $2 \cdot 10^{13}$ $\mathrm{M_\odot Mpc^{-2}h}$. If this configuration was used with our best fit $D_{\rm f}$, then this density contrast would be $\Delta\Sigma = 0.9\text{–}2.2 \cdot 10^{13}$ $\mathrm{M_\odot Mpc^{-2}h}$. This seems to be a good agreement, but one must remember that the uncertainty on our result is still very large.

Scale radius was measured at $R_s = 1.66^{+0.81}_{-0.54}$, and constrained at 3.1σ away from zero. Its upper limit was, however, poorly constrained, with 2σ contours not closing for the value of 4 Mpc/h. Our measurement of filament radius indicates a preference for thick filaments. This is in agreement with observations in Clampitt et al. (2014), where they

report that the data prefers filaments with radii greater than 1 Mpc/h. The upper limit is not well constrained, reaching 4 Mpc with $1-2\sigma$ confidence. This is indeed much higher than N-body simulations predict, for example, Colberg et al. (2005) found only very small number of filaments with radius $R_s > 3$ Mpc/h.

I have used a novel probabilistic approach for combining measurements from multiple cluster pairs to increase the SNR of detection. Within this method, it is very easy to change parameters describing the filaments without transforming the data itself. For example, as I expect the filament mass to depend on its environment, I used filament density contrast tied to density of halos. I also accounted for differences in projected density among filaments due to different line of sight distances between connected halos. I found both these changes to parametrisation to increase the significance of detection.

I also checked that the signal I detected can not be explained by random noise fluctuations and shear signal from large scale structures at different redshift. I performed a random points null test, simulating 32 times the data volume. This test used pairs created from random points in CFHTLenS with artificially added shear signal from dark matter halos. I found that the parameters of filaments measured from this null test are consistent with zero, with large scatter. This scatter, however, was not large enough to explain our main result, thus allowing us to conclude that the signal measured in the main result of this chapter can not solely come from LSS or random noise fluctuations.

Future lensing surveys will open possibilities for studying intercluster dark matter filaments with increased accuracy. With 100-200 halo pairs it will become possible to characterise properties of filaments as a function of their environment, for example halo masses. A comparison with number of galaxies residing in filaments, as well as hot gas, to density of filaments can also be an interesting avenue of investigation.

Chapter 6

Concluding remarks

Weak gravitational lensing is an unbiased probe of matter in the Universe and provides unique opportunities for new discoveries in cosmology and astrophysics. Measurement of galaxy shear, a main weak lensing observable, is very challenging due to the complicated forward process for obtaining galaxy images and the non-linear nature of measurement techniques. Chapters 3 and 4 of this thesis were focused on studying systematics associated with this measurement using model-fitting methods. I particularly focused on effects of noise bias and model bias. Chapter 5 presented a detection of intercluster filaments in CFHTLenS wide fields. I introduced a novel, probabilistic approach to combining measurements from multiple filaments and applied it to CFHTLenS data. In this final chapter I will review the prospects for precision cosmology with weak gravitational lensing, and focus on practical applications of calibration of noise and model bias effects. I will also outline future directions for weak lensing studies of dark matter filaments.

6.1 Shear measurement and its calibration

In Kacprzak et al. (2012) and Refregier et al. (2012) I introduced and characterized a statistical model for *noise bias* in maximum likelihood (ML) model-fitting methods. Noise bias arises due to the fact that the ML shape estimator is a non-linear function of pixel intensities, affected by additive noise. We derived analytic expressions for bias on the fitted parameters, which demonstrate that this bias scales with $1/\mathrm{SNR}^2$ and other parameters

such as galaxy size and morphology. A histogram of ellipticity estimators (see Chapter 3, Fig. 3.1) is not symmetric and its skewness creates a difference between the truth and the mean, thus causing a bias on the shear. I showed that this effect can introduce order of 5% biases on the shear and cannot be neglected for lensing surveys. We found that noise bias depends on the galaxy's true parameters, such as pre-convolved size, ellipticity, and morphology. These parameters are measured with large noise, and probably also significant bias. Any simulation-based calibration scheme will have to infer the true distributions of these parameters. These distributions will also be dependent on the specifications of a survey, and may vary with observing conditions. That may complicate the calibration process.

Another important challenge in shape measurement is the *model bias*. When a galaxy model used to fit an image is not able to capture all its complicated morphological features, shear estimates can be biased (Voigt & Bridle 2009; Bernstein 2010). Using the Shear Reconvolution Analysis technique (Mandelbaum et al. 2011), we can now evaluate the scale of this effect. In Kacprzak et al. (2013) I used a realistic bulge + disc galaxy model, and calculated the model bias using images from the HST COSMOS survey. I measured the multiplicative bias resulting from using simple models for each one out of 20000 galaxies released (Mandelbaum et al. 2011). I found that the standard deviation of multiplicative bias exceeds even the current galaxy survey requirements, while its mean is small. Additionally, I found that there is a slight dependence of model bias on the morphological type of a galaxy. As this is sufficient for current surveys' requirements, further investigation will be necessary to control this effect below the $m < 0.1\%$ level required by WFIRST and Euclid.

Since noise bias estimates have typically assumed perfect knowledge of the galaxy models, it is also important to investigate if the statistical model of the noise bias is still a good description of the problem. In (Kacprzak et al. 2013) I extended the derivation for bias on the ML estimates to the case when an imperfect model is used. Simulation using COSMOS images allowed us to quantify the model bias and its impact in the presence of noise (see Chapter 4, Fig. 4.3). We measured the strength of model-noise bias interaction finding it small enough for current galaxy surveys: with multiplicative bias $m = 0.0008 \pm 0.0009$ for the entire population and $m = -0.004 \pm 0.002$ in the worst case scenario. This effect has not yet been tested to the accuracy required by space based observing programs and more investigation will be required to assure that model and noise bias calibrations

are jointly effective.

I made successful attempts to apply noise bias calibration in a simplified setting. In Kacprzak et al. (2012) I applied a very simple calibration scheme to a single galaxy. In that calibration, we treated the noisy estimates of size and ellipticity parameters as if they were the true parameters. This, potentially, can introduce an error, but we found it to be small, and shear had been calibrated to a good precision. In (Zuntz et al. 2013) I successfully applied the calibration to the GREAT08 challenge. Noise bias calibration has also been successfully applied in practice: Miller et al. (2012) used it to correct the CFHTLenS shear catalogues. For the Dark Energy Survey, noise calibration will be done in a similar way, using precise simulations of large volumes of data. Within the next two years, the control of shear systematics has to be improved from current the $m \sim 5\%$ level to $m \sim 0.3\%$ level. The effectiveness of these calibration schemes will strongly depend on how well the noisy data is represented by the calibration training sets. For example, an approach proposed by Refregier & Amara (2014) uses a Monte Carlo Control Loops approach to creating well matched training sets. In this approach, a space of input parameters to simulations is explored. The quality of match between the simulation and survey data is assessed by comparing statistics of measured, noisy parameter distributions. Distributions which well match the survey data are than perturbed, which allows us to quantify the uncertainty on the input to the simulation.

Another promising approach to shape measurement is to use Bayesian methods. Bernstein & Armstrong (2013) and Sheldon (2014) presented a probabilistic approach, in which the shear is not calculated using, hopefully unbiased, estimators $g = \langle e \rangle$. Instead, more information is stored about each galaxy shear, and calculation of g is more complicated. That approach requires modifications to existing pipelines which use shear information, such as correlation function calculations. However, these methods have demonstrated the ability to avoid noise bias.

6.2 Using colour information to improve the accuracy of shear estimation

A good way to improve the accuracy of a measurement is to include more available data. Measurement of galaxy morphologies from high resolution images in different bands can be also useful for investigation of feasibility of *multi-band shear measurement*. The KIDS,

DES, HSC, WFIRST and LSST surveys will use multiple filter observations. Galaxies have slightly different intrinsic shapes in each band, but the amount of shearing due to lensing is the same in all bands. This fact can potentially be utilized to significantly increase the SNR of ellipticity measurement up to a factor of approximately $\sqrt{N}$, where N is the number of observed bands. In the context of noise bias this can be a very significant gain, due to the fact that the noise bias scales with $1/SNR^2$, as showed earlier. This approach has not yet been demonstrated in the literature. Using high-resolution imaging from the HST, it is possible to investigate whether using a prior in the form of a covariance of shapes in different bands can be used to do a joint estimation of shear.

A critical issue for performing lensing with a wide-filter instrument such as Euclid is the understanding of wavelength dependence of the Point Spread Function (PSF) and galaxy colour gradients. The shape of the PSF varies with wavelength for two main reasons: optics of the telescope introduce an Airy disk with size inversely proportional to the wavelength, and the Modulation Transfer Function (MTF) of the CCDs creates more diffusion for higher energy photons. The Euclid mission will use a wide optical filter, and that can introduce two potential problems. Firstly, if the Spectral Energy Distribution (SED) of a galaxy differs from the one of a star used to estimate the PSF, the shear estimate will be biased beyond the survey requirements. Cypriano et al. (2010) demonstrated methods to ensure that the PSF used for shape deconvolution matches the true PSF smearing the galaxy. Secondly, if the SED of a galaxy varies within the galaxy itself (for example, redder central bulge and bluer spiral arms), then the PSF will depend on the position in the image. Voigt et al. (2011) demonstrated that this problem can not be neglected for Euclid, and proposed a calibration scheme using large numbers of well resolved galaxy images from the HST archive. Semboloni et al. (2012) developed this approach further and showed that the HST archive will contain enough galaxies to calibrate this effect by the time Euclid is operational.

This calibration can be a very sensitive process, as the shear bias depends strongly on galaxy morphology. Precise knowledge of distribution of galaxy properties with redshift will be required. With new tools (Mandelbaum et al. 2011; Rowe et al. 2014) and ~ 60000 galaxy images from the HST it is possible to study effects related to galaxy colours. This avenue of research can potentially have a high impact on the field of shear measurements.

6.3 Weak lensing by dark matter filaments with CFHTLens and future surveys

Weak lensing by dark matter filaments is a relatively new area of research. A range of reasons make the measurement of this signal very challenging. Firstly, the signal from filament lensing is very faint; its strength is close to an order of magnitude lower than lensing by dark matter halos. Current lensing data sets often do not have a necessary number density of sources. Detection of individual pairs is very difficult, although Dietrich et al. (2012) has identified a filament between two massive clusters using Subaru data, with 25 galaxies per square arcmin (priv. comm). Next to a good number of lensing galaxies, this filament had a high projected surface mass density due to to its specific orientation: it was observed "through the barrel", which boosted the density contrast by a factor of 3. Combining measurements from multiple pairs can increase the SNR of detection. However, one must be sure that the pairs selected for such a sample indeed have a filament between, otherwise the measurement can be biased low. Without pre-selection using other data sources (for example X-ray or galaxy counts), it is hard to quantify this effect. Clampitt & Jain (2014) detected signal from filament lensing using more than 200,000 LRG pairs in SDSS, finding the level of $\Delta\Sigma \approx 1 \cdot 10^{12} \mathrm{M}_{\odot}\mathrm{h/Mpc}^2$. Compared to the result from Dietrich et al. (2012), where if the filament was seen face-on, $\Delta\Sigma \approx 2 \cdot 10^{13} \mathrm{M}_{\odot}\mathrm{h/Mpc}^2$, there is an order of magnitude difference. This difference is not unexpected, as Clampitt & Jain (2014) used all LRGs independent of their mass, and probably also included connections which contained voids instead of filaments.

In Chapter 5 I presented a detection of filament lensing using the most massive halos in CFHTLenS wide fields. I chose the halos using the LRG sample from BOSS. As the strength of a filament is likely to be related to density of connected halos, I tied the halo density to filament density using the relation $D_{\mathrm{f}} = \Delta\Sigma^{\mathrm{filament}}_{\mathrm{peak}}/\Delta\Sigma^{\mathrm{halos}}_{R200}$. The radius of a filament was controlled by a scale radius parameter R_{s}. I used a parametric model with two halos and a filament and fitted this model to shear postage stamps cut out from CFHTLenS data. For 19 pairs of filaments connecting halos with at least one of them more massive than $2 \cdot 10^{14} \mathrm{M}_{\odot}/\mathrm{h}$, we measured $D_{\mathrm{f}} = 0.50 \pm 0.13$ and $R_{\mathrm{scale}} = 1.66^{+0.81}_{-0.54}$. For two halos with masses of $M_{200} = 3 \cdot 10^{14} \mathrm{M}_{\odot}/\mathrm{h}$, that corresponds to peak density contrast of a filament of $\Delta\Sigma^{\mathrm{filament}}_{\mathrm{peak}} = 10^{13}\ \mathrm{M}_{\odot}\mathrm{Mpc}^{-2}\mathrm{h}$. I confirmed that the fluctuations due to shape noise, cosmic shear and other large scale structures in the line of sight do

not have enough power to mimic the filament signal for 19 pairs.

By varying the parameters of halo pairs coming into our sample, I found that the value of D_f parameter can vary significantly. For example, restricting halo masses to be both greater than $M_{200} > 2.5 \cdot 10^{14}$ $M_{\odot}h/Mpc$, density parameter D_f increased by a factor of 2. Conversely, when longer pairs were considered, from $10 - 20$ Mpc/h , this parameter halved. Although it is expected that our linear scaling of filament density with density of halos is not correct, I point out that this variation can be due to contamination of our sample by pairs which are disconnected. Given that I selected pairs only using their halo properties (mass, distance and proximity to other halos), I may have included empty regions in the sample. This contamination is an important issue for future filament characterisation studies and can be addressed, as mentioned before, by using non-lensing data, for example, X-ray.

Detection was just a first step in learning about matter distribution in intercluster filaments using weak lensing. Many extensions of this research are possible. For example, it would be very interesting to use other mass-mapping techniques to explore the environment of pairs of halos selected for the filament sample. One could study more about orientation of halo ellipticity with respect to filaments. Alternatively, a more complex model including ellipticity could be used for halos. This would increase the computing time significantly, for grid-based optimisation. MCMC techniques can also be used, but multiplying likelihood surfaces using MCMC samples can pose challenges; before taking a product, probability density has to be reconstructed from samples. This can be a difficult process, especially when measuring signals which are on the edge of a box prior. Although several approaches have been proposed for density estimation with bounded domain, using cyclic or reflective methods, in practice I found them hard to apply to this problem in a way which would not distort the result, when compared to grid-based optimization. Nevertheless, applying grid-based optimization to an 8 parameter model may be tractable using parallel computing, if likelihood evaluation will be distributed on many cores.

Another interesting extension of filament science is to cross-correlate filament regions with the same regions in other data sets. These data sets can include thermal Sunyaev-Zel'dovich (tSZ) effect measurements from Planck (Planck Collaboration 2012), X-ray and measurements of galaxy counts. However, more investigation is needed to assess the feasibility of such measurements.

Upcoming wide field shear surveys will provide an order of magnitude larger shear

catalogues. This opens up a possibility of finding more filament pairs, given a good spectroscopic LRG sample is available in that area. Increasing the number of pairs by 10 would allow us to attempt to characterise the filament signal's dependence on its environment, for example, as a function of connected halo masses. Properties of filaments as a function of their length is also very interesting. Projects such as the Dark Energy Survey and Hyper Suprime Cam will provide 5000 and 2000 square degrees of lensing data respectively in upcoming years. If bright cluster galaxies found in the centre of massive clusters detected in these surveys will be followed up by spectroscopy, interesting new avenues of filament research will be opened.

Appendix A

Im3shape: A maximum-likelihood galaxy shear measurement code for cosmic gravitational lensing

This appendix is a fragment of a paper by Zuntz et al. (2013), to which I contributed very significantly and was therefore second author. I developed and tested core IM3SHAPE modules: convolution, optimization, image rendering. I also designed and applied calibration using simulations. In this appendix I present sections considering noise bias calibration only. Text in this section was written in collaboration with other authors.

A.1 Noise Bias Calibration

At the noise levels in the RealNoise_Blind Challenge we do expect that we need to correct for the noise bias effect described in Refregier et al. (2012), Kacprzak et al. (2012) and references therein. We describe our approach to noise bias calibration in this section.

In Kacprzak et al. (2012) we ran IM3SHAPE on noisy simulations with a range of input parameters to find the behaviour of the noise bias with galaxy properties. We then attempted to remove the bias by matching the measured galaxy properties to the simulation input parameters to predict the expected bias, which we subtracted off. We ran this method on branch 3 of our GREAT08 analysis, which has the same galaxy model

as IM3SHAPE, and found $Q = 166$, with $m = 3.0 \times 10^{-2}$ and $c = 7 \times 10^{-4}$ which is nearly satisfactory for current experiments but well outside our requirements for upcoming experiments.

This straightforward correction fails to do better because the observed properties are themselves noisy (and thus noise-biased), and so our bias estimate is itself biased. In practice, we have found it ineffective to fully calibrate galaxy ellipticity measurements one-by-one using their individual properties, and turn instead to a population-based approach in which a complete class of galaxies is calibrated collectively.

The approach that we take is discussed in Appendix A.3. Briefly, we use the deep data from the low-noise part of the challenge to provide distributions of the galaxy properties, and then perform simulations to compute the mean multiplicative and additive biases for those properties. We then apply these mean biases to the shear estimates from each set.

To predict biases for a given parameter set we simulate galaxies at grid points with scale radii $r_d = \{1, 2.5, 5\}$ and bulge flux fractions $F = \{0, 0.5, 1\}$, and with ellipticity from $e_1 = -0.8$ to 0.8. At each grid point we fit a cubic polynomial to the bias $\hat{e} - e_{\rm true}$ as a function of $e_{\rm true}$, and store the coefficients. To get the bias for intermediate parameters we use a Gaussian radial basis function interpolator. Since the simulations are costly we simulate only at $S/N = 20$, and extrapolate to other values using the result from Refregier et al. (2012) and Kacprzak et al. (2012) that the bias scales as the inverse of the square of the signal-to-noise.

There are several factors about the structure of GREAT08 that are unrealistic in this context. Firstly, not every branch in the main sample has a corresponding low-noise version, whereas in real surveys it should be possible to match populations with more care in most cases. This would tend to make noise calibration harder on GREAT08. Secondly, the sizes of galaxies in GREAT08 are not drawn from a population - they have a single true value. This would tend to make calibration more effective than is realistic. We also note that the information needed to fully perform the calibration in this way was not public at the time of the challenge.

Where possible we match branches using the corresponding low-noise branch; where none exists we use the nearest approximation. See Appendix A.3 for more details.

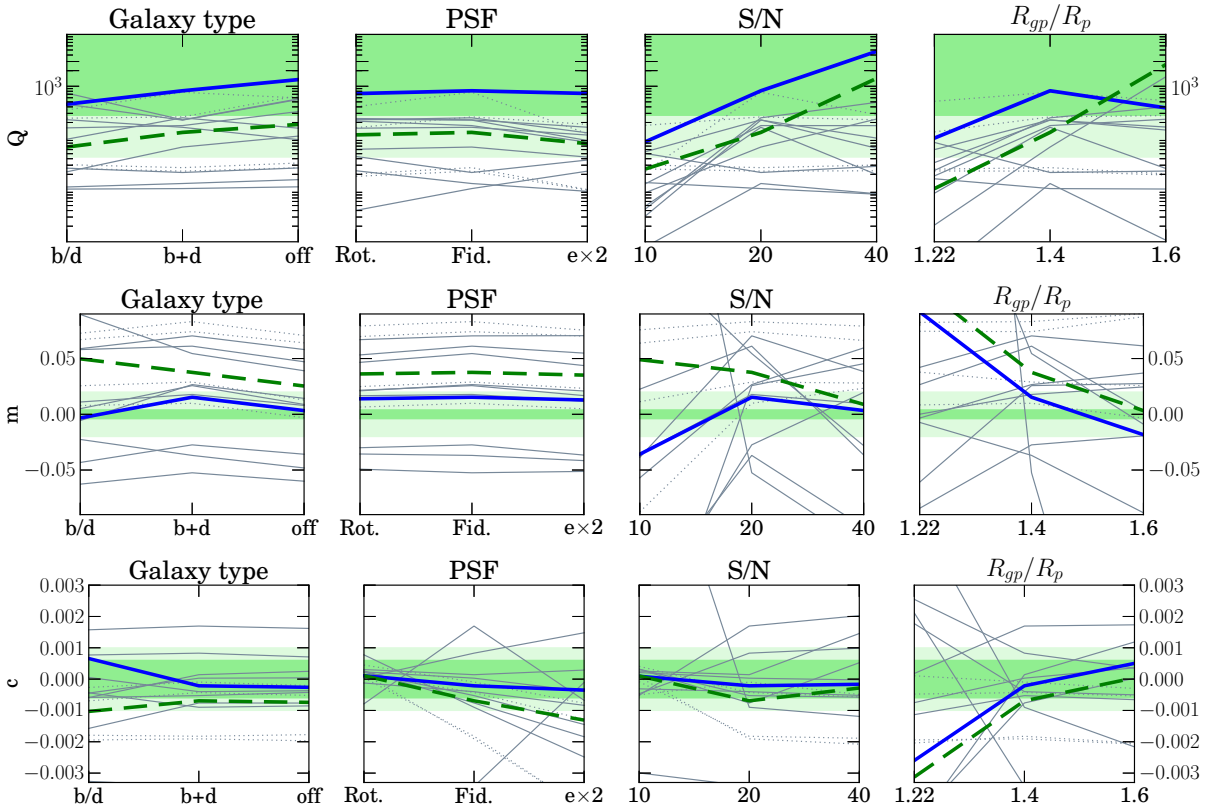

Figure A.1. The Quality factor score Q, multiplicative bias m, and additive bias c obtained in the RealNoise-Blind section of the Great08 challenge. Each column shows scores with a different variation from the fiducial galaxy parameters. The thick solid blue line is our primary result, the IM3SHAPE scores following the matched-population noise bias calibration described in the text. The dashed green line is the score without any calibration. The green bands are the requirements for current and upcoming surveys. The thin grey lines are other entrants at the time of the challenge, with dots indicating stacking methods. The galaxy types are bulge *or* disc (50% of each; labelled b/d), co-elliptical bulge plus disc ($b + d$) and bulge plus disc with offset centroids (off). The PSF values were fiducial (fid), rotated by 90°, and double ellipticity($e \times 2$).

A.2 Scores

Q-factor scores and m and c values for the pre- and post-noise bias calibration results are shown in figure A.1. Even before the calibration the bias on the $S/N = 40$ galaxies was small enough that we achieved a quality factor $Q > 1000$, with m and c small enough for upcoming surveys. At lower signal to noise, as we expect, the pre-calibration Q factor drops to below 100.

After the noise bias calibration procedure described above the Q factors and the c values reach the levels required for upcoming experiments for most of the branches of the challenge. The results are stable with PSF type. The m values are somewhat more variable but, depending on the exact branch, can reach the required levels. In particular the value is good in the b/d branch, where where we expect no model bias.

Our most encouraging result is the stability across the galaxy type branches. The only branch for which we use the correct model is the first galaxy type, where the simulation model is single-component bulge or disc. For all other branches we expect some model somewhat like the ones discussed in section 4.2, but as in that section the effect is not critical for current surveys.

The branches for small and noisy galaxies are clearly more problematic, and we would be forced to remove such galaxies from any current analysis. We believe the extrapolation fails badly in the high-noise case because of a slowdown in the magnitude of the bias with ellipticity, which does not follow the $b \sim (S/N)^{-2}$ relation that we used. Very noisy galaxies have much broader scatter in measured ellipticity, and when these values start to push up to the edge of the space, at $|e| = 1$, this acts as an $m < 0$ bias which partially counteracts the usual $m > 0$ one. Missing this effect can lead to extrapolated values over-correcting the bias.

A.3 Noise bias calibration using deep data

Noise bias is the difference between the expectation of the maximum-likelihood result found by a model fitter like IMSHAPE, which is a biased estimator, and the true underlying value, coming from the non-linear mapping between parameters and image. Its typical value is a few percent, and we need to remove it to reach our target accuracy levels.

The size of the bias depends sensitively on the true parameters of the galaxy, and if these were known we could remove noise bias completely. However, we have access only

to noisy estimates of these parameters and therefore our estimate of the noise bias is itself noisy, and biased. We can think of this as a *bias-on-bias* problem, and we find that failing to account for it means we significantly miss our targets.

We can get around this problem if there is a subset of our observational data that is deeper than the bulk of our sample, and therefore of greater signal-to-noise, to a degree sufficient that it has negligible noise bias. This is often the case in real surveys that seek to detect high redshift supernovae, and it is approximated in the GREAT08 challenge with LowNoise_Blind data set (although the galaxy types in this set do not match those in the main sample exactly). In this case we do not calibrate each galaxy individually, but instead find a mean bias for the population and apply it en masse.

Biases $m(\theta)$ and $c(\theta)$ will afflict each of our galaxies, depending on the true galaxy properties θ. We can calculate mean values of this bias $\hat{m}$ and $\hat{c}$, and apply these evenly to all the galaxies, such that the mean galaxy ellipticity (which is our goal) is correct.

The mean of the bias across the population is given by:

$$\hat{m} = \int m(\theta)p(\theta)\mathrm{d}\theta. \tag{A.1}$$

We find the distribution $p(\theta)$ using fits to the deep data, and $m(\theta)$ using simulations - for each point θ_p in a grid in the parameter space we simulate many galaxies and determine the value $m(\theta_p)$. We can then interpolate between these values to do the integral, and finally apply the mean $\hat{m}$ to all the galaxy estimates.

A similar process is used for the c bias, except that c varies directly with the PSF. This dependence stems from the fact that PSF ellipticity can directly mimic the lensing signal. Linear dependence of shear systematics was shown in, among others, (Paulin-Henriksson et al. 2008) and used as a model in systematics analysis **??**.

We calculate the $\hat{c}$ assuming a fiducial PSF with ellipticity e_0^{psf} aligned with the e_1 direction. The applied value to apply to each galaxy is then:

$$c_1 + ic_2 = \hat{c} \cdot \frac{e^{\mathrm{psf}}}{e_0^{\mathrm{psf}}} \cdot e^{i\theta_{\mathrm{psf}}}. \tag{A.2}$$

Appendix B

Mass and galaxy distributions of four massive galaxy clusters from Dark Energy Survey Science Verification data

This appendix contains my contribution to paper by Melchior et al. (2014) describing mass and light distribution measured from the Dark Energy Survey early science verification data. I have provided simulations and calibration for shear catalogue used in this work. The text was written in collaboration with other authors. The following paragraph is taken from Melchior et al. (2014), section 4.2.1.

Noise-bias calibration

Shape measurements are affected by a prominent bias when the galaxy images become noisy (e.g. Massey et al. 2006). This is a consequence of the observable, the galaxy ellipticity, being non-linearly related to the flux in each pixel and applies to model-fitting methods and moment-based measures of the ellipticity alike (Melchior & Viola 2012; Refregier et al. 2012; Kacprzak et al. 2012).

To calibrate IM3SHAPE's response to noise bias we simulate mock galaxies, using the

GalSim[1] (Rowe et al. 2014) framework. In particular, we adopt the methodology of Mandelbaum et al. (2011) and degrade high-resolution and high-significance images from COSMOS to the DECam resolution and magnitude limit. We approximate the coadd PSF by a circular Moffat (1969) profile with seeing values $\in [0.7, 0.8, 0.9]$ arcsec, spanning the range of most of our observing conditions. Applying exactly the same cuts as for shape catalogues from the coadd images, we have verified that both magnitude and size distributions of the simulated galaxies closely match the observed ones. Adding an artificial shear γ of order 5%, we can infer the shear response

$$m_{\rm n} \equiv \frac{\partial \langle \epsilon \rangle}{\partial \gamma} \tag{B.1}$$

as a function of the signal-to-noise ratio `SNR` and `FWHM_RATIO`. The result is shown in B.1. At high `SNR`, the shear can be measured in an unbiased fashion for all galaxy sizes, whereas the noise bias gets progressively worse for lower `SNR`, scaling roughly as $\texttt{SNR}^{-2}$, consistent with findings of Bernstein & Jarvis (2001). It is counter-intuitive that the smallest galaxies show the least amount of bias. Also, at very low `SNR` the larger galaxies show an intriguing upturn. We interpret both as higher order effects of the noise bias. According to Kacprzak et al. (2012), only even orders of `SNR` can appear in the noise-bias relation, therefore we attempt to parametrise the dependence with the following polynomial model,

$$m_{\rm n} \approx c_0 + c_2 \texttt{SNR}^{-2} + c_4 \texttt{SNR}^{-4}, \tag{B.2}$$

whose best-fit parameters are listed in B.1. The applied noise-bias corrections are taken from these fits in each of the size bins.

[1] `https://github.com/GalSim-developers/GalSim`

FWHM_RATIO	c_0	c_2	c_4	N_{gal} [%]
$\in [1.20, 1.30]$	1.010 ± 0.012	-3.9 ± 3.6	-64 ± 129	24.2
$\in [1.30, 1.40]$	1.001 ± 0.012	-17.2 ± 3.4	$+229 \pm 127$	19.0
$\in [1.40, 1.50]$	1.019 ± 0.013	-30.2 ± 3.5	$+587 \pm 130$	15.3
$\in [1.50, 1.60]$	1.028 ± 0.016	-30.5 ± 4.1	$+498 \pm 150$	12.0
$\in [1.60, 1.70]$	1.019 ± 0.018	-28.9 ± 4.8	$+438 \pm 174$	9.8
$\in [1.70, 1.80]$	1.042 ± 0.021	-28.3 ± 5.4	$+326 \pm 193$	7.3
$\in [1.80, 1.90]$	1.014 ± 0.024	-23.4 ± 6.2	$+271 \pm 214$	5.0

Table B.1. Best-fit parameters of B.2 to the simulated data from B.1. The last column indicates the percentage of all galaxies with shape measurements in any band to fall into the given bin, averaged over all fields and *riz* filters.

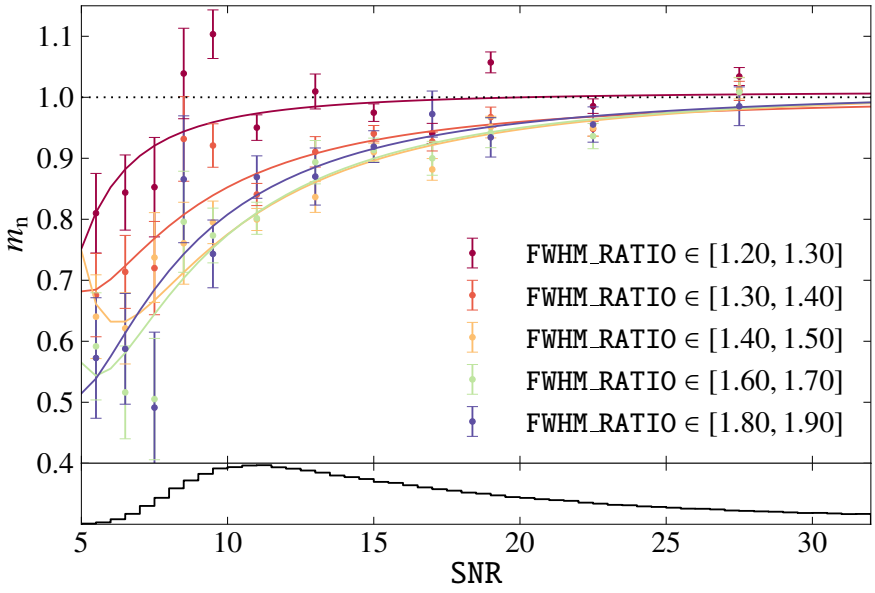

Figure B.1. Noise bias on the multiplicative term m_{n} of the shear response as a function of IM3SHAPE's SNR for different values of the galaxy FWHM_RATIO. The solid lines are even-order polynomial fits to the data (cf. B.2). The bottom panel shows the SNR distribution of galaxies with shape measurement from any single-filter coadd image, averaged over all fields and *riz* filters.

This page was intentionally left blank

Appendix C

GalSim: The modular galaxy image simulation toolkit

In this appendix I present a section in paper by Rowe et al. (2014). I have designed and implemented validation schemes for image rendering methods in GalSim. These schemes tested the FFT-based and photon shooting based drawing engines (C.1.1) as well as reconvolution procedure (C.1.2). I have run and analysed results of these tests and provided the data to the authors. The only exception is the test showing the dependence of rendering accuracy on Sérsic index, which I keep in this section for completeness. Text describing these results was written in collaboration with other authors. Section numbers marked with §correspond to sections in Rowe et al. (2014).

C.1 Numerical validation

In this Section we describe the investigations that were undertaken to validate the accuracy of GalSim image simulations. Although an exhaustive validation of the rendering of every combination of galaxy/PSF profiles and observing conditions is impractical, certain key aspects of GalSim performance are shown here. Emphasis is placed on confirming that GalSim meets the stringent requirements on image transformations for lensing shear and magnification simulation.

In particular, our metric for validating that the rendered images are sufficiently accurate is based on measurements of the size and ellipticity of the rendered profiles, calculated using the adaptive moment routines described in §8.

We define the following "STEP-like" (see Heymans et al. 2006) models for the errors in the estimates of object ellipticity g_1 and g_2 and size σ:

$$\Delta g_i = m_i g_i + c_i, \tag{C.1}$$

$$\Delta\sigma = m_\sigma \sigma + c_\sigma, \tag{C.2}$$

where $i = 1, 2$. The method of estimating the errors Δg_i and $\Delta\sigma$ varies for each of the validation tests described below, but a common component is adaptive moments estimates of rendered object shapes from images (see §8). We will use the formulae above when describing the nature of the errors in each test.

As discussed in Mandelbaum et al. (2014), a well-motivated target for simulations capable of testing weak lensing measurement is to demonstrate consistency at a level well within the overall requirements for shear estimation systematics set by *Euclid* (e.g. Cropper et al. 2013; Massey et al. 2012): $m_i \simeq 2\times10^{-3}$ and $c_i \simeq 2\times10^{-4}$. Such values also place conservative requirements on galaxy size estimation, as the signal-to-noise expected for cosmological magnification measurements has been estimated as $\lesssim 50\%$ relative to shear (e.g. Van Waerbeke 2009; Schmidt et al. 2011; Duncan et al. 2013).

Only if these stringent *Euclid* conditions are met comfortably will simulations be widely usable for testing weak lensing shear estimation, and other precision cosmological applications, in the mid-term future. For each validation test we therefore require that GalSim produce discrepancies that are a factor of 10 or more below the *Euclid* requirements, i.e. $m_x < 2\times10^{-4}$, $c_x < 2\times10^{-5}$, where $x = 1, 2, \sigma$ corresponding to g_1, g_2 and σ, respectively.

The tests in this Section were conducted over a period of extended validation of the GalSim software between July 2013 and the time of writing this paper. During this time period, corresponding approximately to versions 1.0 and 1.1 of GalSim, the routines for rendering objects did not change significantly (except where modifications were found necessary to meet the validation criteria on m_x and c_x defined above).

C.1.1 Equivalence of DFT rendering and photon shooting

One of the principal advantages of the photon shooting method (see §6.3) is that the implementations of the various transformations described in §4.4 are very simple. Photons are just moved from their original position to a new position. Convolutions are similarly straightforward. On the other hand, DFT rendering (see §6.2) needs to deal with issues such as band limiting and aliasing due to folding (cf. §6.2.1).

Thus a powerful test of the accuracy of our DFT implementation is that the the two rendering methods give equivalent results in terms of measured sizes and shapes of the rendered objects. An unlikely conspiracy of complementing errors on both sides would be required for this test to yield false positive results.

Of all the objects in Table 3, Sérsic profiles are the most numerically challenging to render using Fourier methods. Especially for $n \gtrsim 3$, the profiles are extremely cuspy in the centre and have very broad wings, which means that they require a large dynamic range of k values when performing the DFT. They thus provide a good test of our choices for parameters such as `folding_threshold` and `maxk_threshold` (see §6.4) as well as general validation of the DFT implementation strategies.

For our test, we built `Sersic` objects with Sérsic indices in the range $1.5 \leq n \leq 6.2$. The half-light radii and intrinsic ellipticities $|g^{(s)}|$ were drawn from a distribution that matches observed COSMOS galaxies, as described in Mandelbaum et al. (2014). The galaxies were then rotated to a random orientation, convolved with a COSMOS-like PSF (a circular `Airy` profile), and then rendered onto an image via both DFT and photon shooting.

The error estimates were taken to be the difference between the adaptive moments shape and size estimates from the two images:

$$\Delta g_i = g_{i,\mathrm{DFT}} - g_{i,\mathrm{phot}} \tag{C.3}$$

$$\Delta \sigma = \sigma_{\mathrm{DFT}} - \sigma_{\mathrm{phot}} \tag{C.4}$$

For each galaxy model, multiple trials of the photon-shooting images were made, each with very high S/N to avoid noise biases (10^7 photons shot per trial image). The mean and standard error of Δg_i and $\Delta\sigma$ from these trials were used to estimate values and uncertainties for $m_{x,\mathrm{DFT}}$ and $c_{x,\mathrm{DFT}}$ using standard linear regression.

Differences between shape and size estimates are illustrated in Fig. C.1, for $n = 1.5$

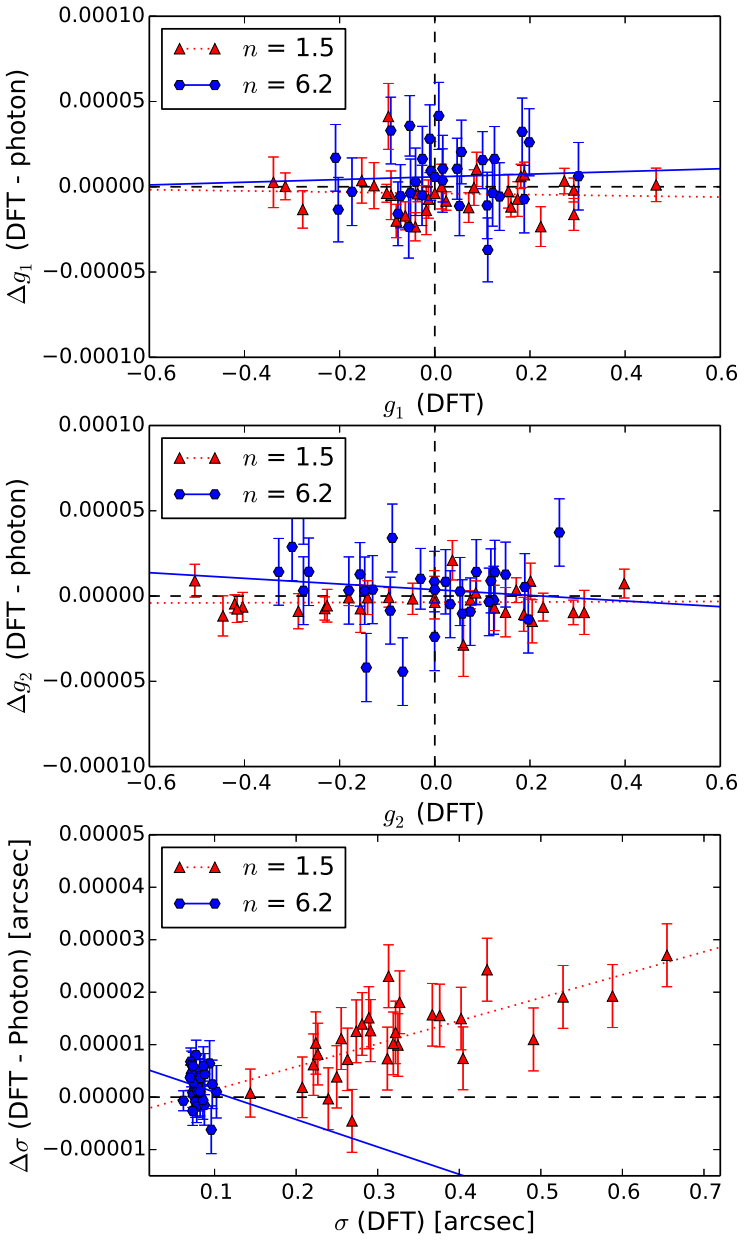

Figure C.1. Difference between measured shears (upper panel: g_1; central panel: g_2; lower panel: σ) for Sérsic profiles simulated using the photon-shooting and DFT rendering methods, plotted against the (shot noise free) shear and size measured from the DFT image. Results are shown for 30 galaxies with realistic size and shape distribution, and Sérsic index values $n = 1.5, 6.2$. (Note that the 'peakiness' of the high-n profiles results in their low σ estimates.) The best-fitting lines are shown, and estimates of the slopes $m_{x,\mathrm{DFT}}$ for these and other values of n are plotted in Fig. C.2.

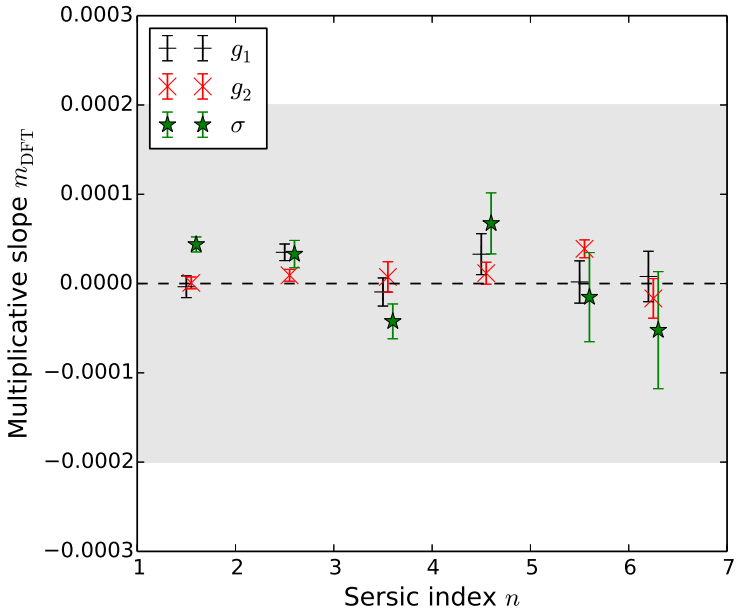

Figure C.2. Estimates of $m_{1,\mathrm{DFT}}$, $m_{2,\mathrm{DFT}}$ and $m_{\sigma,\mathrm{DFT}}$, corresponding to rendering-induced discrepancies in ellipticity g_1, g_2 and size σ, respectively, as a function of Sérsic index n. These slope parameters are defined by Eqns. C.1 and C.2 for the differences between measurements from DFT and photon-shooting-rendered images of Sérsic profiles. The shaded region shows the target for GalSim based on not exceeding one tenth of weak lensing accuracy requirements for Stage IV surveys such as *Euclid* (see §9).

and $n = 6.2$. Fig. C.2 shows derived estimates of $m_{x,\mathrm{DFT}}$ for these and other Sérsic indices tested. Tolerances are met on m-type biases, although discrepancies in ellipticity can be seen to increase somewhat as n increases. It was found that c-type additive biases were consistent with zero for all n indices.

Fig. C.3 shows results from a high-precision investigation of $m_{x,\mathrm{DFT}}$ as a function of `GSParams` parameters (see §6.4), using a randomly selected sample of 270 galaxies from COSMOS at each parameter value and large numbers of photons. Each galaxy was generated in an 8-fold ring test configuration (Nakajima & Bernstein 2006) to further reduce statistical uncertainty. The plot in Fig. C.3 shows the impact of increasing the `folding_threshold` parameter: as expected, the rendering agreement decreases as `folding_threshold` increases, and the representation of object size is most affected. Analogous results were achieved for many of the parameters discussed in §6.4, and the default `GSParams` parameters were found to give conservatively good performance in all tests.

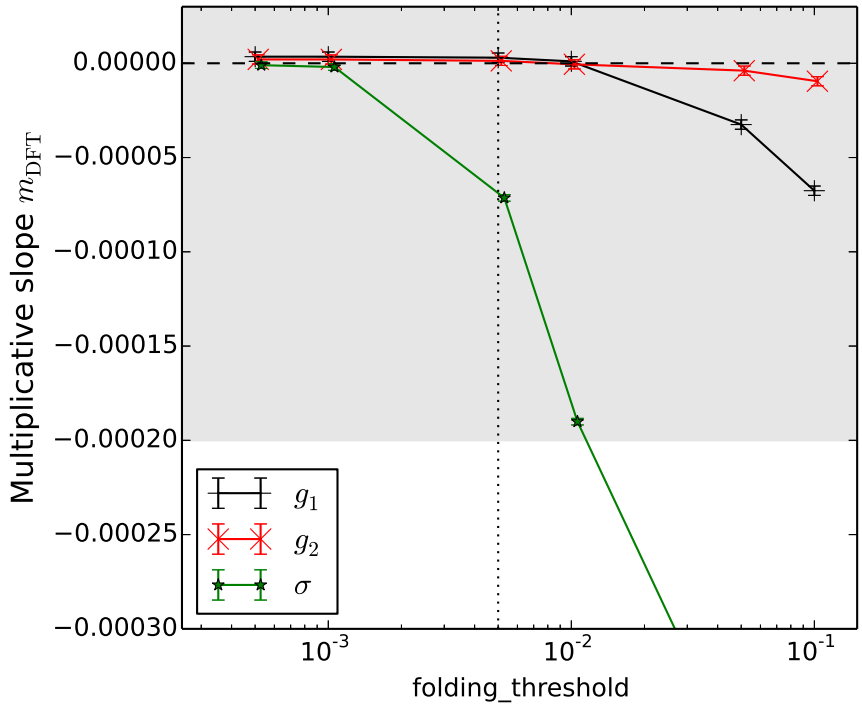

Figure C.3. Estimates of $m_{1,\rm DFT}$, $m_{2,\rm DFT}$ and $m_{\sigma,\rm DFT}$, corresponding to rendering-induced discrepancies in ellipticity g_1, g_2 and size σ, respectively, as a function of the `GSParams` parameter `folding_threshold`. Each point shows the average from the randomly-selected sample of 270 unmodified COSMOS galaxy models described in §9.1. The rendering parameter `folding_threshold` is described in §6.4 and takes a default value of 5×10^{-3}, indicated by the dotted line. As for Fig. C.2 these parameters are defined by the model of Eqns. C.1 and C.2. The shaded region shows the target for GalSim based on not exceeding one tenth of weak lensing accuracy requirements for Stage IV surveys (see §9).

C.1.2 Accuracy of reconvolution

As a final demonstration of GalSim high precision operation, we tested that we can accurately apply the reconvolution algorithm of §6.5 (Mandelbaum et al. 2011). The aim is to represent the appearance of a test object following an applied shear $g_{\rm applied}$, when viewed at lower resolution.

This test was carried out using `Sersic` profiles convolved by a known COSMOS-like PSF (a circular `Airy` profile), rendered at high resolution (0.03 arcsec/pixel). These images, along with images of the PSF, were then used as inputs to initialize `RealGalaxy` objects, mimicking the use of real COSMOS galaxy images. In the usual manner these objects were sheared and reconvolved by a broader (ground-based or Stage IV space-based

survey) PSF, then rendered at lower resolution.

Because of the use of an underlying parametric Sérsic profile, the rendering of which has been validated in §9.1, we can also render the convolved, sheared object *directly* at lower resolution to provide a reference for comparison. We quantify any error in the effectively applied shear due to the reconvolution process as $m_{i,\mathrm{reconv}}$ and $c_{i,\mathrm{reconv}}$, defined according to Eqn. C.1.

The test was done for 200 profiles whose parameters were selected from the real COSMOS galaxy catalogue described in Mandelbaum et al. (2014), using random galaxy rotations in an 8-fold ring test configuration (Nakajima & Bernstein 2006).

Since galaxies with different light profiles might be more or less difficult to accurately render using reconvolution, we must consider not only the mean values of m and c, but also investigate their ranges, which could identify galaxy types for which the method fails to work sufficiently accurately even if it is successful for most galaxies.

Fig. C.4 shows the standard deviation of $m_{i,\mathrm{reconv}}$ as a function of the `folding_threshold` parameter described in §6.4. Near the GalSim default value of 5×10^{-3}, our requirement $m_{i,\mathrm{reconv}} < 2 \times 10^{-4}$ is met comfortably in the ensemble average. Across the sample of 200 COSMOS galaxies a small fraction (3/200) exceeded our requirement for the default `folding_threshold` value for $m_{2,\mathrm{reconv}}$. However, we do not believe that this represents enough of a concern to change the default `GSParams` settings. Provided that a representative training set of galaxy models (such as the COSMOS sample), of sufficient size, is used, the variation in $m_{i,\mathrm{reconv}}$ seen in Fig. C.4 should not prevent simulations using the reconvolution algorithm from being accurate to Stage IV requirements for weak lensing.

If greater accuracy is required, users wishing to reduce the impact of these effects can modify the values of the `GSParams` according to their needs. In this case, reducing `folding_threshold` by a factor of 10 brings $m_{2,\mathrm{reconv}}$ within requirements for all 200 galaxies tested. Additive biases $c_{i,\mathrm{reconv}}$ were found to be extremely small (and consistent with zero) in all cases.

These results show that the approximations inherent in the reconvolution process do not significantly interfere with GalSim ability to render accurate images suitable for weak lensing simulations, for a realistic range of galaxy profiles drawn from COSMOS.

C.1.3 Limitations

While results presented in this Section are encouraging, with the default settings providing accuracy that comfortably exceeds our requirements by a factor of 5–10 in many cases, it must be remembered that no set of tests presented in this article could be sufficient to positively validate GalSim performance for all possible future applications. Users of GalSim are strongly advised to conduct their own tests, tailored to their specific requirements.

One specific caveat worthy of mention is the adoption of a circular PSF for all tests presented. A circular `Airy` was chosen as a simple approximation to the PSF found in COSMOS and other *HST* images. GalSim makes no distinction between those objects describing PSFs and those describing galaxies when rendering convolutions of multiple profiles. However, while unlikely, it is possible that a subtle bug or other coding issue might only be activated for cases where *both* galaxy and PSF break circular symmetry.

Another caveat is that we only used a particular set of COSMOS galaxies for the training sample. It is plausible that galaxy models drawn from a population with a different redshift distribution to the COSMOS sample, or imaged in a filter other than F814W, might have sufficiently different morphological characteristics to fail the rendering requirements adopted in this work.

In many cases, therefore, users may find it necessary to modify the tests presented here, especially where the inputs and requirements of their analyses differ significantly from the assumptions presented here. Some of the tests in this Section will hopefully serve as a useful starting point for these investigations.

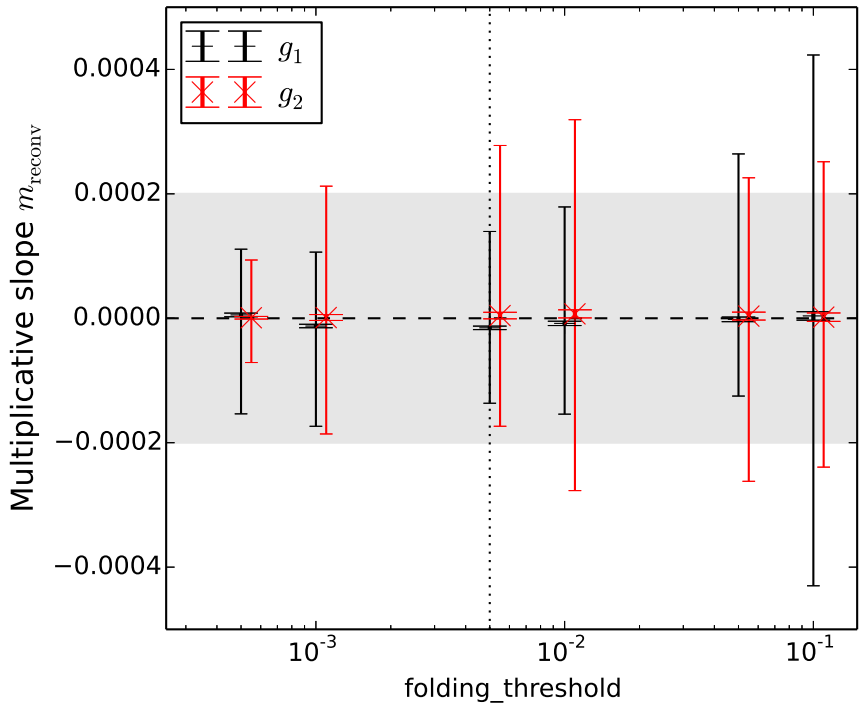

Figure C.4. Multiplicative slope $m_{i,\mathrm{reconv}}$ for the reconvolution test of §C.1.2, for g_1 and g_2, as a function of the `folding_threshold` parameter (cf. §6.4). The exterior error bars show the full range of values for the 200 models tested, and the points and interior error bars show the mean and standard error. The shaded region shows the target for GalSim based on not exceeding one tenth of weak lensing accuracy requirements for Stage IV surveys (see §9). The default value of `folding_threshold` is indicated by the dotted line.

This page was intentionally left blank

Acknowledgements

First and foremost, I would like to give my thanks to my supervisor, Prof. Sarah Bridle. Your incredible passion and grit was an inspiration for me throughout four years of study. Practical advice, as well as general discussions, were invaluable for completion of work in this thesis. I thank you for enabling me to speak at many meetings, collaborating with other scientists, and take on leadership roles in the DES collaboration; you have always encouraged me to reach higher. I value and admire your incredible leadership; it has been a great opportunity to be able to learn from you. You have made my transition from machine learning to cosmology very smooth.

I would also like to thank prof. John-Shawe Taylor. I greatly appreciate our rare, but significant, conversations. It's been very refreshing indeed to get advice on my work from machine learning point of view. I hope for further collaboration in the future.

I would like to give thanks to my collaborators: Barnaby Rowe; for incredibly useful practical advice, Joe Zuntz, for great collaboration on IM3SHAPE project (among others), Michael Hirsch, for many inspiring chats; Alex Refregier and Adam Amara; for discussions about nitty-gritty of model fitting and Mike Jarvis for help with work on DES simulations. I also thank Catherine Heymans for incredible enthusiasm and tons of good advice on the filaments project, as well as Bhuv Jain and Joseph Clampitt for many great chats on filaments. It's been a great pleasure to work with you; I feel very lucky to be able to learn from such an exceptional group of scientists.

Thank you all, Group A, for creating unique atmosphere at UCL. I especially thank my friends and office mates, Luke, Silvia, Camilla, Foteini, Roger, Cris, Patrick, Stephanie, Niall, Emily, Farah, Alex, Mikako, Boris, Morgan, Donnacha and many others. It has been an unforgettable time.

I have been lucky to have excellent teachers during my school education. I especially thank Stanislaw Lis, Bożena Szalewicz and Danuta Matusiak for developing my passion for mathematics and physics.

Finally, biggest thanks go to my parents and family. Thank you Mum and Dad, for everything. Your continuous support during my PhD studies was invaluable. None of this would be possible without your love, patience and encouragement.

I acknowledge support from the European Research Council in the form of a Starting Grant with number 240672. I thank Gary Bernstein for suggesting the calibration approach and for many fruitful discussions. I thank Alexandre Refregier, Adam Amara, Mike Jarvis, Catherine Heymans, Cris Sabiu and Caroline Pung for helpful discussions. I thank Mike Jarvis, Rachel Mandelbaum, Gary Bernstein and other developers of GalSim software for excellent support. The authors acknowledge the use of the UCL Legion High Performance Computing Facility (Legion@UCL), and associated support services, in the completion of this work. I thank Dugan Witherick for help with Legion Cluster. I thank Joseph Clampitt, Bhuvnesh Jain, Niall Maccran, Sam Cusworth, Emily Hall and Sarah Bridle for helpful discussions on Chapter 5.

Bibliography

Albrecht, A., Amendola, L., Bernstein, G., Clowe, D., Eisenstein, D., Guzzo, L., Hirata, C., Huterer, D., Kirshner, R., Kolb, E. & Nichol, R., 2009, *Findings of the Joint Dark Energy Mission Figure of Merit Science Working Group*
URL: *http://arxiv.org/abs/0901.0721*

Albrecht, A., Bernstein, G., Cahn, R., Freedman, W. L., Hewitt, J., Hu, W., Huth, J., Kamionkowski, M., Kolb, E. W., Knox, L., Mather, J. C., Staggs, S. & Suntzeff, N. B., 2006, *Report of the Dark Energy Task Force*
URL: *http://arxiv.org/abs/astro-ph/0609591*

Amara, A. & Refregier, A., 2007, *Monthly Notices of the Royal Astronomical Society*, **391**(1), 228
URL: *http://dx.doi.org/10.1111/j.1365-2966.2008.13880.x*

Bacon, D., Refregier, A., Clowe, D. & Ellis, R., 2000*a*, *Monthly Notices of the Royal Astronomical Society*, **325**(3), 1065
URL: *http://dx.doi.org/10.1046/j.1365-8711.2001.04507.x*

Bacon, D., Refregier, A. & Ellis, R., 2000*b*, *Monthly Notices of the Royal Astronomical Society*, **318**(2), 625
URL: *http://dx.doi.org/10.1046/j.1365-8711.2000.03851.x*

Bacon, D. J., Goldberg, D. M., Rowe, B. T. P. & Taylor, A. N., 2006, *Monthly Notices of the Royal Astronomical Society*, **365**(2), 414
URL: *http://dx.doi.org/10.1111/j.1365-2966.2005.09624.x*

Bartelmann, M. & Schneider, P., 1999, *Physics Reports*, **340**(4-5), 291
URL: *http://dx.doi.org/10.1016/s0370-1573(00)00082-x*

Bartelmann, M., Viola, M., Melchior, P. & Schäfer, B. M., 2011, *Calibration biases in measurements of weak lensing*
URL: *http://arxiv.org/abs/1103.5923*

Battye, R. A. & Moss, A., 2014, *Evidence for massive neutrinos from CMB and lensing observations*
URL: *http://arxiv.org/abs/1308.5870*

Benjamin, J., Van Waerbeke, L., Heymans, C., Kilbinger, M., Erben, T., Hildebrandt, H., Hoekstra, H., Kitching, T. D., Mellier, Y., Miller, L., Rowe, B., Schrabback, T., Simpson, F., Coupon, J., Fu, L., Harnois-Déraps, J., Hudson, M. J., Kuijken, K., Semboloni, E., Vafaei, S. & Velander, M., 2012, *CFHTLenS tomographic weak lensing: Quantifying accurate redshift distributions*
URL: *http://arxiv.org/abs/1212.3327*

Bernstein, G. M., 2010, *Shape measurement biases from underfitting and ellipticity gradients*
URL: *http://arxiv.org/abs/1001.2333*

Bernstein, G. M. & Armstrong, R., 2013, *Bayesian Lensing Shear Measurement*
URL: *http://arxiv.org/abs/1304.1843*

Bernstein, G. M. & Jarvis, M., 2001, *The Astronomical Journal*, **123**(2), 583
URL: *http://dx.doi.org/10.1086/338085*

Bond, N., Strauss, M. & Cen, R., 2010, *Crawling the Cosmic Network: Exploring the Morphology of Structure in the Galaxy Distribution*
URL: *http://arxiv.org/abs/0903.3601*

Bonnet, H. & Mellier, Y., 1995,
aap, **303**, 331

Bridle, S., Balan, S. T., Bethge, M., Gentile, M., Harmeling, S., Heymans, C., Hirsch, M., Hosseini, R., Jarvis, M., Kirk, D., Kitching, T., Kuijken, K., Lewis, A., Paulin-Henriksson, S., Scholkopf, B., Velander, M., Voigt, L., Witherick, D., Amara, A., Bernstein, G., Courbin, F., Gill, M., Heavens, A., Mandelbaum, R., Massey, R., Moghaddam, B., Rassat, A., Refregier, A., Rhodes, J., Schrabback, T., Shawe-Taylor, J., Shmakova, M., van Waerbeke, L. & Wittman, D., 2009*a*, *Results of the GREAT08 Challenge: An image analysis competition for cosmological lensing*
URL: *http://arxiv.org/abs/0908.0945*

Bridle, S., Shawe-Taylor, J., Amara, A., Applegate, D., Balan, S. T., Berge, J., Bernstein, G., Dahle, H., Erben, T., Gill, M., Heavens, A., Heymans, C., High, F. W., Hoekstra, H., Jarvis, M., Kirk, D., Kitching, T., Kneib, J.-P., Kuijken, K., Lagatutta, D., Mandelbaum, R., Massey, R., Mellier, Y., Moghaddam, B., Moudden, Y., Nakajima, R., Paulin-Henriksson, S., Pires, S., Rassat, A., Refregier, A., Rhodes, J., Schrabback, T., Semboloni, E., Shmakova, M., van Waerbeke, L., Witherick, D., Voigt, L. & Wittman, D., 2009*b*, *The Annals of Applied Statistics*, **3**(1), 6
URL: *http://dx.doi.org/10.1214/08-aoas222*

Bridle, S. L., Kneib, J. P., Bardeau, S. & Gull, S. F., 2002, in P. Natarajan (ed.), *The Shapes of Galaxies and their Dark Halos*, pp. 38–46
URL: *http://dx.doi.org/10.1142/9789812778017_0006*

Bruzual, G. & Charlot, S., 2003, *Monthly Notices of the Royal Astronomical Society*, **344**(4), 1000
URL: *http://dx.doi.org/10.1046/j.1365-8711.2003.06897.x*

Carroll, S. M., 2000, *The Cosmological Constant*
URL: *http://arxiv.org/abs/astro-ph/0004075*

Cautun, M., van de Weygaert, R., Jones, B. J. T. & Frenk, C. S., 2014, *Evolution of the cosmic web*
URL: *http://arxiv.org/abs/1401.7866*

Chang, C., Jarvis, M., Jain, B., Kahn, S. M., Kirkby, D., Connolly, A., Krughoff, S., Peng, E. & Peterson, J. R., 2013, *The Effective Number Density of Galaxies for Weak Lensing Measurements in the LSST Project*
URL: *http://arxiv.org/abs/1305.0793*

Clampitt, J. & Jain, B., 2014, *Lensing Measurements of the Mass Distribution in SDSS Voids*
URL: *http://arxiv.org/abs/1404.1834*

Clampitt, J., Jain, B. & Takada, M., 2014, *Detection of Stacked Filament Lensing Between SDSS Luminous Red Galaxies*
URL: *http://arxiv.org/abs/1402.3302*

Clifton, T., Ferreira, P. G., Padilla, A. & Skordis, C., 2012, *Modified Gravity and Cosmology*
URL: *http://arxiv.org/abs/1106.2476*

Clowe, D., Bradac, M., Gonzalez, A. H., Markevitch, M., Randall, S. W., Jones, C. & Zaritsky, D., 2006, *The Astrophysical Journal*, **648**(2), L109
URL: *http://dx.doi.org/10.1086/508162*

Coc, A., Uzan, J.-P. & Vangioni, E., 2013, *Standard Big-Bang Nucleosynthesis after Planck*
URL: *http://arxiv.org/abs/1307.6955*

Colberg, J. M., Krughoff, K. S. & Connolly, A. J., 2005, *Monthly Notices of the Royal Astronomical Society*, **359**(1), 272
URL: *http://dx.doi.org/10.1111/j.1365-2966.2005.08897.x*

Colless, M., 1998, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **357**(1750), 105
URL: *http://dx.doi.org/10.1098/rsta.1999.0317*

Copeland, E. J., Sami, M. & Tsujikawa, S., 2006, *International Journal of Modern Physics D*, **15**(11), 1753
URL: *http://dx.doi.org/10.1142/s021827180600942x*

Cropper, M., Hoekstra, H., Kitching, T., Massey, R., Amiaux, J., Miller, L., Mellier, Y., Rhodes, J., Rowe, B., Pires, S., Saxton, C. & Scaramella, R., 2013, *Defining a weak lensing experiment in space*
URL: *http://arxiv.org/abs/1210.7691*

Cypriano, E. S., Amara, A., Voigt, L. M., Bridle, S. L., Abdalla, F. B., Refregier, A., Seiffert, M. & Rhodes, J., 2010, *Cosmic shear requirements on the wavelength-dependence of telescope point spread functions*
URL: *http://arxiv.org/abs/1001.0759*

Dietrich, J. P., Werner, N., Clowe, D., Finoguenov, A., Kitching, T., Miller, L. & Simionescu, A., 2012, *Nature*, **487**(7406), 202
URL: *http://dx.doi.org/10.1038/nature11224*

Dodelson, S., Shapiro, C. & White, M., 2005, *Physical Review D*, **73**(2)
URL: *http://dx.doi.org/10.1103/physrevd.73.023009*

Duffy, A. R., Schaye, J., Kay, S. T. & Vecchia, C. D., 2011, *Monthly Notices of the Royal Astronomical Society: Letters*, **390**(1), L64
URL: *http://dx.doi.org/10.1111/j.1745-3933.2008.00537.x*

Duncan, C., Joachimi, B., Heavens, A., Heymans, C. & Hildebrandt, H., 2013, *On the complementarity of galaxy clustering with cosmic shear and flux magnification*
URL: *http://arxiv.org/abs/1306.6870*

Eisenstein, D. J., Zehavi, I., Hogg, D. W., Scoccimarro, R., Blanton, M. R., Nichol, R. C., Scranton, R., Seo, H., Tegmark, M., Zheng, Z., Anderson, S., Annis, J., Bahcall, N., Brinkmann, J., Burles, S., Castander, F. J., Connolly, A., Csabai, I., Doi, M., Fukugita, M., Frieman, J. A., Glazebrook, K., Gunn, J. E., Hendry, J. S., Hennessy, G., Ivezic, Z., Kent, S., Knapp, G. R., Lin, H., Loh, Y., Lupton, R. H., Margon, B., McKay, T., Meiksin, A., Munn, J. A., Pope, A., Richmond, M., Schlegel, D., Schneider, D., Shimasaku, K., Stoughton, C., Strauss, M., SubbaRao, M., Szalay, A. S., Szapudi, I., Tucker, D., Yanny, B. & York, D., 2005, *The Astrophysical Journal*, **633**(2), 560
URL: *http://dx.doi.org/10.1086/466512*

Erben, T., Hildebrandt, H., Miller, L., van Waerbeke, L., Heymans, C., Hoekstra, H., Kitching, T. D., Mellier, Y., Benjamin, J., Blake, C., Bonnett, C., Cordes, O., Coupon, J., Fu, L., Gavazzi,

R., Gillis, B., Grocutt, E., Gwyn, S. D. J., Holhjem, K., Hudson, M. J., Kilbinger, M., Kuijken, K., Milkeraitis, M., Rowe, B. T. P., Schrabback, T., Semboloni, E., Simon, P., Smit, M., Toader, O., Vafaei, S., van Uitert, E. & Velander, M., 2012, *CFHTLenS: The Canada-France-Hawaii Telescope Lensing Survey - Imaging Data and Catalogue Products*
URL: *http://arxiv.org/abs/1210.8156*

Erben, T., van Waerbeke, L., Mellier, Y., Schneider, P., Cuillandre, J. C., Castander, F. J. & Dantel-Fort, M., 1999, *Mass-detection of a matter concentration projected near the cluster Abell 1942: Dark clump or high-redshift cluster?*
URL: *http://arxiv.org/abs/astro-ph/9907134*

Evans, A. K. D. & Bridle, S., 2009
URL: *http://dx.doi.org/10.1088/0004-637x/695/2/1446*

Fu, L., Kilbinger, M., Erben, T., Heymans, C., Hildebrandt, H., Hoekstra, H., Kitching, T. D., Mellier, Y., Miller, L., Semboloni, E., Simon, P., Van Waerbeke, L., Coupon, J., Harnois-Déraps, J., Hudson, M. J., Kuijken, K., Rowe, B., Schrabback, T., Vafaei, S. & Velander, M., 2014, *CFHTLenS: Cosmological constraints from a combination of cosmic shear two-point and three-point correlations*
URL: *http://arxiv.org/abs/1404.5469*

Fu, L., Semboloni, E., Hoekstra, H., Kilbinger, M., van Waerbeke, L., Tereno, I., Mellier, Y., Heymans, C., Coupon, J., Benabed, K., Benjamin, J., Bertin, E., Doré, O., Hudson, M. J., Ilbert, O., Maoli, R., Marmo, C., McCracken, H. J. & Ménard, B., 2007
URL: *http://dx.doi.org/10.1051/0004-6361:20078522*

Gaztanaga, E., 2003, *The Astrophysical Journal*, **589**(1), 82
URL: *http://dx.doi.org/10.1086/374616*

Gentile, M., Courbin, F. & Meylan, G., 2012*a*, *A simple fitting method (gfit) for galaxy shape measurement in weak lensing surveys*
URL: *http://arxiv.org/abs/1211.4847*

Gentile, M., Courbin, F. & Meylan, G., 2012*b*, *Astronomy & Astrophysics*, **549**, A1+
URL: *http://dx.doi.org/10.1051/0004-6361/201219739*

Goobar, A., Hannestad, S., Mortsell, E. & Tu, H., 2006, *Journal of Cosmology and Astroparticle Physics*, **2006**(06), 019
URL: *http://dx.doi.org/10.1088/1475-7516/2006/06/019*

Heymans, C., Grocutt, E., Heavens, A., Kilbinger, M., Kitching, T. D., Simpson, F., Benjamin, J., Erben, T., Hildebrandt, H., Hoekstra, H., Mellier, Y., Miller, L., Van Waerbeke, L., Brown,

M. L., Coupon, J., Fu, L., Harnois-Deraps, J., Hudson, M. J., Kuijken, K., Rowe, B., Schrabback, T., Semboloni, E., Vafaei, S. & Velander, M., 2013, *CFHTLenS tomographic weak lensing cosmological parameter constraints: Mitigating the impact of intrinsic galaxy alignments*
URL: *http://arxiv.org/abs/1303.1808*

Heymans, C., Van Waerbeke, L., Bacon, D., Berge, J., Bernstein, G., Bertin, E., Bridle, S., Brown, M. L., Clowe, D., Dahle, H., Erben, T., Gray, M., Hetterscheidt, M., Hoekstra, H., Hudelot, P., Jarvis, M., Kuijken, K., Margoniner, V., Massey, R., Mellier, Y., Nakajima, R., Refregier, A., Rhodes, J., Schrabback, T. & Wittman, D., 2006, *Monthly Notices of the Royal Astronomical Society*, **368**(3), 1323
URL: *http://dx.doi.org/10.1111/j.1365-2966.2006.10198.x*

Heymans, C., Van Waerbeke, L., Miller, L., Erben, T., Hildebrandt, H., Hoekstra, H., Kitching, T. D., Mellier, Y., Simon, P., Bonnett, C., Coupon, J., Fu, L., Harnois-D'eraps, J., Hudson, M. J., Kilbinger, M., Kuijken, K., Rowe, B., Schrabback, T., Semboloni, E., van Uitert, E., Vafaei, S. & Velander, M., 2012, *CFHTLenS: The Canada-France-Hawaii Telescope Lensing Survey*
URL: *http://arxiv.org/abs/1210.0032*

Higuchi, Y., Oguri, M. & Shirasaki, M., 2014, *Monthly Notices of the Royal Astronomical Society*, **441**(1), 745
URL: *http://dx.doi.org/10.1093/mnras/stu583*

Hildebrandt, H., Erben, T., Kuijken, K., van Waerbeke, L., Heymans, C., Coupon, J., Benjamin, J., Bonnett, C., Fu, L., Hoekstra, H., Kitching, T. D., Mellier, Y., Miller, L., Velander, M., Hudson, M. J., Rowe, B. T. P., Schrabback, T., Semboloni, E. & Benitez, N., 2011, *CFHTLenS: Improving the quality of photometric redshifts with precision photometry*
URL: *http://arxiv.org/abs/1111.4434*

Hirata, C. M., Mandelbaum, R., Seljak, U., Guzik, J., Padmanabhan, N., Blake, C., Brinkmann, J., Budavari, T., Connolly, A., Csabai, I., Scranton, R. & Szalay, A. S., 2004, *Monthly Notices of the Royal Astronomical Society*, **353**(2), 529
URL: *http://dx.doi.org/10.1111/j.1365-2966.2004.08090.x*

Hirata, C. M. & Seljak, U., 2003, *Monthly Notices of the Royal Astronomical Society*, **343**(2), 459
URL: *http://dx.doi.org/10.1046/j.1365-8711.2003.06683.x*

Hoekstra, H. & Jain, B., 2008, *Annual Review of Nuclear and Particle Science*, **58**(1), 99
URL: *http://dx.doi.org/10.1146/annurev.nucl.58.110707.171151*

Hosseini, R. & Bethge, M., 2009, *Spectral Stacking: Unbiased Shear Estimation for Weak Gravitational Lensing*, Technical report

Hu, W., 1999, *The Astrophysical Journal*, **522**(1), L21
URL: *http://dx.doi.org/10.1086/312210*

Hu, W., 2002, *Physical Review D*, **66**(8), 083515+
URL: *http://dx.doi.org/10.1103/physrevd.66.083515*

Hubble, E., 1929, *Proceedings of the National Academy of Science*, **15**, 168
URL: *http://dx.doi.org/10.1073/pnas.15.3.168*

Hudson, M. J., Gillis, B. R., Coupon, J., Hildebrandt, H., Erben, T., Heymans, C., Hoekstra, H., Kitching, T. D., Mellier, Y., Miller, L., Van Waerbeke, L., Bonnett, C., Fu, L., Kuijken, K., Rowe, B., Schrabback, T., Semboloni, E., van Uitert, E. & Velander, M., 2013, *CFHTLenS: Co-evolution of galaxies and their dark matter haloes*
URL: *http://arxiv.org/abs/1310.6784*

Huff, E. M., Eifler, T., Hirata, C. M., Mandelbaum, R., Schlegel, D. & Seljak, U., 2011, *Seeing in the dark -- II. Cosmic shear in the Sloan Digital Sky Survey*
URL: *http://arxiv.org/abs/1112.3143*

Jauzac, M., Jullo, E., Kneib, J.-P., Ebeling, H., Leauthaud, A., Ma, C.-J., Limousin, M., Massey, R. & Richard, J., 2012, *A Weak-Lensing Mass Reconstruction of the Large-Scale Filament Feeding the Massive Galaxy Cluster MACSJ0717.5+3745*
URL: *http://arxiv.org/abs/1208.4323*

Kacprzak, T., Bridle, S., Rowe, B., Voigt, L., Zuntz, J., Hirsch, M. & MacCrann, N., 2013, *Sérsic galaxy models in weak lensing shape measurement: model bias, noise bias and their interaction*
URL: *http://arxiv.org/abs/1308.4663*

Kacprzak, T., Zuntz, J., Rowe, B., Bridle, S., Refregier, A., Amara, A., Voigt, L. & Hirsch, M., 2012, *Measurement and Calibration of Noise Bias in Weak Lensing Galaxy Shape Estimation*
URL: *http://arxiv.org/abs/1203.5049*

Kaiser, N., 1992,
apj, **388**, 272
URL: *http://dx.doi.org/10.1086/171151*

Kaiser, N., 1996, *The Astrophysical Journal*, **498**(1), 26
URL: *http://dx.doi.org/10.1086/305515*

Kaiser, N., 1999, *The Astrophysical Journal*, **537**(2), 555
URL: *http://dx.doi.org/10.1086/309041*

Kaiser, N., Squires, G. & Broadhurst, T., 1994, *The Astrophysical Journal*, **449**, 460+
URL: *http://dx.doi.org/10.1086/176071*

Kaiser, N., Wilson, G. & Luppino, G. A., 2000, *Large-Scale Cosmic Shear Measurements*
URL: *http://arxiv.org/abs/astro-ph/0003338*

Kawasaki, M., Kohri, K. & Moroi, T., 2005, *Physical Review D*, **71**(8)
URL: *http://dx.doi.org/10.1103/physrevd.71.083502*

Kilbinger, M., Fu, L., Heymans, C., Simpson, F., Benjamin, J., Erben, T., Harnois-Deraps, J., Hoekstra, H., Hildebrandt, H., Kitching, T. D., Mellier, Y., Miller, L., Van Waerbeke, L., Benabed, K., Bonnett, C., Coupon, J., Hudson, M. J., Kuijken, K., Rowe, B., Schrabback, T., Semboloni, E., Vafaei, S. & Velander, M., 2013, *CFHTLenS: Combined probe cosmological model comparison using 2D weak gravitational lensing*
URL: *http://arxiv.org/abs/1212.3338*

Kirk, D., Lahav, O., Bridle, S., Jouvel, S., Abdalla, F. B. & Frieman, J. A., 2013, *Optimising Spectroscopic and Photometric Galaxy Surveys: Same-sky Benefits for Dark Energy and Modified Gravity*
URL: *http://arxiv.org/abs/1307.8062*

Kitching, T., Balan, S., Bernstein, G., Bethge, M., Bridle, S., Courbin, F., Gentile, M., Heavens, A., Hirsch, M., Hosseini, R., Kiessling, A., Amara, A., Kirk, D., Kuijken, K., Mandelbaum, R., Moghaddam, B., Nurbaeva, G., Paulin-Henriksson, S., Rassat, A., Rhodes, J., Schölkopf, B., Shawe-Taylor, J., Gill, M., Shmakova, M., Taylor, A., Velander, M., van Waerbeke, L., Witherick, D., Wittman, D., Harmeling, S., Heymans, C., Massey, R., Rowe, B., Schrabback, T. & Voigt, L., 2011, *The Annals of Applied Statistics*, **5**(3), 2231
URL: *http://dx.doi.org/10.1214/11-aoas484*

Kitching, T. D., Balan, S. T., Bridle, S., Cantale, N., Courbin, F., Eifler, T., Gentile, M., Gill, M. S. S., Harmeling, S., Heymans, C., Hirsch, M., Honscheid, K., Kacprzak, T., Kirkby, D., Margala, D., Massey, R. J., Melchior, P., Nurbaeva, G., Patton, K., Rhodes, J., Rowe, B. T. P., Taylor, A. N., Tewes, M., Viola, M., Witherick, D., Voigt, L., Young, J. & Zuntz, J., 2013, *Monthly Notices of the Royal Astronomical Society*, **423**(4), 3163
URL: *http://dx.doi.org/10.1111/j.1365-2966.2012.21095.x*

Kuijken, K., 1999, *Weak weak lensing: correcting weak shear measurements accurately for PSF anisotropy*
URL: *http://arxiv.org/abs/astro-ph/9904418*

Lahav, O. & Liddle, A. R., 2014, *The Cosmological Parameters 2014*
URL: *http://arxiv.org/abs/1401.1389*

Lewis, A., 2009, *Monthly Notices of the Royal Astronomical Society*, **398**(1), 471
URL: *http://dx.doi.org/10.1111/j.1365-2966.2009.15161.x*

Liddle, A. R. & Lyth, D. H., 2000, *Cosmological Inflation and Large-Scale Structure* (Cambridge Univ. Press, Cambridge)

Liu, X., Pan, C., Li, R., Shan, H., Wang, Q., Fu, L., Fan, Z., Kneib, J.-P., Leauthaud, A., Van Waerbeke, L., Makler, M., Moraes, B., Erben, T. & Charbonnier, A., 2014, *Cosmological Constraints From Weak Lensing Peak Statistics With CFHT Stripe-82 Survey*
URL: *http://arxiv.org/abs/1412.3683*

Lourakis, M. I. A., Jul. 2004, *levmar: Levenberg-Marquardt nonlinear least squares algorithms in C/C++*, [web page]
verb+http://www.ics.forth.gr/~lourakis/levmar/+, [Accessed on 31 Jan. 2005.]

Luppino, G. A. & Kaiser, N., 1996, *The Astrophysical Journal*, **475**(1), 20
URL: *http://dx.doi.org/10.1086/303508*

MacCrann, N., Zuntz, J., Bridle, S., Jain, B. & Becker, M. R., 2014, *Cosmic Discordance: Are Planck CMB and CFHTLenS weak lensing measurements out of tune?*
URL: *http://arxiv.org/abs/1408.4742*

Mandelbaum, R., Hirata, C. M., Leauthaud, A., Massey, R. J. & Rhodes, J., 2011, *Precision simulation of ground-based lensing data using observations from space*
URL: *http://arxiv.org/abs/1107.4629*

Mandelbaum, R., Rowe, B., Bosch, J., Chang, C., Courbin, F., Gill, M., Jarvis, M., Kannawadi, A., Kacprzak, T., Lackner, C., Leauthaud, A., Miyatake, H., Nakajima, R., Rhodes, J., Simet, M., Zuntz, J., Armstrong, B., Bridle, S., Coupon, J., Dietrich, J. P., Gentile, M., Heymans, C., Jurling, A. S., Kent, S. M., Kirkby, D., Margala, D., Massey, R., Melchior, P., Peterson, J., Roodman, A. & Schrabback, T., 2014, *The Third Gravitational Lensing Accuracy Testing (GREAT3) Challenge Handbook*
URL: *http://arxiv.org/abs/1308.4982*

Markevitch, M., 2005, *Chandra observation of the most interesting cluster in the Universe*
URL: *http://arxiv.org/abs/astro-ph/0511345*

Martin, J., 2012, *Everything You Always Wanted To Know About The Cosmological Constant Problem (But Were Afraid To Ask)*
URL: *http://arxiv.org/abs/1205.3365*

Massey, R., Heymans, C., Berge, J., Bernstein, G., Bridle, S., Clowe, D., Dahle, H., Ellis, R., Erben, T., Hetterscheidt, M., High, F. W., Hirata, C., Hoekstra, H., Hudelot, P., Jarvis, M.,

Johnston, D., Kuijken, K., Margoniner, V., Mandelbaum, R., Mellier, Y., Nakajima, R., Paulin-Henriksson, S., Peeples, M., Roat, C., Refregier, A., Rhodes, J., Schrabback, T., Schirmer, M., Seljak, U., Semboloni, E. & Van Waerbeke, L., 2006, *Monthly Notices of the Royal Astronomical Society*, **376**(1), 13
URL: *http://dx.doi.org/10.1111/j.1365-2966.2006.11315.x*

Massey, R., Hoekstra, H., Kitching, T., Rhodes, J., Cropper, M., Amiaux, J., Harvey, D., Mellier, Y., Meneghetti, M., Miller, L., Paulin-Henriksson, S., Pires, S., Scaramella, R. & Schrabback, T., 2012, *Origins of weak lensing systematics, and requirements on future instrumentation (or knowledge of instrumentation)*
URL: *http://arxiv.org/abs/1210.7690*

Massey, R. & Refregier, A., 2006, *Monthly Notices of the Royal Astronomical Society*, **363**(1), 197
URL: *http://dx.doi.org/10.1111/j.1365-2966.2005.09453.x*

Mather, J. C., Cheng, E. S., Eplee, R. E., Isaacman, R. B., Meyer, S. S., Shafer, R. A., Weiss, R., Wright, E. L., Bennett, C. L., Boggess, N. W., Dwek, E., Gulkis, S., Hauser, M. G., Janssen, M., Kelsall, T., Lubin, P. M., Moseley, S. H., Murdock, T. L., Silverberg, R. F., Smoot, G. F. & Wilkinson, D. T., 1990,
apjl, **354**, L37
URL: *http://dx.doi.org/10.1086/185717*

Maturi, M. & Merten, J., 2013, *Weak lensing detection of intra-cluster filaments with ground based data*
URL: *http://arxiv.org/abs/1306.0015*

Mead, J. M. G., King, L. J. & McCarthy, I. G., 2009, *Probing the cosmic web: inter-cluster filament detection using gravitational lensing*
URL: *http://arxiv.org/abs/0910.2417*

Melchior, P., Boehnert, A., Lombardi, M. & Bartelmann, M., 2009, *Limitations for shapelet-based weak-lensing measurements*
URL: *http://arxiv.org/abs/0906.5092*

Melchior, P., Suchyta, E., Huff, E., Hirsch, M., Kacprzak, T., Rykoff, E., Gruen, D., Armstrong, R., Bacon, D., Bechtol, K., Bernstein, G. M., Bridle, S., Clampitt, J., Honscheid, K., Jain, B., Jouvel, S., Krause, E., Lin, H., MacCrann, N., Patton, K., Plazas, A., Rowe, B., Vikram, V., Wilcox, H., Young, J., Zuntz, J., Abbott, T., Abdalla, F., Allam, S. S., Banerji, M., Bernstein, J. P., Bernstein, R. A., Bertin, E., Buckley-Geer, E., Burke, D. L., Castander, F. J., da Costa, L. N., Cunha, C. E., Depoy, D. L., Desai, S., Diehl, H. T., Doel, P., Estrada, J., Evrard, A. E., Neto, A. F., Fernandez, E., Finley, D. A., Flaugher, B., Frieman, J. A., Gaztanaga, E.,

Gerdes, D., Gruendl, R. A., Gutierrez, G. R., Jarvis, M., Karliner, I., Kent, S., Kuehn, K., Kuropatkin, N., Lahav, O., Maia, M. A. G., Makler, M., Marriner, J., Marshall, J. L., Merritt, K. W., Miller, C. J., Miquel, R., Mohr, J., Neilsen, E., Nichol, R. C., Nord, B. D., Reil, K., Roe, N. A., Roodman, A., Sako, M., Sanchez, E., Santiago, B. X., Schindler, R., Schubnell, M., Sevilla-Noarbe, I., Sheldon, E., Smith, C., Soares-Santos, M., Swanson, M. E. C., Sypniewski, A. J., Tarle, G., Thaler, J., Thomas, D., Tucker, D. L., Walker, A., Wechsler, R., Weller, J. & Wester, W., 2014, *Mass and galaxy distributions of four massive galaxy clusters from Dark Energy Survey Science Verification data*
URL: *http://arxiv.org/abs/1405.4285*

Melchior, P. & Viola, M., 2012, *Monthly Notices of the Royal Astronomical Society*, **424**(4), 2757
URL: *http://dx.doi.org/10.1111/j.1365-2966.2012.21381.x*

Melchior, P., Viola, M., Schäfer, B. M. & Bartelmann, M., 2012, *Monthly Notices of the Royal Astronomical Society*, **412**(3), 1552
URL: *http://dx.doi.org/10.1111/j.1365-2966.2010.17875.x*

Miller, L., Heymans, C., Kitching, T. D., Van Waerbeke, L., Erben, T., Hildebrandt, H., Hoekstra, H., Mellier, Y., Rowe, B. T. P., Coupon, J., Dietrich, J. P., Fu, L., Harnois-Deraps, J., Hudson, M. J., Kilbinger, M., Kuijken, K., Schrabback, T., Semboloni, E., Vafaei, S. & Velander, M., 2012, *Monthly Notices of the Royal Astronomical Society*, **429**(4), 2858
URL: *http://dx.doi.org/10.1093/mnras/sts454*

Miller, L., Kitching, T. D., Heymans, C., Heavens, A. F. & Van Waerbeke, L., 2007, *Monthly Notices of the Royal Astronomical Society*, **382**(1), 315
URL: *http://dx.doi.org/10.1111/j.1365-2966.2007.12363.x*

Moffat, A. F. J., 1969,
aap, **3**, 455

Nakajima, R. & Bernstein, G., 2006, *The Astronomical Journal*, **133**(4), 1763
URL: *http://dx.doi.org/10.1086/511957*

Navarro, J. F., Frenk, C. S. & White, S. D. M., 1997, *The Astrophysical Journal*, **490**(2), 493
URL: *http://dx.doi.org/10.1086/304888*

Okura, Y. & Futamase, T., 2010, *Elliptical Weighted HOLICs for Weak Lensing Shear Measurement. part1:Definitions and isotropic PSF correction*
URL: *http://arxiv.org/abs/1010.0466*

Okura, Y. & Futamase, T., 2012, *Elliptical Weighted HOLICs for Weak Lensing Shear Measurement part3:Random Count Noise Effect for Image's Moments in Weak Lensing Analysis*
URL: *http://arxiv.org/abs/1208.3564*

Olive, K. A., Steigman, G. & Walker, T. P., 1999, *Physics Reports*, **333-334**, 389
URL: *http://dx.doi.org/10.1016/s0370-1573(00)00031-4*

Paulin-Henriksson, S., Amara, A., Voigt, L., Refregier, A. & Bridle, S. L., 2008, *Astronomy and Astrophysics*, **484**(1), 67
URL: *http://dx.doi.org/10.1051/0004-6361:20079150*

Peacock, J. A., Schneider, P., Efstathiou, G., Ellis, J. R., Leibundgut, B., Lilly, S. J. & Mellier, Y., 2006, *Report by the ESA-ESO Working Group on Fundamental Cosmology*
URL: *http://arxiv.org/abs/astro-ph/0610906*

Peng, C. Y., Ho, L. C., Impey, C. D. & Rix, H.-W., 2002, *The Astronomical Journal*, **124**(1), 266
URL: *http://dx.doi.org/10.1086/340952*

Penzias, A. A. & Wilson, R. W., 1965,
apj, **142**, 419
URL: *http://dx.doi.org/10.1086/148307*

Perlmutter, S., Aldering, G., Goldhaber, G., Knop, R. A., Nugent, P., Castro, P. G., Deustua, S., Fabbro, S., Goobar, A., Groom, D. E., Hook, I. M., Kim, A. G., Kim, M. Y., Lee, J. C., Nunes, N. J., Pain, R., Pennypacker, C. R., Quimby, R., Lidman, C., Ellis, R. S., Irwin, M., McMahon, R. G., Ruiz-Lapuente, P., Walton, N., Schaefer, B., Boyle, B. J., Filippenko, A. V., Matheson, T., Fruchter, A. S., Panagia, N., Newberg, H. J. M. & Couch, W. J., 1998, *The Astrophysical Journal*, **517**(2), 565
URL: *http://dx.doi.org/10.1086/307221*

Planck Collaboration, 2012, *Planck intermediate results. VIII. Filaments between interacting clusters*
URL: *http://arxiv.org/abs/1208.5911*

Planck Collaboration, 2014*a*, *Astronomy & Astrophysics*, **571**, A1+
URL: *http://dx.doi.org/10.1051/0004-6361/201321529*

Planck Collaboration, 2014*b*, *Planck 2013 results. XVI. Cosmological parameters*
URL: *http://arxiv.org/abs/1303.5076*

Refregier, A., 2001, *Monthly Notices of the Royal Astronomical Society*, **338**(1), 35
URL: *http://dx.doi.org/10.1046/j.1365-8711.2003.05901.x*

Refregier, A., 2003, *Annual Review of Astronomy and Astrophysics*, **41**(1), 645
URL: *http://dx.doi.org/10.1146/annurev.astro.41.111302.102207*

Refregier, A. & Amara, A., 2014, *A Way Forward for Cosmic Shear: Monte-Carlo Control Loops*
URL: *http://arxiv.org/abs/1303.4739*

Refregier, A., Kacprzak, T., Amara, A., Bridle, S. & Rowe, B., 2012, *Noise bias in weak lensing shape measurements*
URL: *http://arxiv.org/abs/1203.5050*

Riess, A. G., Filippenko, A. V., Challis, P., Clocchiattia, A., Diercks, A., Garnavich, P. M., Gilliland, R. L., Hogan, C. J., Jha, S., Kirshner, R. P., Leibundgut, B., Phillips, M. M., Reiss, D., Schmidt, B. P., Schommer, R. A., Smith, R. C., Spyromilio, J., Stubbs, C., Suntzeff, N. B. & Tonry, J., 1998, *The Astronomical Journal*, **116**(3), 1009
URL: *http://dx.doi.org/10.1086/300499*

Rowe, B., Bacon, D., Massey, R., Heymans, C., Häußler, B., Taylor, A., Rhodes, J. & Mellier, Y., 2013, *Monthly Notices of the Royal Astronomical Society*, **435**(1), 822
URL: *http://dx.doi.org/10.1093/mnras/stt1353*

Rowe, B., Jarvis, M., Mandelbaum, R., Bernstein, G. M., Bosch, J., Simet, M., Meyers, J. E., Kacprzak, T., Nakajima, R., Zuntz, J., Miyatake, H., Dietrich, J. P., Armstrong, R., Melchior, P. & Gill, M. S. S., 2014, *GalSim: The modular galaxy image simulation toolkit*
URL: *http://arxiv.org/abs/1407.7676*

Rubin, V. C., Ford, W. K. J. & Thonnard, 1980,
apj, **238**, 471
URL: *http://dx.doi.org/10.1086/158003*

Schmidt, F., Leauthaud, A., Massey, R., Rhodes, J., George, M. R., Koekemoer, A. M., Finoguenov, A. & Tanaka, M., 2011, *The Astrophysical Journal*, **744**(2), L22+
URL: *http://dx.doi.org/10.1088/2041-8205/744/2/l22*

Schneider, P., 1995, *Cosmological Applications of Gravitational Lensing*
URL: *http://arxiv.org/abs/astro-ph/9512047*

Schneider, P. & Seitz, C., 1994, *Steps towards Nonlinear Cluster Inversion Through Gravitational Distortions. I. Basic Considerations and Circular Clusters*
URL: *http://arxiv.org/abs/astro-ph/9407032*

SDSS Collaboration, 2014, *The Tenth Data Release of the Sloan Digital Sky Survey: First Spectroscopic Data from the SDSS-III Apache Point Observatory Galactic Evolution Experiment*
URL: *http://arxiv.org/abs/1307.7735*

Seitz, C. & Schneider, P., 1996, *Steps towards nonlinear cluster inversion through gravitational distortions: III. Including a redshift distribution of the sources*
URL: *http://arxiv.org/abs/astro-ph/9601079*

Semboloni, E., Hoekstra, H., Huang, Z., Cardone, V., Cropper, M., Joachimi, B., Kitching, T., Kuijken, K., Lombardi, M., Maoli, R., Mellier, Y., Miller, L., Rhodes, J., Scaramella, R., Schrabback, T. & Velander, M., 2012, *On the shear estimation bias induced by the spatial variation of colour across galaxy profiles*
URL: *http://arxiv.org/abs/1211.5025*

Sérsic, J. L., 1963, *Boletin de la Asociacion Argentina de Astronomia La Plata Argentina*, **6**, 41

Sheldon, E. S., 2014, *An Implementation of Bayesian Lensing Shear Measurement*
URL: *http://arxiv.org/abs/1403.7669*

Simpson, F., Heymans, C., Parkinson, D., Blake, C., Kilbinger, M., Benjamin, J., Erben, T., Hildebrandt, H., Hoekstra, H., Kitching, T. D., Mellier, Y., Miller, L., Van Waerbeke, L., Coupon, J., Fu, L., Harnois-Déraps, J., Hudson, M. J., Kuijken, K., Rowe, B., Schrabback, T., Semboloni, E., Vafaei, S. & Velander, M., 2012, *Monthly Notices of the Royal Astronomical Society*, **429**(3), 2249
URL: *http://dx.doi.org/10.1093/mnras/sts493*

Sousbie, T., 2013, *DisPerSE: robust structure identification in 2D and 3D*
URL: *http://arxiv.org/abs/1302.6221*

Spergel, D. N., Verde, L., Peiris, H. V., Komatsu, E., Nolta, M. R., Bennett, C. L., Halpern, M., Hinshaw, G., Jarosik, N., Kogut, A., Limon, M., Meyer, S. S., Page, L., Tucker, G. S., Weiland, J. L., Wollack, E. & Wright, E. L., 2003, *The Astrophysical Journal Supplement Series*, **148**(1), 175
URL: *http://dx.doi.org/10.1086/377226*

Springel, V., White, S. D. M., Jenkins, A., Frenk, C. S., Yoshida, N., Gao, L., Navarro, J., Thacker, R., Croton, D., Helly, J., Peacock, J. A., Cole, S., Thomas, P., Couchman, H., Evrard, A., Colberg, J. & Pearce, F., 2005, *Nature*, **435**(7042), 629
URL: *http://dx.doi.org/10.1038/nature03597*

Takada, M. & Jain, B., 2005, *Monthly Notices of the Royal Astronomical Society*, **348**(3), 897
URL: *http://dx.doi.org/10.1111/j.1365-2966.2004.07410.x*

Tegmark, M., Eisenstein, D., Strauss, M., Weinberg, D., Blanton, M., Frieman, J., Fukugita, M., Gunn, J., Hamilton, A., Knapp, G., Nichol, R., Ostriker, J., Padmanabhan, N., Percival, W., Schlegel, D., Schneider, D., Scoccimarro, R., Seljak, U., Seo, H., Swanson, M., Szalay, A.,

Vogeley, M., Yoo, J., Zehavi, I., Abazajian, K., Anderson, S., Annis, J., Bahcall, N., Bassett, B., Berlind, A., Brinkmann, J., Budavari, T., Castander, F., Connolly, A., Csabai, I., Doi, M., Finkbeiner, D., Gillespie, B., Glazebrook, K., Hennessy, G., Hogg, D., Ivezic, Z., Jain, B., Johnston, D., Kent, S., Lamb, D., Lee, B., Lin, H., Loveday, J., Lupton, R., Munn, J., Pan, K., Park, C., Peoples, J., Pier, J., Pope, A., Richmond, M., Rockosi, C., Scranton, R., Sheth, R., Stebbins, A., Stoughton, C., Szapudi, I., Tucker, D., Berk, D. V., Yanny, B. & York, D., 2006, *Physical Review D*, **74**(12), 123507+
URL: *http://dx.doi.org/10.1103/physrevd.74.123507*

Tegmark, M., Strauss, M., Blanton, M., Abazajian, K., Dodelson, S., Sandvik, H., Wang, X., Weinberg, D., Zehavi, I., Bahcall, N., Hoyle, F., Schlegel, D., Scoccimarro, R., Vogeley, M., Berlind, A., Budavari, T., Connolly, A., Eisenstein, D., Finkbeiner, D., Frieman, J., Gunn, J., Hui, L., Jain, B., Johnston, D., Kent, S., Lin, H., Nakajima, R., Nichol, R., Ostriker, J., Pope, A., Scranton, R., Seljak, U., Sheth, R., Stebbins, A., Szalay, A., Szapudi, I., Xu, Y. & Others, ., 2004, *Physical Review D*, **69**(10), 103501+
URL: *http://dx.doi.org/10.1103/physrevd.69.103501*

Tempel, E., Stoica, R. S., Martinez, V. J., Liivamägi, L. J., Castellan, G. & Saar, E., 2013, *Detecting filamentary pattern in the cosmic web: a catalogue of filaments for the SDSS*
URL: *http://arxiv.org/abs/1308.2533*

Thomas, S. A., Abdalla, F. B. & Lahav, O., 2010, *Physical Review Letters*, **105**(3), 031301+
URL: *http://dx.doi.org/10.1103/physrevlett.105.031301*

Van Waerbeke, L., 2009, *Shear and Magnification: Cosmic Complementarity*
URL: *http://arxiv.org/abs/0906.1583*

Van Waerbeke, L., Mellier, Y., Erben, T., Cuillandre, J. C., Bernardeau, F., Maoli, R., Bertin, E., Cracken, H. J. M., Le Fevre, O., Fort, B., Dantel-Fort, M., Jain, B. & Schneider, P., 2000, *Detection of correlated galaxy ellipticities on CFHT data: first evidence for gravitational lensing by large-scale structures*
URL: *http://arxiv.org/abs/astro-ph/0002500*

Velander, M., Kuijken, K. & Schrabback, T., 2011, *Monthly Notices of the Royal Astronomical Society*, **412**(4), 2665
URL: *http://dx.doi.org/10.1111/j.1365-2966.2010.18085.x*

Viola, M., Melchior, P. & Bartelmann, M., 2010, *Biases in, and corrections to, KSB shear measurements*
URL: *http://arxiv.org/abs/1006.2470*

Voigt, L. M. & Bridle, S. L., 2009, *Limitations of model fitting methods for lensing shear estimation*
URL: *http://arxiv.org/abs/0905.4801*

Voigt, L. M., Bridle, S. L., Amara, A., Cropper, M., Kitching, T. D., Massey, R., Rhodes, J. & Schrabback, T., 2011, *The impact of galaxy colour gradients on cosmic shear measurement*
URL: *http://arxiv.org/abs/1105.5595*

Wittman, D. M., Tyson, J. A., Kirkman, D., Dell'Antonio, I. & Bernstein, G., 2000, *Detection of weak gravitational lensing distortions of distant galaxies by cosmic dark matter at large scales*
URL: *http://arxiv.org/abs/astro-ph/0003014*

Wright, C. O. & Brainerd, T. G., 1999, *Gravitational Lensing by NFW Halos*
URL: *http://arxiv.org/abs/astro-ph/9908213*

Zel'dovich, Y. B., 1970,
aap, **5**, 84

Zhang, Y., Dietrich, J. P., McKay, T. A., Sheldon, E. S. & Nguyen, A. T. Q., 2013, *Studying Inter-Cluster Galaxy Filaments Through Stacking GMBCG Galaxy Cluster Pairs*
URL: *http://arxiv.org/abs/1304.6696*

Zuntz, J., Kacprzak, T., Voigt, L., Hirsch, M., Rowe, B. & Bridle, S., 2013, *IM3SHAPE: A maximum-likelihood galaxy shear measurement code for cosmic gravitational lensing*
URL: *http://arxiv.org/abs/1302.0183*