THE UNIVERSITY *of* EDINBURGH

Optimal Weak Lensing Tomography for CFHTLenS

Emma Liana Grocutt

A thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy
to the
University of Edinburgh
2012

Abstract

Weak gravitational lensing is a powerful astronomical tool for constraining cosmological parameters that is entering its prime. Lensing occurs because gravitational fields deflect light rays and measuring this deflection through a statistic known as cosmic shear allows us to directly measure the properties of dark matter and dark energy on large scales. In principle, gravitational lensing is a clean probe of the cosmology of the Universe, as it depends on gravity alone and not on incomplete astrophysical models or approximations. In practice, however, there are several factors that limit the accuracy and precision of lensing measurements. These include accurate measurement of galaxy shapes, correctly accounting for distortions to galaxy images due to the point spread function of the telescope, the presence of intrinsic alignments (IAs) of galaxy shapes due to physical processes, and inaccuracies in commonly-used galaxy photometric redshift information. These effects may all introduce systematic errors in lensing measurements which must be carefully accounted for to ensure that cosmological constraints from lensing are unbiased and as precise as possible.

The Canada-France-Hawaii-Telescope Lensing Survey (CFHTLenS) is the largest weak lensing survey completed to date, covering 154 square degrees of the sky in 5 optical bands, with photometric redshift information for every survey galaxy. With lensing measurements from more galaxies than ever before, the statistical uncertainties on parameter estimates will be the lowest ever achieved from weak lensing. If left unaccounted for, sources of systematic error would dominate over the statistical uncertainty, potentially biasing parameter estimates catastrophically. A technique known as tomography in which galaxies are sorted into bins based on their redshift can help constrain cosmological parameters more precisely. This is because utilising the redshifts of survey galaxies retains cosmological information that would otherwise be lost, such as the behaviour

of dark energy and the growth of structure over time. Tomography, however, increases the demand for systematics-free galaxy catalogues as the technique is strongly sensitive to the IA signal and photometric redshift errors. Therefore, future lensing analyses will require a more sophisticated treatment of these effects to extract maximal information from the lensing signal. A thorough understanding of the error on lensing measurements is necessary in order to produce meaningful cosmological constraints. One of the key features of cosmic shear is that it is highly correlated over different angular scales, meaning that error estimates must take into account the covariance of the data over different angular scales, and in the case of tomography, between different redshift bins. The behaviour and size of the (inverse) covariance matrix is one of the limiting factors in such a cosmological likelihood analysis, so constructing an accurate, unbiased estimate of the covariance matrix inverse is essential to cosmic shear analysis.

This thesis presents work to optimise tomographic weak lensing analysis and achieve the tightest parameter constraints possible for a CFHTLenS-like survey. N-body simulations and Gaussian shear fields incorporating an IA model (known as the 'non-linear alignment' model) with a free parameter are used to estimate fully tomographic covariance matrices of cosmic shear for CFHTLenS. We simultaneously incorporate for the first time the error contribution expected from the non-linear alignment model for IAs and realistic photometric redshift uncertainties as measured from the CFHTLenS. We find that non-Gaussian simulations that incorporate nonlinearity on small scales are needed to ensure the covariance is not underestimated, and that the covariance matrix is shot-noise dominated for almost all tomographic correlations. The number of realisations of the simulations used to estimate the covariance places a hard limit on the maximum number of tomographic bins that one can use in an analysis. Given the available number of lines of sight generated from CFHTLenS-like simulations, we find that up to ~ 15 tomographic bins may be utilised in a likelihood analysis.

The estimated tomographic covariance matrices are used in a least-squares likelihood analysis in order to find the combination of both angular and tomographic bins that gives the tightest constraints on some key cosmological parameters. We find that the optimum binning is somewhat degenerate, with around 6 tomographic and 8 angular bins being optimal, and limited by the

available number of realisations of the simulations used to estimate the covariance. We also investigate the bias on best-fit parameter estimates that occurs if IAs or photometric redshift errors are neglected. With our choice of IA model, the effect of neglecting IAs on the best-fit cosmological parameters is not significant for a CFHTLenS-like survey, although this may not be true if the IA signal differs substantially from the model, or for future wide-field surveys with much smaller statistical uncertainties. Similarly, neglecting photometric redshift errors does not result in significant bias, although we apply similar caveats.

Finally, we apply the results of this optimisation to the CFHTLenS cosmic shear data, performing a preliminary analysis of the shear correlation function to produce both 2D and optimal tomographic cosmological constraints. From 6-bin tomography, we constrain the matter density parameter $\Omega_m = 0.419^{+0.123}_{-0.090}$, the amplitude of the matter power spectrum $\sigma_8 = 0.623^{+0.101}_{-0.084}$ and the amplitude parameter of the non-linear alignment model, $A = -1.161^{+1.163}_{-0.597}$. We perform this analysis to test the validity and limitations of the optimal binning on real data and find that 6-bin tomography improves parameter constraints considerably, albeit not as much as when performed on simulated data. This analysis represents an important step in the development of techniques to optimise the recovery of lensing information and hence cosmological constraints, while simultaneously accounting for potential sources of bias in shear analysis.

Declaration

Except where otherwise stated, the research undertaken in this thesis was the unaided work of the author. Where the work was done in collaboration with others, a significant contribution was made by the author.

E. Grocutt
30th July 2012

Acknowledgements

I acknowledge the support of the Science and Technologies Facilities Council and the European Research Council in funding this thesis. This thesis would not have been possible without the guidance and support of my supervisors Catherine Heymans, Alan Heavens and Tom Kitching. I would like to thank Martin Kilbinger and Michael Brown for providing me with invaluable code for this project, and in particular to Martin for many useful discussions. I thank the CFHTLenS collaboration for providing me with data. Others who have supported me academically include the Weak Lensing group at the IfA, Edinburgh, and in particular Andy Taylor who taught me a great deal about lens theory.

I thank those who have supported me on personal level throughout my PhD: my Mum, Grandma, brother and especially my Dad for talking me off a ledge many times. I also thank my friends, in particular the closest friend I have had during my PhD, Rebecca Bowler. I also thank Joe Harkness for his kindness and patience through a difficult period, and Dave Taylor for helping me make the transition into my PhD. I would also like to thank Paula Wilkie, Ken Rice and Jane Patterson for having my back.

Contents

List of Figures

List of Tables

Chapter 1

Introduction to Cosmology

This thesis concerns the observation of weak gravitational lensing within the context of cosmology. In this chapter I present an outline of the relevant cosmological background to this work. The standard ΛCDM model of cosmology will be described in terms of its key components, and a review of the evidence for the model will be presented. The Hot Big Bang model is assumed, as to date it is the widely-accepted paradigm describing the early Universe. This chapter is intended to introduce the relevant background and concepts that are essential in understanding the goals of this thesis. We extend this to an introduction to weak lensing in Chapter 2.

1.1 The Standard Model

Cosmology as it stands today is characterised by the two main components of our Universe: **dark matter** and **dark energy**, which make up $\sim 96\%$ of the Universe's mass-energy content and are collectively known as the dark sector. The discovery of these two components has led to the development of a model of our Universe known as the standard, concordance or ΛCDM model. This model is well supported by a large body of observational evidence but is, however, phenomenological and one of the most pertinent challenges in modern cosmology is that of trying to develop a testable theoretical model to explain the dark sector. There are also many other puzzles facing cosmologists today, such as the nature of inflation, the resolution of the coincidence problem and the reconciliation of the energy of the cosmological constant (the 'Λ' in ΛCDM) and its predicted energy

from particle physics. In order to develop theories to solve these puzzles, high-quality observations from a number of different cosmological probes are necessary to allow us to test hypotheses and measure the underlying physics. The ability to constrain cosmology using weak lensing is now a decade old, and with several deep weak lensing surveys upcoming in the next decade which promise to be wider than ever before, weak lensing measurement is becoming established as an invaluable technique for understanding the nature of the cosmos.

1.1.1 The Inflationary Universe

One of the most important discoveries in cosmology was made by Edwin Hubble in the 1920s when he deduced that the velocity of recession of distant galaxies was directly proportional to their distance from us [1]. Through his observations, Hubble not only proved that the Universe was bigger than previously thought due to the determination of distances to other galaxies, but the velocity–distance law showed that the entire Universe was expanding [2, 3]. Hubble's work and the discovery in 1964 by Penzias and Wilson of an almost uniform radio signal over the whole sky — the **cosmic microwave background** (CMB) — led to the acceptance of the Big Bang model of cosmology, in which the Universe expanded from hot, dense initial conditions at some finite time in the past and is still expanding today.

The CMB signal was discovered in 1964 by Penzias & Wilson [4] and is a relic from the early Universe, which was initially filled with a plasma of charged particles and radiation in thermal equilibrium [5]. As the Universe expanded, it also cooled and once the temperature was low enough neutral atoms could form from the charged particles in a epoch known as **recombination**. These atoms were no longer able to absorb thermal radiation effectively, which allowed the radiation to travel through the Universe unimpeded — the Universe became transparent. The photons that were allowed to free-stream at the epoch of recombination have been propagating through the Universe ever since, becoming increasingly redshifted due to expansion until their blackbody temperature peaks today in the microwave at 2.73 K. The CMB signal is remarkably uniform down to scales of approximately 1 part in 10^5 (ignoring the well-understood dipole anisotropy in the CMB due to the peculiar motion of the solar system) which leads to the **horizon problem** in the standard Big Bang model.

In order to understand why the horizon problem exists, we much first look at the **cosmological principle** that the Big Bang model holds to be true: *On the largest scales, the Universe is both isotropic and homogeneous.* Isotropy means that the Universe looks the same in all directions for a given observer (Hubble's law for the recession velocity of galaxies supports this; the law holds whichever line of sight is chosen) and homogeneity is satisfied if the statistical properties of the Universe are the same everywhere. That is, there are no preferred observers and a translation in space would not lead to a radically different view of the Universe. Homogeneity cannot be proven correct as we cannot observe the Universe from a cosmologically-significant distance away from the Milky Way galaxy[1]. Although there are some cosmological models that do not assume homogeneity (such as models that use the Lemaître, Tolman & Bondi metric to place observers in the centre of voids), it is a necessary assumption to make if one wants to avoid giving humanity a privileged position in the Universe[2].

There is therefore a problem with the standard Big Bang model if one assumes the cosmological principle holds: how can the CMB, and indeed the Universe in general, look uniform over the entire sky if the Universe has a finite age? If the CMB photons were emitted at the epoch of recombination, then any two areas of sky separated by more than a small angle ($\sim 2\,^\circ$) will not have had time to have causal contact[3] with each other and should not be in equilibrium. The solution to this horizon problem (and several others) is solved by the addition of **inflation** to the Big Bang model. The hypothesised inflationary epoch was a period of exponential expansion that occurred between $t \sim 10^{-35}$ to 10^{-33} seconds after the initial singularity, leading to huge inflation of the then-causally connected regions, meaning regions on all parts of the sky today can be in causal contact and hence in equilibrium. Inflation also solves another cosmological problem known as the **flatness problem**, which is described below in section 1.1.2. It should be noted that despite its ability to neatly solve some key cosmological problems, inflation as a theory has some significant issues. First, the inflationary model is phenomenological in nature. The inflationary field does not correspond

[1] We can, however, test for violations of homogeneity for example through the Sunyaev-Zel'dovich effect of clusters on CMB photons, see [6] for a review.

[2] This is the Copernican principle. If the Universe is isotropic and the Copernican principle holds, then it is necessarily homogeneous.

[3] Causally connected regions are limited by the horizon scale, which is defined by the maximum distance information can travel in time t, the age of the Universe.

to any known physical field and theoretical models of inflation must be fitted *ad hoc* to the data and fine-tuned to reproduce observations. Second, in order for inflation to occur, the Universe is required to undergo a phase change to bring an end to the inflationary period. Mechanisms for such a phase change have been theorised, but as yet suffer from a lack of observational evidence to support them. Additionally, inflationary theory exploits changes in the level of the vacuum energy to work (see §1.3), but since the origin and behaviour of the vacuum energy is not fully understood, any theory that depends on it is on somewhat shaky foundations [7].

Inflation, despite its problems as a model, currently represents cosmologists' 'best guess' at the nature and conditions present in the very early Universe. Thus, a cosmological model has emerged which ties together the initial state of the Universe, an epoch of inflation, an era of recombination and ongoing expansion, while also preserving the cosmological principle.

1.1.2 Cosmological Formalism

General Relativity

We describe in this section the mathematical formalism underpinning the standard cosmological model. To define a spacetime metric that adequately describes the cosmological model, we use the expression for a line element in general relativity (GR),

$$\mathrm{d}s^2 = c^2\mathrm{d}\tau^2 = g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu \tag{1.1}$$

where $g_{\mu\nu}$ is the metric tensor (see e.g. [7]). $\mathrm{d}s^2$ is the separation between two points in spacetime. The metric is used to describe the curved space-time of GR. In special relativity this must reduce to

$$c^2\mathrm{d}\tau^2 = c^2\mathrm{d}t^2 - \mathrm{d}x^2 - \mathrm{d}y^2 - \mathrm{d}z^2 \tag{1.2}$$

in Minkowski coordinates, where τ is the proper time and t is the coordinate time (see [8] for an explanation of proper and coordinate time in the context of relativity).

A geodesic describes the shortest path between two spacetime events, which

is the path which a freely-falling particle would take. The integrated path length of a geodesic is, from Eqn. 1.1,

$$s = \int (g_{\mu\nu} \mathrm{d}x^\mu \mathrm{d}x^\nu)^{1/2}. \tag{1.3}$$

Geodesics are stationary paths, meaning that a small deviation from the path does not cause any change in length (to first order) in s. The geodesic equation can be rewritten using the Lagrangian, $\mathcal{L}$, a function of the coordinates x^μ and their derivatives the velocities U^μ. Thus if $\delta s = 0$ for stationary paths, in its most general form the geodesic equation is [9]

$$\delta \int (g_{\mu\nu} U^\mu U^\nu)^{1/2} \mathrm{d}\tau = 0. \tag{1.4}$$

Different types of particles will follow different geodesics. Massless particles such as photons ($v = c$) will follow null geodesics with $\int \mathrm{d}s^2 = 0$. Massive particles will necessarily have $v < c$, and will follow time-like geodesics with $\int \mathrm{d}s^2 > 0$. Hypothetical particles with $v > c$ would follow space-like geodesics, with $\int \mathrm{d}s^2 < 0$. The physical interpretation of this is that particles moving along null geodesics (at speed c) will experience no time passing, whereas massive particles will experience a time interval. Space-like particles, if they were physical, would effectively move backwards through time [8].

The **energy-momentum tensor**, or stress-energy tensor describes the energy and momentum of spacetime in terms of its density ρ and flux,

$$T^{\mu\nu} = \left(\rho + \frac{p}{c^2}\right) U^\mu U^\nu - p g^{\mu\nu}. \tag{1.5}$$

The energy-momentum tensor treats the components contributing to the energy and density as perfect fluids (an inviscid fluid described only by its energy density and isotropic pressure p). The importance of this will become clear later on in section 1.3.1. The energy-momentum tensor vanishes in the absence of a component with a pressure and/or a density.

In GR, the energy-momentum of spacetime is identified as the source of its curvature. The Einstein **Field Equation** equates spacetime curvature with the

energy-momentum within that spacetime,

$$G^{\mu\nu} = -\frac{8\pi G}{c^4} T^{\mu\nu}. \tag{1.6}$$

Here, the Einstein tensor $G^{\mu\nu}$ incorporates the curvature information through a tensor known as the Ricci curvature tensor, $R^{\mu\nu}$ and the curvature scalar R,

$$G^{\mu\nu} = R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R, \tag{1.7}$$

with $R_{\beta\gamma} = R^{\alpha}_{\beta\gamma\alpha}$ and $R = R^{\mu}_{\mu} = g^{\mu\nu} R_{\mu\nu}$. The Ricci curvature tensor is a contraction of the Riemann curvature tensor, which provides a full description of the curvature of spacetime at each point compared to a Euclidean space [7].

The Robertson-Walker Metric

In order to describe an expanding, isotropic and homogeneous space-time we may use the Robertson-Walker (RW) space-time metric, which can be written in several different forms. Here we choose a form that makes the comoving coordinates[4] dimensionless, as in [7],

$$\mathrm{d}s^2 = c^2 \mathrm{d}\tau^2 = c^2 \mathrm{d}t^2 - a^2(t)[\mathrm{d}r^2 + S_k^2(r)\mathrm{d}\Omega^2], \tag{1.8}$$

where $\mathrm{d}s^2$ is the invariant space-time interval between two events, $a(t)$ is the time-dependent **scale factor** which is a dimensionless measure of the size of the Universe ($a(t_0)$ is normalised to 1 at the present day) and r is the time-independent comoving radial coordinate. Spherical symmetry allows the spatial part of the metric to be decomposed into a radial and a transverse part ($\mathrm{d}\Omega^2 = \mathrm{d}\theta^2 + \sin^2 \mathrm{d}\phi^2$). Finally, we define the function

$$S_k(r) = \begin{cases} \sin r & (k = 1) \\ \sinh r & (k = -1) \\ r & (k = 0). \end{cases}$$

[4]Comoving observers are at rest with respect to the matter in their vicinity and thus observe an isotropically expanding conformal Universe. Comoving coordinates are useful distance measures as they do not change over time and are defined by the worldlines of particles moving with cosmic expansion (see later in this section).

k corresponds to the curvature parameter of the Universe: $k = +1$ for a closed Universe with positive spatial curvature, $k = -1$ for an open Universe with negative spatial curvature and $k = 0$ for a flat Universe that has no curvature.

The Friedmann Equation

Einstein's gravitational field equations (Eqn. 1.6) relate the evolution of the scale factor to the pressure, p, and energy density, ρ, of the mass-energy components of the Universe. The Friedmann equation is derived from this and describes the evolution of the scale factor in relation to p and ρ from the field equations for a Universe described by the RW metric [10, 11],

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2}. \tag{1.9}$$

Its derivative is the acceleration equation

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right). \tag{1.10}$$

It should be noted that the energy density is the total energy density arising from the matter, radiation and vacuum energy densities,

$$\rho = \rho_m + \rho_r + \rho_\Lambda. \tag{1.11}$$

The significance of the vacuum energy will be discussed in section 1.3.

If the scale factor in Eqn. 1.8 satisfies the Friedmann equation, the RW metric is known as the Friedmann-Lemaître-Robertson-Walker (FLRW) metric. The Friedmann equation can be used to define the **Hubble parameter**, $H(t)$, which describes the rate of change of the scale factor,

$$H(t) = \frac{\dot{a}}{a}. \tag{1.12}$$

The value of the Hubble parameter at the present day is referred to as the Hubble constant and is often written as $H(t_0) = H_0$. In the absence of peculiar velocities, the Hubble parameter relates the velocity of recession of a galaxy, v, with its

proper distance[5] from the observer, D_{prop},

$$H_0 D_{\text{prop}} = v. \tag{1.13}$$

H_0 has units of $\text{kms}^{-1}\text{Mpc}^{-1}$, so it is often useful to express the Hubble constant in terms of the dimensionless Hubble parameter, h, such that

$$H_0 = 100\, h_0\, \text{kms}^{-1}\text{Mpc}^{-1}. \tag{1.14}$$

Recent observations suggest that $h_0 \sim 0.74$ [12]. Galaxies whose recession velocities dominate over their peculiar velocities are said to be moving with the Hubble flow.

With the definition of the Hubble parameter it is possible to rewrite the Friedmann equation in the form

$$\frac{8\pi G\rho}{3H^2} - \frac{kc^2}{\dot{a}^2} = 1, \tag{1.15}$$

such that there exists a **critical density** needed for a flat ($k = 0$) Universe,

$$\rho_c = \frac{3H^2}{8\pi G}. \tag{1.16}$$

This allows us to further define the **density parameter** as simply the ratio of the density (of ρ_m, ρ_r, ρ_Λ or the total density ρ) to the critical density,

$$\Omega \equiv \frac{\rho_x}{\rho_c} = \frac{8\pi G\rho_x}{3H^2}. \tag{1.17}$$

It is obvious that for a flat Universe the total density parameter will be given as

$$\Omega \equiv \Omega_m + \Omega_r + \Omega_\Lambda = 1. \tag{1.18}$$

A Universe with spatial curvature will have $\Omega \neq 1$ and Ω will also contain a contribution from the curvature, Ω_c ($\Omega_c = 0$ if and only if $k = 0$).

[5]Proper distance is defined as the light travel time between two points multiplied by c (see later in this section).

Cosmological Redshift

The Universe is expanding as a consequence of the initial conditions (and more recently the effect of dark energy, see section 1.3.1)[6]. Photons reaching us from cosmological distances will have had their wavelengths stretched. This can be shown from the RW metric for a null geodesic ($\mathrm{d}s^2 = 0$),

$$r = c \int_{t_e}^{t_o} \frac{\mathrm{d}t}{a(t)}, \tag{1.19}$$

for a light ray emitted at time t_e and received at the later time t_o. The comoving coordinate r is constant; altering the limits on the integral (equivalent to emitting photons from r at later times) cannot alter the integral itself. This means that the photons will undergo time dilation dependent on how much the Universe has expanded between t_e and t_o, which leads to the condition

$$\frac{\mathrm{d}t_e}{\mathrm{d}t_o} = \frac{a(t_e)}{a(t_o)} \tag{1.20}$$

Redshift is usually denoted by z and is defined in terms of the ratio of the shift in the frequency of light,

$$1 + z = \frac{\nu_e}{\nu_o}. \tag{1.21}$$

Combining Eqn. 1.20 and Eqn. 1.21 gives us the Lemaître wavelength rule:

$$1 + z = \frac{\nu_e}{\nu_o} = \frac{1}{a(t)} \tag{1.22}$$

since $a(t_o) = 1$. The redshift is essentially a measure of the scale factor at the time of emission. It is useful to note that on small scales where $v \ll c$, redshift can be expressed in terms of the recession velocity,

$$1 + z \approx 1 + \frac{v}{c} \tag{1.23}$$

[6]To state that space or spacetime itself is expanding is incorrect, since spacetime can always be locally described by a flat, static Minkowski metric. Redshift is the result of the accumulation of a series of infinitesimal Doppler shifts due to the recession of galaxies moving with the Hubble flow. For a discussion see [13, 14].

and equivalently the velocity-distance law approximates to the Hubble law,

$$H_0 D_{\text{prop}} = v \approx cz. \tag{1.24}$$

Currently, the highest redshift galaxies discovered to date have $z \sim 8.6$ [15].

Cosmological Distance Measures

On cosmological scales, spacetime is curved and a measure of the 'distance' between two points becomes ambiguous. One must carefully define exactly what is meant when discussing the separation of two objects to avoid confusion. The most common distance measures in astronomy are described below.

Comoving distance is given in Eqn. 1.19 and is the distance between the worldlines of two events measured at $t = t_0$. It can be written

$$\mathrm{d}D_{\text{com}} = \mathrm{d}r = \frac{c\mathrm{d}t}{a} = \frac{c\mathrm{d}a}{a^2 H(a)} \tag{1.25}$$

so that

$$D_{\text{com}}(z_1, z_2) = \int_{a_1(z_2)}^{a_2(z_1)} \frac{c\mathrm{d}a}{a^2 H(a)} = \int_{z_1}^{z_2} \frac{c\mathrm{d}z}{H(z)}. \tag{1.26}$$

For an observer at the origin ($z_1 = 0$) the comoving distance to an object with redshift z will be

$$D_{\text{com}}(z) = \int_0^z \frac{c\mathrm{d}z'}{H(z')}. \tag{1.27}$$

The comoving distance has the useful property of being time-invariant through its direct dependence on the scale factor.

The **angular diameter distance** relates an object's actual size to its apparent angular size,

$$D_{\text{ang}}(z_1, z_2) = \frac{\mathrm{d}A}{\mathrm{d}\Omega}, \tag{1.28}$$

where $\mathrm{d}A$ is the cross-sectional area of an object and $\mathrm{d}\Omega$ is the solid angle the object subtends as seen by an observer. From the RW metric, for an object at redshift z_2, $\mathrm{d}A$ is equivalent to a sphere with radius $R = a(z_2) S_k[D_{\text{com}}(z_1, z_2)]$, surface area $4\pi R^2 = 4\pi a^2(z_2) S_k^2[D_{\text{com}}(z_1, z_2)]$; the angle subtended at z_1 will be 4π so

$$\frac{\mathrm{d}A}{\mathrm{d}\Omega} = \frac{4\pi a^2(z_2) S_k^2[D_{\text{com}}(z_1, z_2)]}{4\pi}. \tag{1.29}$$

Inserting this into Eqn. 1.28 tells us that for an observer at the origin, the angular diameter distance is given by

$$D_{\rm ang}(z) = \frac{S_k[D_{\rm com}(z)]}{1+z}. \tag{1.30}$$

Another distance measure commonly used is the **luminosity distance**. This is a measure of the distance an object would be at if the inverse-square law for brightness always held. Luminosity distance is defined by the observed flux F and luminosity L of a source:

$$D_{\rm lum} = \left(\frac{L}{4\pi F}\right)^{\frac{1}{2}}. \tag{1.31}$$

Luminosity distance can be written in terms of $D_{\rm ang}$ and $D_{\rm com}$,

$$D_{\rm lum}(z_1, z_2) = \left[\frac{a(z_1)}{a(z_2)}\right]^2 D_{\rm ang}(z_1, z_2) = \frac{a(z_1)^2}{a(z_2)} S_k[D_{\rm com}(z_1, z_2)]. \tag{1.32}$$

Eqn. 1.32 illustrates why the inverse-square law does not hold at all distances; photons undergo a cosmological redshift of $a(z_1)/a(z_2)$, and a shift due to time dilation of the same magnitude. The expansion of the Universe also means that any sphere receiving photons will increase in surface area by a factor of $[a(z_1)/a(z_2)]^2$, altering the flux by a factor of $[a(z_1)/a(z_2)]^4$ and hence $D_{\rm lum}$ by $[a(z_1)/a(z_2)]^2$. Again, one can write this for an observer at the origin:

$$D_{\rm lum}(z) = (1+z) S_k[D_{\rm com}(z)]. \tag{1.33}$$

Finally, the **proper distance**, $\mathrm{d}D_{\rm prop}$, between two points is defined as the light travel time (multiplied by c) between them,

$$\mathrm{d}D_{\rm prop} = c\mathrm{d}t = c\frac{\mathrm{d}a}{aH(a)}. \tag{1.34}$$

Integrating this gives

$$D_{\rm prop}(z_1, z_2) = \int_{z_1}^{z_2} \frac{c\mathrm{d}z}{H(z)(1+z)} \tag{1.35}$$

for two events at redshifts z_1 and z_2. This measure is not often used as it is

time-dependent and difficult to determine [7, 9].

Every distance measure depends on the underlying cosmology, leading to inherent uncertainties in any cosmological distance measurement. For this reason, it is common for astronomers to use the redshifts of objects to describe their positions, as these depend only on the RW metric (and hence GR) being correct. It is interesting to note that Taylor expanding any one of these distance measures to first order yields $D(z) = \frac{cz}{H_0} + O(z^2)$, so locally the Hubble law applies.

The Flatness Problem

Combining Eqn. 1.9 with Eqn. 1.18 and rearranging gives us the following expression for the density parameter:

$$\Omega(t) - 1 = \frac{kc^2}{H^2(t)a^2}. \tag{1.36}$$

From this a fine-tuning problem is apparent. If the value of Ω is equal to 1 initially, then it will remain equal to 1 indefinitely. However, this is not the case if $\Omega \neq 1$ initially; using the solutions for a and H for matter and radiation in section 1.3.1, Ω can be written

$$\begin{aligned} |\Omega(t) - 1|_r &\propto t \\ |\Omega(t) - 1|_m &\propto t^{2/3}. \end{aligned} \tag{1.37}$$

One can see that in general, $|\Omega - 1|$ is amplified with time. Initial values of $\Omega > 1$ will cause Ω to increase with time, and vice versa. Another way to understand the flatness problem is that unless k is exactly equal to 1 (i.e. the Universe is initially perfectly flat), then in a Universe like ours with initially large fractions of matter and/or radiation, any curvature will become more exaggerated over time, driving the density to either a very large or very small value. Is it then cosmic coincidence that we observe the total density parameter to be consistent with 1? The inflationary model ensures that we do not have to rely on such a coincidence; during the inflationary period the scale factor has the de Sitter form:

$$a(t) = e^{\left(\frac{\Lambda_I}{3}\right)^{1/2} t} \tag{1.38}$$

meaning that at early times,

$$|\Omega(t) - 1| \propto e^{-\left(\frac{4\Lambda_I}{3}\right)^{1/2} t}. \tag{1.39}$$

Λ_I is the approximate energy density of the Inflationary field, or the 'Inflaton'. This result has the effect of driving Ω extremely close to 1 at early times (as the exponential expansion has the effect of smoothing out deviations from flatness), meaning that subsequent expansion at late times has very little effect on the value of $|\Omega - 1|$ in comparison. Thus the addition of inflation to the Big Bang paradigm ensures a solution to the flatness problem.

1.2 Structure Formation

Structure forms because of gravitational collapse. Inflation in the early Universe grew quantum fluctuations to sizes larger than the horizon scale, producing 'seed' fluctuations from the initially homogeneous Universe, around which matter could begin to collapse. Once a region begins to collapse, the relative density of that region increases causing further collapse in a form of positive feedback [7, 16]. The density of a region compared to the mean density $\bar{\rho}$ at space coordinate $\boldsymbol{x}$ can be expressed as a density contrast,

$$\delta(\boldsymbol{x}) = \frac{\rho(\boldsymbol{x}) - \bar{\rho}}{\bar{\rho}}. \tag{1.40}$$

$\delta(\boldsymbol{x})$ is the dimensionless density perturbation of the underlying matter distribution. Two types of density perturbations can occur: **adiabatic** and **isocurvature** [17, 7]. Adiabatic perturbations occur if a fluid is adiabatically compressed in space, whereas isocurvature perturbations occur if the entropy density is perturbed but the energy density is not. If $\delta(\boldsymbol{x}) \ll 1$, adiabatic density fluctuations will grow due to the effect of gravity,

$$\begin{aligned} \delta(a) &\propto a^2, \quad a < a_{\text{eq}} \\ \delta(a) &\propto a, \quad a > a_{\text{eq}}. \end{aligned} \tag{1.41}$$

Where a_{eq} denotes the time of transition between radiation and matter domination, i.e. when matter and radiation densities are equal. This behaviour

is a consequence of the fact that energy densities of matter and radiation are affected by adiabatic perturbations in different ways. The energy density for radiation is $\propto T^4 \propto a^{-4}$, whereas the number density is $\propto T^3 \propto a^{-3}$ (See §1.3.1). Poisson's equation relates the Newtonian potential Φ to the density: $-k^2\Phi/a^2 \propto \rho\delta$, meaning that $\Phi \propto \delta/a$ for matter and $\Phi \propto \delta/a^2$ for radiation [7, 18]. This is valid for an Einstein-de-Sitter (EdS, matter dominated) model with zero curvature and dark energy. For isocurvature perturbations, at early times any expansion acts to preserve the initial density:

$$\begin{aligned} \delta(a) &\propto \text{constant}, \quad a < a_{\text{eq}} \\ \delta(a) &\propto a(t)^{-1}, \quad a > a_{\text{eq}}. \end{aligned} \tag{1.42}$$

In both cases, the overall shape of the spectrum of perturbations is preserved, but the amplitude changes. At later times, when $a \gg a_{\text{eq}}$, the density perturbations will behave according to some growth function $g(a)$,

$$\delta(a) = \delta_0 a \frac{g(a)}{g(a_0)}, \tag{1.43}$$

where $g(a)$ is defined in [19] as a function of the matter density parameters Ω_m and Ω_Λ.

Before the era of matter domination, the expansion timescale is shorter than the collapse timescale, meaning that the fast, radiation-driven $\rho_r \propto a^{-4}$ (see section 1.3.1) expansion will prevent the growth of matter perturbations that are within the horizon size, d_H. Larger scale perturbations with $\lambda > d_H$ will remain unaffected by this and will continue to grow [7, 20]. The growth of fluctuations with $\lambda < d_H$ is suppressed by the factor

$$f_{\text{sup}} = \left(\frac{a_{\text{ent}}}{a_{\text{eq}}}\right)^2, \tag{1.44}$$

where a_{ent} is the time at which a perturbation enters the horizon.

As the horizon scale expands with time and more regions are causally connected, more perturbations on larger scales enter the horizon and their growth is stalled. Once matter domination begins, collisionless cold dark matter (see section 1.3.2) can collapse under gravity (although baryonic matter is ionic and still bound and oscillating with the photons). It is later, at the era

of recombination, that baryons can decouple from photons and fall into the potential wells created by the dark matter collapse. This process gives rise to a picture of structure formation in which baryonic galaxies form at the centre of large dark matter haloes. As the horizon scale increases with time, larger and larger fluctuations become causally connected, allowing superstructures such as filaments to form [21]. This formation model relies on dark matter being cold and non-relativistic and is known as 'bottom-up' formation; if dark matter were hot and relativistic, the small scale fluctuations would be completely removed by diffusion. This would result in only the largest scale fluctuations remaining, and a 'top-down' model of structure formation, which does not fit well with observations [20, 22].

1.2.1 The Matter Power Spectrum

The cosmological principle ensures that the statistical properties of the density field are homogeneous. The density field is assumed to be a Gaussian random field since the Fourier modes of the initial quantum (seed) fluctuations are uncorrelated. Such a field can be described by a **3D matter power spectrum** [23, 7],

$$P_\delta(|\mathbf{k}|) = \int \mathrm{d}^3\boldsymbol{x} e^{-i\boldsymbol{x}.\boldsymbol{k}} C_{\delta\delta}(|\boldsymbol{x}|), \tag{1.45}$$

where k is the wavenumber and is related to the comoving wavelength of fluctuation λ by $k = 2\pi/\lambda$. $P_\delta(|\mathbf{k}|)$ is often abbreviated to $P_\delta(k)$ where $k = |\boldsymbol{k}|$. We show how this equation is derived from the two-point correlation function of a Gaussian field in §2.3.3.

The power spectrum is often quoted in a dimensionless form:

$$\Delta^2(k) = \frac{k^3 P_\delta(k)}{2\pi^2}. \tag{1.46}$$

To determine the shape and behaviour of the power spectrum, one must consider the behaviour of the density fluctuations that give rise to it. For an EdS model, a density fluctuation enters the horizon at time t_{ent} when

$$\lambda = d_H(a_{\mathrm{ent}}) = \frac{c}{a_{\mathrm{ent}} H(a_{\mathrm{ent}})}. \tag{1.47}$$

At matter-radiation equality, $\rho_r = \rho_m$ and hence $a_{\mathrm{eq}}^{-4}\Omega_R = a_{\mathrm{eq}}^{-3}\Omega_M$ (where capital

subscripts denote the values at $t = t_0$), meaning that a fluctuation that enters the horizon will have wavelength dependences on the scale factor as follows:

$$\begin{aligned} \lambda &\propto a_{\rm ent} \quad (t_{\rm ent} \ll t_{\rm eq}) \\ \lambda &\propto \sqrt{a_{\rm ent}} \quad (t_{\rm eq} \ll t_{\rm ent} \ll t_0). \end{aligned} \tag{1.48}$$

Given the primordial power spectrum $P_i(k)$, the growth equation, Eqn. 1.41 shows how the power spectrum has changed by $t_{\rm ent}$:

$$\begin{aligned} P_{\delta,{\rm ent}}(k) &\propto a_{\rm ent}^4 P_i(k) \propto k^{-4} P_i(k) \quad (t_{\rm ent} \ll t_{\rm eq}) \\ P_{\delta,{\rm ent}}(k) &\propto a_{\rm ent}^2 P_i(k) \propto k^{-2} P_i(k) \quad (t_{\rm eq} \ll t_{\rm ent} \ll t_0). \end{aligned} \tag{1.49}$$

The total power of the density fluctuations at $t_{\rm ent}$ is assumed to be scale invariant; that is, $k^3 P_{\rm ent}(k) =$ constant. This is the Harrison-Peebles-Zel'dovich spectrum [24, 25]. It implies that the primordial power spectrum scales with k as $P_i(k) \propto k^{n_s}$ where n_s is known as the slope of the primordial power spectrum and is expected to be 1 [7]. CMB results find $n_s = 0.967 \pm 0.014$, consistent with unity [26].

Eqn. 1.44 shows how the growth of a perturbation will be suppressed as a function of scale factor. Defining k_0 to be the wavenumber corresponding to the horizon size at $a_{\rm eq}$, the suppression factor may be written

$$f_{\rm sup} = \left(\frac{k_0}{k}\right)^2. \tag{1.50}$$

Since the matter power spectrum is related to the primordial power spectrum through the suppression factor such that $P_\delta(k \gg k_0) = f_{\rm sup}^2 P_i(k)$, the power spectrum at the point of matter-radiation equality is therefore

$$\begin{aligned} P_\delta(k) &\propto k \quad k \ll k_0 \\ P_\delta(k) &\propto k^{-3} \quad k \gg k_0. \end{aligned} \tag{1.51}$$

The general form of this can be seen in Fig. 1.1. After $t_{\rm eq}$ the shape of the power spectrum is modified by non-linear effects (see section 1.2.2) and it grows in amplitude according to the growth factor.

1.2.2 The Non-Linear Power Spectrum

On small scales (at large k) the linear evolution of the power spectrum begins to break down. Non-linearity corrections must be applied to take into account the fact that collapsing structures do not go on collapsing forever and do, in fact, stabilise. This has the effect of increasing the amplitude of the power spectrum at large k (see Fig. 1.1). The modification to the power spectrum due to non-linearities has been estimated from N-body simulations; [27] describe an analytic function that relates the non-linear power spectrum to the linear one via a fitting function f_{nl}:

$$\Delta^2_{nl}(k_{nl}) = f_{nl}[\Delta^2_l(k_l)]. \tag{1.52}$$

where f_{nl} is defined in Eqn. 21 in [27], l denotes 'linear' and nl denotes 'non-linear'. The linear wavenumber k_l is related to the non-linear wavenumber k_{nl} via

$$k_l = [1 + \Delta^2_{nl}(k_{nl})]^{-1/3} k_{nl}. \tag{1.53}$$

More recently, the 'halo model' was developed by [28], in which the density field is decomposed into clumps of matter with varying mass and some density profile. Large-scale clustering of mass arises through correlations between different haloes. The halo model non-linear power spectrum is defined as

$$\Delta^2_{nl}(k) = \Delta^2_Q(k) + \Delta^2_H(k), \tag{1.54}$$

where Q is a quasi-linear term that denotes the power generated from the correlation of large-scale haloes with each other and H represents the power resulting from self-correlation of dark matter haloes. The halo model has been shown to be a better fit than the Peacock & Dodds formula in [27] when applied to N-body simulations. The 'halofit' code is publicly available[7] and can be used to provide accurate halo model predictions of the non linear matter power spectrum. Fig. 1.1 compares the expected linear power spectrum the the non-linear one predicted by halofit.

[7] from `http://www.roe.ac.uk/~jap/haloes/`

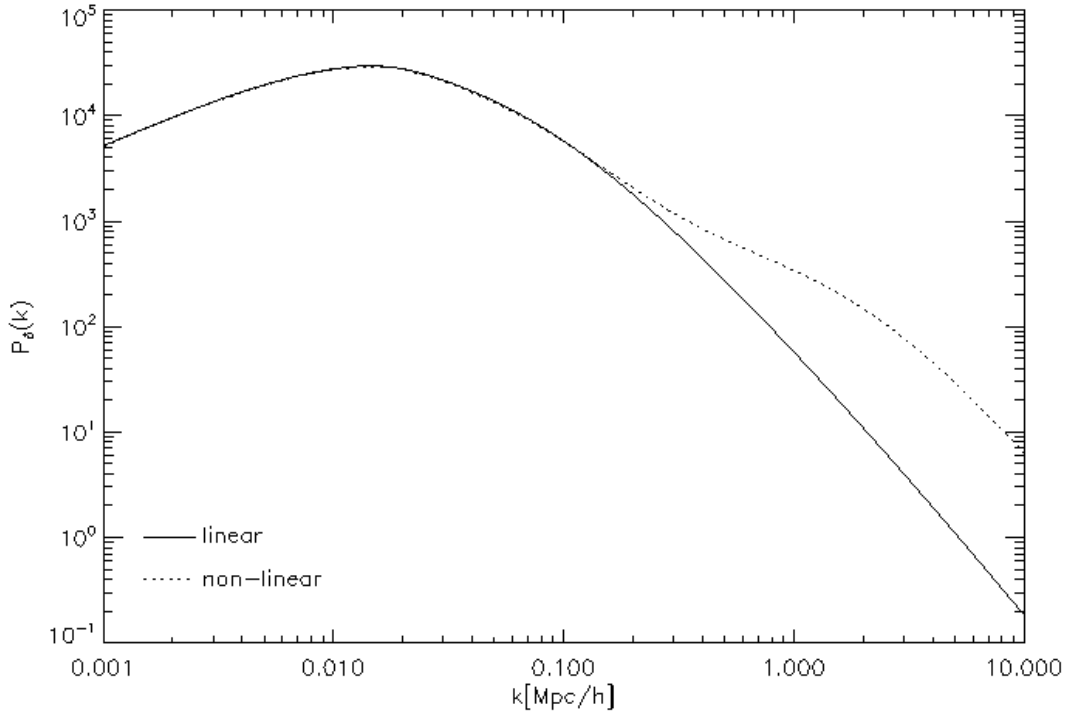

Figure 1.1: The linear and non-linear matter power spectrum prediction for the standard model from halofit.

1.2.3 Normalisation and σ_8

To measure the amplitude of the power spectrum, the usual approach is to use the normalised quantity σ_8, defined as the rms variation of the density field in a tophat sphere of radius $8h^{-1}$Mpc [29, 30, 31]. Generally this can be written

$$\sigma^2(R,z) = \int_0^\infty \frac{\mathrm{d}k}{k}\Delta^2(k,z)W^2(kR) \tag{1.55}$$

where W is a weighting function defined as the Fourier transform of a spherical tophat filter,

$$W(k,R) = \frac{3J_1(kR)}{kR}. \tag{1.56}$$

J_1 is a first order Bessel function and $\sigma_8 = \sigma(8h^{-1}\,\mathrm{Mpc})$. σ_8 describes the linear clustering of matter and its amplitude can be measured with a number of methods including weak lensing [9, 32].

1.3 The ΛCDM Universe

Until the 1990s, the cosmological formalism that we have developed in the preceding section was thought to be more or less complete. The Universe was thought to resemble an EdS model: matter-dominated at late times, inflationary, flat[8] and undergoing continued *decelerating* expansion. There are two main components to the accepted paradigm that have been alluded to but not yet described: dark energy and dark matter. Together, dark energy ('Λ') and cold dark matter ('CDM') make up ∼96% of the density of the ΛCDM Universe.

The joint discoveries of [34] and [35] that distant type Ia supernovae (SNe Ia) were fainter than expected changed our picture of the Universe. SNe Ia are used as **standard candles** in astronomy because their light curves are very well understood such that the observed magnitude of a SN Ia is dependent on its luminosity distance. The SNe Ia observations in 1998 showed that at a given luminosity distance, objects were more redshifted than would be expected for a

[8]The evidence for a flat Universe comes from probes such as the Wilkinson Microwave Anisotropy Probe (WMAP) and the balloon-borne BOOMERanG, which measure temperature anisotropies in the CMB. The size of these fluctuations is well–known from plasma physics, allowing the angular scale of the anisotropies and hence the curvature of the Universe to be measured. The WMAP 7-year results find the curvature density parameter to be $\Omega_k = -0.0057^{+0.0067}_{-0.0068}$ (68% confidence limit, CL), consistent with a flat Universe [26, 33].

decelerating expansion; the SNe Ia are receding at a higher velocity than expected and the expansion of the Universe on cosmic scales must be *accelerating*. It is dark energy that is posited to be the cause of the acceleration.

The identity of dark energy is as yet unknown, but the simplest explanation is vacuum energy, or zero point energy, denoted by Λ. Vacuum energy arises from quantum field theory and can be thought of as the potential energy of particle–anti particle pairs that appear and annihilate in the quantum vacuum within very short timescales limited by the Uncertainty Principle. Vacuum energy has been shown to be real and measurable by the Casimir effect [36]. The energy of the vacuum is constant throughout space and time hence its alternative name, the **cosmological constant**. A convincing particle physics explanation for its magnitude currently eludes us, however. The cosmological constant enters the field equation (Eqn. 1.6) in the form

$$G^{\mu\nu} + \Lambda g^{\mu\nu} = -\frac{8\pi G}{c^4} T^{\mu\nu} \tag{1.57}$$

and is defined as [7]

$$\Lambda = \frac{8\pi G \rho_\Lambda}{3c^2}. \tag{1.58}$$

Originally, Einstein included a cosmological constant term as a correction to the geometric part of the field equation, as the GR formalism predicted a non-static Universe whereas Einstein believed the Universe was static. Without Λ, gravity would cause a universe initially in dynamic equilibrium to contract. Λ acts to produce gravitational repulsion, allowing Einstein's universe to remain static. When evidence of cosmological redshift emerged (see section 1.1.1), Einstein withdrew his support for the cosmological constant. However later scientists such as Zel'dovich, Lemaître and Eddington began to associate the cosmological constant with vacuum energy, leading to its continued use to represent vacuum energy today [9].

1.3.1 Dark Energy & Expansion

In order to produce accelerated expansion at late times, dark energy must have negative pressure. According to GR the pressure within a substance contributes to its gravity just as its mass density does, hence the inclusion of a pressure term in the Friedmann equation. If dark energy has a negative pressure sign then

it acts on cosmic scales to push objects apart; in the vacuum energy model, as spacetime expands there will an increasing amount of dark energy between two comoving objects and the cosmological expansion will be driven ever faster. The amount of negative pressure that dark energy must have is determined by the cosmological **equation of state** of dark energy, which can be derived if dark energy is assumed to be a perfect fluid. We may treat the Universe as a closed, adiabatic system and apply the first law of thermodynamics to it,

$$\mathrm{d}U = -p\mathrm{d}V, \tag{1.59}$$

where U is the energy of the system of volume V and under pressure p. The volume will simply scale as a^3 since a has the dimensions of length, and U can be written in terms of the energy density, $U = \rho c^2 V$. Using these relations and the first law we find the fluid, or continuity, equation [7, 37]

$$c^2 \frac{\mathrm{d}\rho}{\mathrm{d}a} + 3\left(\frac{p + \rho c^2}{a}\right) = 0. \tag{1.60}$$

This describes how the density and pressure of a fluid are related to one another, and how they will evolve over the age of the Universe. In the context of cosmology, the equation of state for a perfect fluid is as

$$p = w\rho c^2 \tag{1.61}$$

in which w is known as the **equation of state parameter** and will vary for each of the contributors to the mass–energy content of the Universe:

$$\begin{aligned} w_m &\approx 0 \text{ (matter)} \\ w_r &= \frac{1}{3} \text{ (radiation)} \\ w_\Lambda &\approx -1 \text{ (cosmological constant).} \end{aligned} \tag{1.62}$$

It is useful to treat the components of the Universe as separate fluids in this way, as they are both physically and mathematically distinct; they represent different physical phenomena whose densities evolve differently and independently.

For dark energy to have enough negative pressure to drive the acceleration of the Universe, it must overcome the positive pressure from radiation such that

$w_\Lambda < -\frac{1}{3}$. Solving Eqn. 1.60 using the equation of state gives the scale-factor dependence of the density of each component,

$$\begin{aligned} \rho_m &\propto a^{-3} \\ \rho_r &\propto a^{-4} \\ \rho_\Lambda &= \text{constant.} \end{aligned} \tag{1.63}$$

So only a value of $w = -1$ gives us the constant energy density required for a cosmological constant. Eqn. 1.63 tells us that in the early Universe, radiation dominated the mass–energy density. As ρ_r falls faster than ρ_m by a factor of a (due to the fact that radiation undergoes redshifting with the expansion of the Universe and matter does not), eventually a time came when $\rho_r = \rho_m$ and the matter density began to dominate. Since a is always increasing, the matter density eventually drops enough for $\rho_m = \rho_\Lambda$ and dark energy begins to dominate the mass–energy content of the Universe. It is this dark energy dominated era that we occupy at the present day.

The time evolution of the density of each component can be found by substituting the identities in Eqn. 1.63 into the Friedmann equation (assuming flatness):

$$\begin{aligned} &a(t) \propto t^{2/3} && \text{(matter)} \\ &a(t) \propto t^{1/2} && \text{(radiation)} \\ &a(t) \propto e^{\left(\frac{8\pi G\rho_\Lambda}{3}\right)^{1/2} t} && \text{(cosmological constant).} \end{aligned} \tag{1.64}$$

From this, using the definition of the Hubble parameter in Eqn. 1.12 one can determine the behaviour of the Hubble parameter in universes dominated by matter, radiation and dark energy respectively:

$$\begin{aligned} &H(t) = 2/3t && \text{(matter)} \\ &H(t) = 1/2t && \text{(radiation)} \\ &H(t) = \text{constant} && \text{(cosmological constant).} \end{aligned} \tag{1.65}$$

A matter-dominated universe behaves like an Einstein-de Sitter universe, and would undergo decelerating expansion. A radiation-dominated universe would

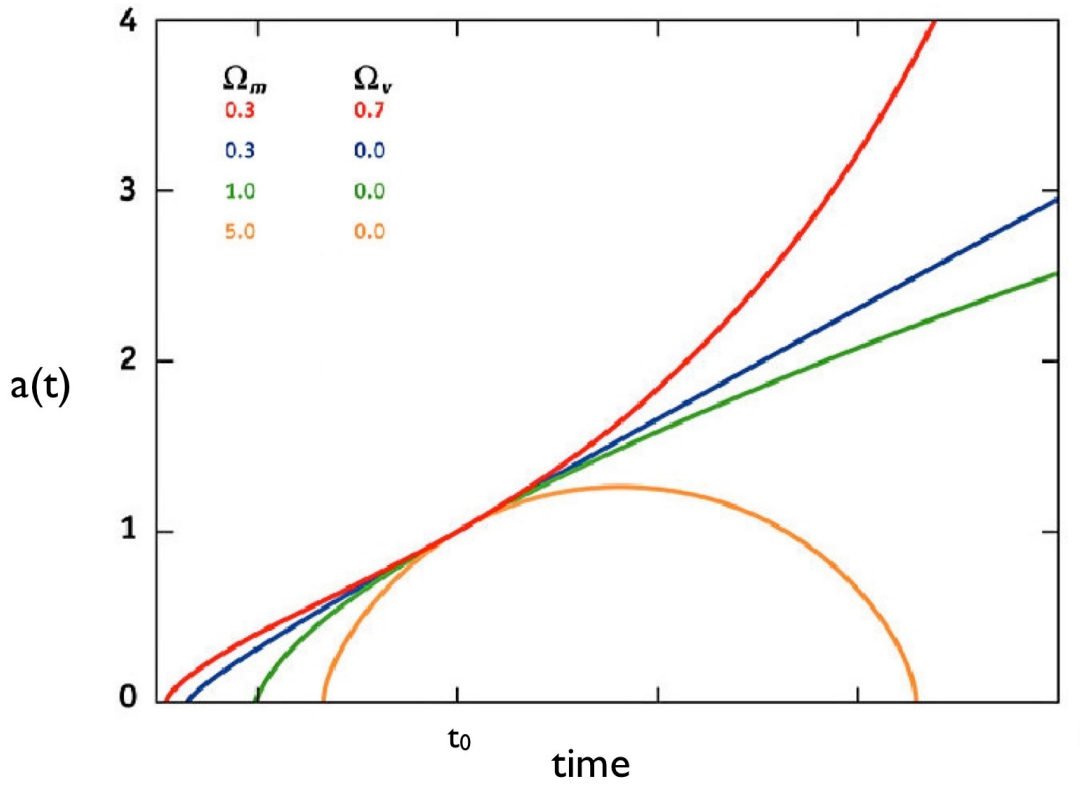

Figure 1.2: Behaviour of the scale factor for different cosmologies. ΛCDM is shown in red [38].

also experience decelerated expansion (as our early universe did). The solution for the dark-energy dominated universe asymptotes towards a **de Sitter** solution; in such a model the expansion of the universe accelerates exponentially. As dark energy is the dominant contributor to the density in the present-day Universe, it appears that our Universe will behave like a de Sitter universe in the future. The combined effects of the different density components of our Universe on the expansion rate can be seen in Fig. 1.2.

The Hubble Parameter

It is worth noting that a cosmological constant term will alter the Friedmann equations:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2} + \frac{\Lambda}{3}, \tag{1.66}$$

where

$$\Omega_{de} = \frac{\Lambda}{3H^2}. \tag{1.67}$$

The Friedmann equation can also be rewritten using Eqns. 1.63 and 1.64 as [39, 7]

$$H^2(a) = \frac{8\pi G}{3}(\rho_m a^{-3} + \rho_r a^{-4} + \rho_{de} e^{-3\int_a^1 [1+w(a')]\mathrm{d}(\ln a')}) - \frac{kc^2}{a^2}, \tag{1.68}$$

where $w(a)$ is a generic scale-dependent function for the dark energy equation of state parameter that is dependent on the nature of dark energy. This can be combined with Eqns. 1.17 and 1.36 to give

$$H^2(a) = H_0^2 \left[\Omega_m a^{-3} + \Omega_r a^{-4} + \Omega_{de} e^{-3\int_a^1 [1+w(a')]\mathrm{d}(\ln a')} + (1-\Omega)a^{-2}\right]. \tag{1.69}$$

This can be expressed in terms of redshift:

$$H^2(a) = H_0^2 \left[\Omega_m (1+z)^3 + \Omega_r (1+z)^4 + \Omega_{de} e^{3\int_0^z [1+w(z')]/(1+z')\mathrm{d}z'} + (1-\Omega)(1+z)^2\right]. \tag{1.70}$$

These equations are useful as they relate the distance of an object to its redshift. They also highlight how distance measures are dependent on the density components and geometry of the Universe.

Evidence for Dark Energy

In addition to the SNe Ia observations of 1998, further evidence for the existence of dark energy comes from follow up surveys of SNe Ia (from e.g. [40]) and from CMB measurements. The most precise CMB measurements come from the Wilkinson Microwave Anisotropy Probe. WMAP is a space-based NASA mission that launched in July 2001 and has now stopped collecting data. WMAP was launched to follow in the footsteps of the Cosmic Background Explorer (COBE) and detect small anisotropies in the CMB. COBE had a large angular resolution of 7 degrees, and was able to detect anisotropy in the CMB at the level of ~ 1 part in 10^3. WMAP has much higher resolution ($\sim 13'$) and as a result has been able to measure anisotropies of the order $\sim 10^{-5}$. Data releases and analyses have been staggered at two year intervals, with the most recent data release (as of the time of writing) being the WMAP 7-year results, which are the most precise yet. Overall, WMAP data produce a consistently good fit to the ΛCDM paradigm with

some of the most precise parameter estimates of any cosmological probe. WMAP is superseded by the Planck spacecraft, launched in 2009, which will measure the power spectrum of the CMB to even smaller scales ($\sim$ 3 x WMAP). The **power spectrum of the CMB** has been measured to unprecedented precision by WMAP7 and used to show that the Universe is close to flat ($\Omega = 1.02 \pm 0.02$) as well as constraining the matter density to be $\Omega_m h^2 = 0.1345^{+0.0056}_{-0.0055}$ and the baryon density to be $\Omega_b = 0.0455 \pm 0.0028$ (all 68% CL) [26]. As a independent probe, the CMB thus shows that there is a large amount of missing mass-energy density that the matter content of the Universe cannot account for alone.

The **integrated Sachs-Wolfe effect** as measured from the CMB also points to an accelerating Universe, since potential wells will decay in an accelerating Universe. Large scale structure in the Universe creates potential wells in the region of galaxy clusters. Photons that enter such a potential well will gain some energy, but will not lose all that energy again upon leaving the potential well if the well has been stretched out — made shallower — by the expansion of the Universe. This leads to 'hot spots' and 'cold spots' on the sky depending on whether the photon passes through a potential well or through a potential hill (due to voids) respectively. This is the integrated Sachs-Wolfe effect and it occurs on large angular scales in the CMB. At the present time, cosmological measurements from the integrated Sachs-Wolfe effect are much less precise than from the CMB [41].

WiggleZ[9] [42], SDSS[10] [43] and BOSS[11] observations of hundreds of thousands of galaxies have measured the growth rate of **large scale structure** and the matter power spectrum (see section 1.2.1). These spectroscopic surveys allow the redshift of each survey galaxy to be determined. Assuming the galaxies are a good tracer for the underlying dark matter (as observations of galaxies and clusters implies, see section 1.3.2), constraints can be placed upon the matter power spectrum at different redshifts. WiggleZ find that $\Omega_m = 0.27 \pm 0.04(68\% CL)$, consistent with CMB measurements. Recent results from [44] combined BOSS results with WMAP 7yr data to find $\Omega_m = 0.298 \pm 0.017$. Again, when combined with measurements of the curvature of the Universe, this provides evidence that the dark energy density is non-zero.

[9] `http://wigglez.swin.edu.au/site/`.
[10] Sloan Digital Sky Survey, `http://www.sdss.org/`.
[11] `http://cosmology.lbl.gov/BOSS/`.

Observing the effect of **baryon acoustic oscillations** (BAOs) on the CMB power spectrum and the clustering of galaxies can constrain the density and equation of state of dark energy. In the early Universe, before the surface of last scattering that gave rise to the CMB, baryonic matter would have fallen into the potential wells of the dark matter. As more baryons fell into the potential wells, pressure within the wells would have accumulated to the point where the infalling baryons would start to 'bounce' out of the wells. Since this early Universe consisted of a relativistic plasma of particles, the sound speed was high and the oscillations in the in-fall of baryons were analogous to sound waves. The growth of these perturbations is governed by the Jeans scale: perturbations larger than the Jeans length are free to grow through gravity and smaller perturbations are prevented from collapse by pressure (as described in § 1.2). In the time between the formation of these perturbations and the epoch of recombination when photons can free-stream, modes of different wavelength can complete different numbers of oscillation periods, producing a series of maxima and minima in the CMB power spectrum [9, 45]. The density perturbations caused by BAOs cause matter (both baryonic and dark matter, since dark matter and baryons are coupled through gravitational interactions) to collect in a shell at the point where the photons began to free-stream. The distance from the centre of an overdense region to the radius at which photons free-stream is known as the sound horizon. The sound horizon hence represents the characteristic length scale at which galaxy pairs preferentially form. Since there are multiple superposed perturbations, there will be multiple length scales at which we expect to see overdensities of galaxy pairs. These length scales can be detected statistically by measuring the correlation function at different scales - one compares the measured density of galaxy pairs at a separation θ with that expected from randomly distributed galaxies [46, 45]. BAOs have been detected by the SDSS-III programme [47] and WiggleZ [48] who find that $w < -1/3$ (i.e. the Universe's expansion is accelerating) with 99.8% likelihood.

Finally, **weak lensing** is a relatively recently-exploited cosmological probe which is sensitive to the amount and behaviour of dark energy as the weak lensing distortions depend on the distance traveled and the growth of structure [49]. Weak lensing measurement is the focus of this thesis and will be discussed further in Chapter 2. We briefly compare lensing and BAOs as probes of cosmology. Both

lensing and BAOs are sensitive to the redshift and hence have constraining power on redshift-dependent parameters such as the dark energy equation of state. Both phenomena are sensitive to inaccuracies in photometric redshift distributions (see §2.3.5), although BAOs are particularly sensitive to this systematic [45]. The accuracy of BAO measurements depend on the reliability of measurements of the galaxy bias parameter b and the baryon density Ω_b, which affect the predicted BAO signal, whereas lensing arises solely as a consequence of general relativity and does not rely on complex astrophysics, making it a somewhat 'cleaner' probe. One key advantage of studying BAOs, however, is that nonlinearities in the signal (from e.g. nonlinear clustering, which affect both the BAO and lensing signal on small scales) induce predictable shifts in the oscillation scale. This means that nonlinearities can be modelled both analytically and through numerical simulations, allowing for accurate calibration of these effects. BAOs and weak lensing both have the ability to provide useful cosmological constraints, and will demand somewhat different survey configurations and analysis techniques.

Figs. 1.3-1.4 illustrate the complimentary of the different cosmological probes discussed. Fig. 1.3 shows how orthogonal constraints from weak lensing are with CMB constraints on the Ω_m-σ_8 contour. Although the joint Ω_m-σ_8 constraints from lensing are substantially larger than that from the CMB, the degeneracy in the results lies orthogonal to that of the CMB and hence combining the data sets results in even more precise constraints than from either probe alone. Since the lensing and CMB power spectra are independent of one another, comparing the results directly allows us to check for consistency between them. A similar result can be seen in Fig. 1.4 for the dark energy equation of state parameters, w_0 and w_a for a variety of cosmological probes; different cosmological probes produce joint constraints that vary in both overall precision and degeneracy, allowing very tight combined constraints to be established.

The Identity of Dark Energy

Recent observations suggest that $\Omega_\Lambda = 0.725 \pm 0.016$ and $-1.10 \leq w_\Lambda \leq 0.14$ [26], consistent with vacuum energy and the ΛCDM model. The nature of the dark energy equation of state is unknown; it is often parametrised as $w(a) = w_0 + w_a(1-a)$, where w_a is some time-dependent parameter [52, 26]. This formalism is simply a first order Taylor expansion designed to describe any time evolution

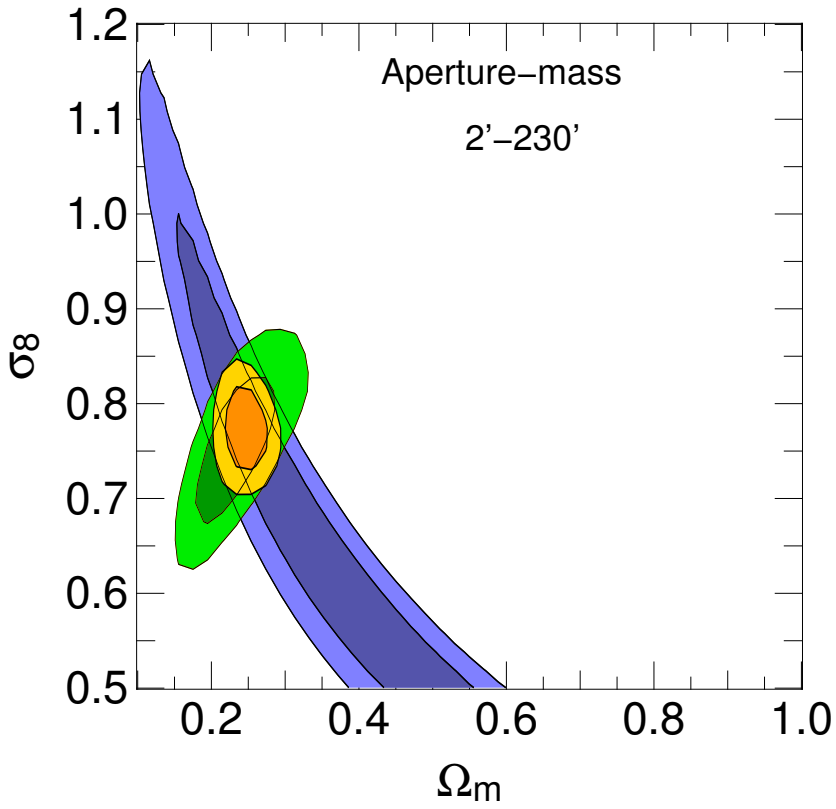

Figure 1.3: Typical 1 and 2σ constraints on the matter density parameter Ω_m and the amplitude of the matter power spectrum σ_8 from CFHTLS weak lensing (see §2.3.6, blue contours) and WMAP 3yr CMB measurements (green contours), as well as the combined constraints from both data sets (yellow contours) [50]. The constraints from lensing, while larger than those from the CMB, are orthogonal to CMB constraints and hence can be used to further narrow the parameter estimates. The weak lensing statistic used is the aperture mass statistic (see §4.1.1) measured over scales of $2' \leq \theta \leq 230'$.

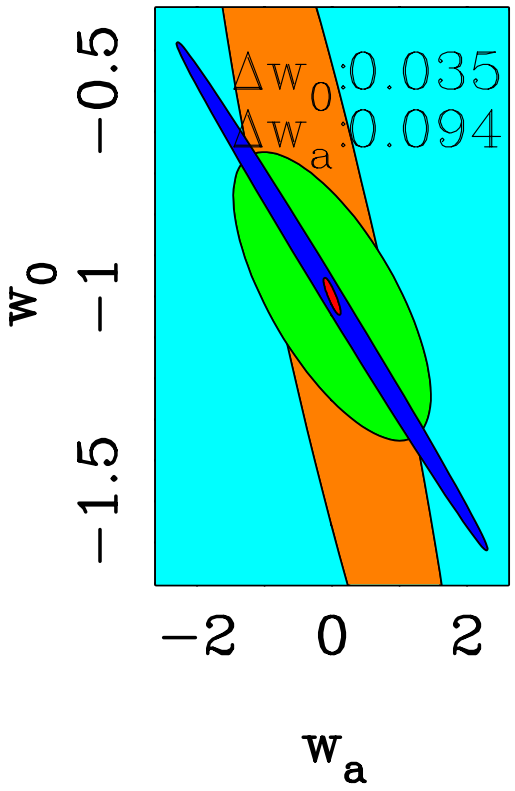

Figure 1.4: Predicted 1σ Gaussianised constraints on the dark energy equation of state parameters w_0 and w_a from SNe Ia (light blue, fills plot), Planck CMB constraints (dark blue strip), weak lensing (green ellipse) and BAOs (orange strip), with the combined constraint from all probes shown as the small red ellipse [51]. The values of Δw_0 and Δw_a on the plot refer to the marginalised uncertainties on the dark energy parameters from the combined constraint.

of the equation of state. In the case of $a = a_0 = 1$ (present day), $w(a) = w_0$. There are alternatives to the cosmological constant which have been proposed as an explanation for the accelerating expansion of the Universe. These models include alternative gravity models (which explains accelerated expansion through a breakdown of GR on larger-than-supercluster scales) and dynamical models in which the dark energy density varies over time due to a time–dependent equation of state parameter, $w(t)$ (see e.g. [53, 54]). Such behaviour is not ruled out by observations, especially when considering that very little is known about how w varies with redshift (and hence time). For a comprehensive, recent review see e.g. [55].

Even though there are many alternative theories to the cosmological constant, the ΛCDM paradigm consistently stands up to observational tests. This model is not without its problems, however. One such difficulty is the mismatch between particle physics theory and astronomical observations; the predicted value for the vacuum energy from quantum field theory is $\sim 10^{110}$ ergcm^{-3} but the dark energy density from observations is just $\sim 10^{-10}$ ergcm^{-3} [7, 8]. This 120 order–of–magnitude discrepancy is one of the largest problems facing the ΛCDM model. Possible explanations for how the cosmological constant could be suppressed to such a small, but crucially non–zero value include anthropic arguments (if the cosmological constant were much higher, structure formation would have been impossible due to extended inflation [56]) and predictions from string theory (string theory is a framework that attempts to reconcile GR and quantum mechanics by positing that fundamental particles are composed of 1-dimensional oscillating 'strings' [57]; for an introduction see e.g. [58]). A further reason for potentially doubting the validity of the concordance Universe is known as the **coincidence problem**. The observed values of Ω_Λ and Ω_m are approximately equal in the present era, yet we know that at some point in the past the matter density would have dominated over the dark energy density by many orders of magnitude, and in the future dark energy will dominate the energy density by increasing orders of magnitude. The transition to dark energy domination has occurred relatively recently in our Universe's history (so recent, in fact, that it is possible to observe SNe Ia that are distant enough that they appear to still be decelerating in their recession [59]). This has led many to ask the question: why are we observers at this special, transitional epoch? Is it random chance

that allows us to exist at a rare time in the Universe's evolution when the detection of dark energy is possible, or is there some underlying mechanism or even new physics that leads to the current measurement of the energy density? Many hypotheses have been put forward, including anthropic arguments [60, 61] and alternative dark energy models like those mentioned above, but as yet a fundamental physical explanation for dark energy eludes us.

1.3.2 Dark Matter

The final key component in the ΛCDM model is **cold dark matter** (CDM). From the density parameter we know that around 30% of the mass–energy content of the Universe is in the form of matter. It is now known, however, that the majority of that matter is not baryonic[12], but is an as–yet unknown type of non–baryonic matter termed dark matter. As the name suggests, dark matter does not emit radiation and its existence can as yet only be inferred from its gravitational effect on baryons.

Evidence for Dark Matter

The first evidence for dark matter was discovered by Fritz Zwicky in 1933 during a study of the Coma cluster [63] . Using the Virial theorem, it was shown that the cluster had insufficient mass to hold itself together unless some invisible matter was present in the cluster to contribute to its mass. More recently, observations of a cluster that has come to be known as the Bullet cluster [64] combined X–ray measurements of hot, intra-cluster gas (which acts as a good tracer of the baryonic mass) and **gravitational lensing** measurements (which trace the total cluster mass). The Bullet cluster actually consists of two galaxy clusters that have just collided, and a measured offset between the X-ray gas and the total cluster masses as a result of the collision indicates that there is a large component of non-baryonic matter in the system. Additionally, gravitational lensing is directly sensitive to the amount and distribution of the dark matter (see Chapter 2), and lensing measurements of galaxies around galaxy clusters show that the lensing effect is too strong to be caused by the visible baryonic matter alone [65, 66].

[12] A baryon is a massive particle made up of three fundamental particles known as quarks; protons and neutrons are baryons, and they make up most of the mass of visible matter in the Universe [62].

Evidence for dark matter associated with individual galaxies also exists and comes primarily from the measured **rotation curve** of a galaxy, which profiles the rotational velocity of visible matter as a function of radial distance from the centre of the galaxy. From Newtonian physics (which is valid because the gravitational field is weak), the expected shape of the rotation curve is

$$v(r) = \sqrt{\frac{GM(< r)}{r}}, \tag{1.71}$$

where $v(r)$ is the velocity of an object orbiting at radius r and $M(< r)$ is the mass contained within that orbit. At large r, most of the mass of a galaxy will be contained within the orbit and the velocity of stars in the region should decrease as $v \propto r^{-1/2}$. Observations of globular clusters in galaxies show that $v \sim$ constant with radius, suggesting that a dense outer halo of dark matter exists that extends far beyond the visible edge of the galaxy. Galactic dark matter halos are typically calculated to be around five to ten times the mass of the visible galaxy [7, 9].

Nucleosynthesis in the hot, early Universe restricts the baryon fraction of the Universe to $\Omega_b h^2 = 0.0224 \pm 0.0009$ [67]. Combined with measurements from large scale structure that constrain $\Omega_m \sim 0.27$, (see section 1.3.1), this provides further evidence that the majority of matter in the Universe is non-baryonic.

BAOs also imply the existence of dark matter. The angular power spectrum of the CMB contains information on the modes of these oscillations via their sound speeds. By measuring the ratio of the amplitude of the peaks from different modes it is possible to measure the baryon fraction [20, 68]. Currently, such measurements have yielded a value for the baryon density parameter of $\Omega_b \approx 0.04$. The significance of this result is that only $\sim 15\%$ ($\sim$0.04/0.27) of the matter in our Universe is baryonic; the nature of the other 85% is still very much unknown.

The Identity of Dark Matter

Three main types of non–baryonic dark matter have been proposed: hot dark matter (consisting of particles that decoupled from radiation when ultrarelativistic), warm dark matter (decoupled when relativistic) and cold dark matter (decoupled when non-relativistic). The standard model is based on cold dark matter because its existence seems necessary to explain the formation of structure in the Universe, as a hot dark matter Universe does not fit observations well. A warm dark matter

Universe could be plausible and is an area worthy of future investigation [7].

A number of different dark matter candidates have been proposed. Dark matter is believed to be a fluid that interacts with baryonic matter through gravity only, and with itself only weakly. The leading candidates are known as **Weakly Interacting Massive Particles**, or WIMPS. WIMPS have a large mass, and hence are slow moving and 'cold'. They interact with each other through the weak nuclear force, and with all matter through gravity. The identity of such a particle is as yet unknown, but particle physics suggests it may be a stable supersymmetric particle such as a neutralino. Particles accelerators such as the Large Hadron Collider may detect WIMPs indirectly if they are produced in high-energy collisions [69]; direct detection may also be possible using deep underground cryogenic detectors by looking for dark matter particles scattering off atomic nuclei (distinguished from background particles which scatter off electrons) [70, 71].

1.4 Summary

The hot big bang model and ΛCDM paradigm have been highly successful at describing our Universe. Over the course of the twentieth century, a cosmological picture has emerged that describes the evolution of the Universe and its major constituents, and makes accurate, testable predictions. Multiple cosmological probes have contributed complementary evidence towards validating the standard model, however there are still components of the model - notably the identity of dark matter and dark energy - that are not well understood. Solving these mysteries is the driving force behind much of cosmological research today. Weak gravitational lensing is one probe that cosmologists are using to constrain the identities of dark matter and dark energy. Over the past decade, weak lensing has begun to develop into a competitive tool for precision cosmology, complementing observations made with other cosmological probes such as BAOs, SNe and the CMB.

Precise and accurate measurements of cosmological parameters are an essential tool for discriminating between different cosmological models and unveiling the identities of dark matter and dark energy. As we shall see in the following chapters, weak lensing measurements are coming to the forefront of cosmology

as a tool for precisely constraining these parameters. In order to make sure the next generation of lensing telescopes are able to use their survey data as effectively as possible and place tight constraints on cosmological parameters, we must investigate ways to efficiently and optimally use the available data, as well as ensure that such data is as free from systematic sources of error as possible. The work in this thesis was therefore conducted with the intention of minimising the uncertainty on estimates of key cosmological parameters measured from weak lensing data, so that we might move one small step closer to a well-understood, comprehensive model for our Universe. This thesis was conducted using weak lensing data and simulations, and with this in mind, gravitational lens theory and the principles of observing weak lensing will be discussed in the next chapter.

Chapter 2

Cosmology with Weak Gravitational Lensing

Photons propagating through an inhomogeneous density field will be deflected such that their path is curved as seen by an observer. This is the phenomenon known as gravitational lensing. It occurs because light rays follow null geodesics in spacetime, and if spacetime is curved (non-Euclidean) then these geodesics are no longer straight. The elegance and utility of lensing lies in the fact that it is underpinned by the well-understood physics of general relativity, and observable lensing quantities are easily relatable to cosmological parameters. In this chapter, we introduce the concept of gravitational lensing and review the basic lensing formalism with a focus on the weak lensing regime which is the topic of this thesis. We also discuss some useful weak lensing statistics such as the shear correlation function, and introduce the concept of tomography, which is central to this thesis. Finally, we review some of the common sources of systematic error in weak lensing data and review the status of some prominent current and upcoming weak lensing surveys.

2.1 What is Lensing?

The deflection of light in a gravitational field is predicted by both Newtonian dynamics and GR, and was used to test GR successfully in 1919 (GR predicts twice the Newtonian deflection angle [7]). The theory of gravitational lensing is eloquently described by GR and in the past three decades lensing has emerged

as a powerful probe of cosmology. Part of the reason that gravitational lensing is so useful to astronomers is because the deflection of light has no dependence on the underlying nature of the lensing matter. This means that lensing can be used to constrain cosmological parameters whilst making very few assumptions about the nature of astrophysical objects, leading in principle to fewer sources of uncertainty in our analysis.

The easiest way to understand why 'mass bends light' is to consider the *falling lift* thought experiment. This experiment utilises the **strong equivalence principle**[1] and its application to inertial frames of reference. An inertial frame is one in which Newton's laws of motion are valid; particles which are not acted on by any net external force move in straight lines with constant velocity. Consider a lift in an inertial reference frame, S. If a person inside the lift shines a torch horizontally, an observer in that frame would expect to see the beam of light shine horizontally across the lift and reach a point at the same height on the opposite wall (see the left hand side of Fig. 2.1). Since the frame we chose is inertial, then an external observer in *any* inertial reference frame, S', will witness the light beam travelling horizontally across the lift. Now consider repeating the experiment in a lift that is in free fall under a gravitational acceleration $\boldsymbol{g}$. The equivalence principle tells us that a freely–falling frame is locally equivalent to an inertial frame so an observer inside the lift would again see the light beam tracing a straight, horizontal path. However an external observer in a different frame will now see the light take a curved path relative to the lift (right hand side of Fig. 2.1). The universality of free fall tells us that the lift could be falling in a gravitational field or being accelerated by some other mechanism in the absence of a gravitational field. An observer outside the freely–falling frame will thus observe light rays being deflected as they pass through a gravitational field. This is the effect known as gravitational lensing.

2.1.1 Lensing Regimes

Here we review the three different classes of gravitational lens, the last of which forms the basis of this report.

[1]The strong equivalence principle applies to all the laws of physics, and states that: *In a locally inertial frame, all of special relativity applies.* This can also be expressed by saying that *the laws of physics, and the outcome of any local experiment, are the same for all locally inertial, freely–falling frames.*

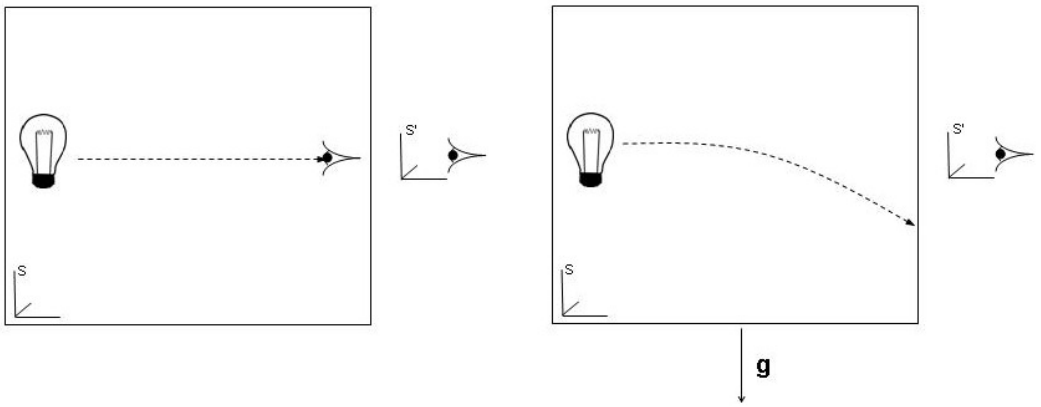

Figure 2.1: *Left*: A light ray in a lift in an inertial frame will be seen to travel along a straight line by observers in all inertial frames. *Right*: A light ray in a lift that is in free–fall will appear to take a curved path as seen by observers in inertial frames.

Microlensing

Microlensing events occur when a dark, massive object such as a brown dwarf passes in front of a background star. If the alignment between observer, lens and background star is precise, the star will be significantly magnified due to the lensing effect (see §2.2.5). The magnification of the background object is detectable with high-resolution telescopes such as the Hubble Space Telescope (HST) by the characteristic high–kurtosis, sharp peak in its brightness [72]. Microlensing events are useful for exploring the non-luminous baryonic matter in the galaxy since the mass of the lens can be determined from lensing theory in the absence of any light from the lens itself. Microlensing has notably been used to detect extrasolar planets: if the lens is a star with a planet in orbit, the extra mass of the planet will produce a second smaller peak in the brightness profile of the lensed star. This has led to the discovery of extrasolar planets with masses as low as 1.7 times the mass of the Earth [73, 74]. Other objects that are capable of inducing detectable microlensing events include white dwarfs, neutron stars and black holes which in this context are sometimes known as MAssive Compact Halo Objects (MACHOs). MACHOs were once proposed as a baryonic

solution to the identity of dark matter, but microlensing observations by EROS[2] have shown that MACHOs can account for at most $\sim 10\%$ of the missing galactic dark matter [75]. This number is further constrained by the peaks in the power spectrum of the CMB that limit $\Omega_b \sim 0.04$ (see §1.3.2).

Strong Lensing

Strong gravitational lensing occurs when massive objects such as clusters of galaxies act as gravitational lenses. It is this lensing regime that produces the most extreme distortions to astronomical images. An example of a galaxy cluster strongly lensing background galaxies is shown in Fig. 2.2. The cluster lens produces multiple sheared and magnified images of the background galaxies. The number of multiple images can be used to estimate the volume of space between the observer and the cluster, which is dependent on cosmological parameters and in particular the cosmological constant. If a strongly lensed source varies with time, then its image will vary with time as well. However, due to the curvature of spacetime in the presence of mass the light rays that form the multiple images will have different path lengths. This leads to time delays between the observed images so that, provided the angles in the lens system are known (as they often are), the exact distances between images and objects are known thus allowing an estimation of h_0 [76, 77]. The arcs in Fig. 2.2 can only be produced by massive objects such as galaxy clusters, so the number, size and geometry of these arcs can yield information on the massive lens. It is also possible to reconstruct the mass distribution of the lens in order to improve constraints on the matter density [78]. Strong lensing clusters can yield valuable cosmological information, however detection of strong lensing events is rare and relies on chance alignments of galaxies and massive clusters, thus limiting its use as a cosmological probe.

Weak Lensing

The final type of gravitational lensing is distinct from the other two as it does not rely on the chance alignment of astrophysical objects to be detected and is essentially ubiquitous. Weak lensing occurs when the lensing mass is much lower than in the strong lensing regime so the resulting image distortions are much

[2] `http://eros.in2p3.fr`.

Figure 2.2: Strong lensing by Abell2218 cluster produces arcs and multiple galaxy images (NASA/STScI).

smaller and multiple galaxy images and visible arcs are not produced. Weak lensing arises because any dark matter along the line of sight to a galaxy will produce a small lensing effect. Since there will be some dark matter between us and every other galaxy, weak lensing can be observed along any line of sight to a galaxy. Individual light rays experience a series of tiny deflections as they pass through the varying gravitational fields due to the dark matter distribution, which sum to give a net distortion (this is the Born approximation, see §2.2.1).

The distortion of galaxy images due to weak lensing includes a change in their observed ellipticities. This effect is known as **cosmic shear** and galaxy images are typically sheared by only $\sim$1%. This makes such distortions impossible to spot by eye and detailed statistical treatment of many galaxy images is needed to accurately measure this effect. Fluctuations in the matter density in different regions of space ensure that the weak lensing signal will vary over different parts of the sky, so measuring cosmic shear can yield direct information on the matter density, Ω_m, and the matter power spectrum for which the amplitude at $z = 0$ can be parameterised by σ_8. Weak lensing is sensitive to the combination of Ω_m and σ_8, leading to a characteristic degeneracy in the measurement of these two parameters. The degeneracy can be partially broken by measuring the lensing signal at different redshifts (for example, by using a technique known

as tomography; this will be discussed further in §2.3.7 and Chapters 4 & 5) [79] or by comparing the lensing signal in the linear and non-linear regimes [80]. Combined with CMB measurements, strong constraints on Ω_m and σ_8 can be achieved since the degeneracies in the parameters from each measurement are almost orthogonal in the Ω_m–σ_8 plane [81, 82]. The measurement of cosmic shear is well-established and has been used to measure the bias parameter b [83, 84] and the 3D non–linear matter power spectrum $P_\delta(k)$ [85], and can be used to probe dark energy [82, 86].

2.2 Gravitational Lens Theory

A full treatment of gravitational lensing can be described in the context of GR in the space-time of the Robertson-Walker metric. However most astrophysical situations permit a simpler, approximate treatment, called gravitational lens theory, which we describe here.

2.2.1 The Lens Equation

A typical gravitational lensing situation is sketched in Fig. 2.3, where a mass at distance D_d deflects light rays from a source at distance D_s. The distance between the lens plane and the source plane is given by D_{ds} and all distances are angular diameter distances. In the absence of lensing, we would expect the image to be located at $\boldsymbol{\beta}$, but the deflection of angle $\hat{\boldsymbol{\alpha}}$ caused by the mass at D_d produces an image at $\boldsymbol{\theta}$. Since all the angles involved in typical lensing situations are small, we can use the small angle approximation throughout. This allows us to approximate the path of the photons as straight lines with a kink in the source plane, as their true path will be smoothly curved by the lens[3]. From Fig. 2.3 we can then write

$$\boldsymbol{\theta} D_s = \boldsymbol{\beta} D_s + \hat{\boldsymbol{\alpha}} D_{ds}. \tag{2.1}$$

[3]In general in lensing theory we assume that the deflected light ray can be approximated to a straight line in the neighbourhood of the deflecting mass. This is known as the Born approximation and is valid as long as the deviation of the light ray from a true straight line is small compared to the scale over which the mass distribution changes. The Born approximation has been shown to be relatively robust in ray-tracing simulations for shear measurement purposes [85, 87].

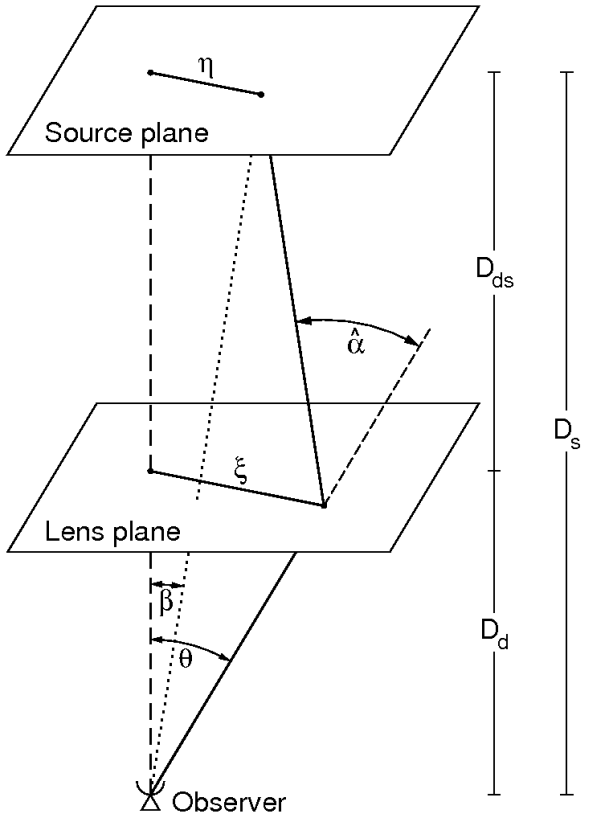

Figure 2.3: Diagram showing the relevant angular diameter distances and angles between the observer, the lens plane, and the source plane in a lensing system [29].

We now define the **reduced deflection angle**, $\boldsymbol{\alpha} = \hat{\boldsymbol{\alpha}}\frac{D_{ds}}{D_s}$ so that the lens equation is given by

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \boldsymbol{\alpha}. \tag{2.2}$$

2.2.2 The Deflection Angle

The deflection angle describes the magnitude and direction of the deflection a light ray when it is weakly lensed. There are several ways to derive the deflection angle and in this section we adopt the approach of [88] and [89] to find the deflection angle of a light ray by a point mass. We begin by considering Fermat's principle: *light takes the path along which the travel time is extremal.* This can

be expressed mathematically,

$$\delta \int_A^B n\left[\boldsymbol{x}l\right] \frac{\mathrm{d}l}{c} = 0, \tag{2.3}$$

where n is the refractive index, $\boldsymbol{x}$ is the trajectory of the photon, dl is an infinitesimal displacement along the ray and the gravitational field is assumed to be weak and the lens small compared to cosmological distances. In the presence of lensing the RW metric defined in Eqn. 1.8 must be weakly perturbed by the gravitational potential, Φ, such that $|\Phi| \ll c^2$,

$$\mathrm{d}s^2 = \left(1 + \frac{2\Phi}{c^2}\right) c^2 \mathrm{d}t^2 - \left(1 - \frac{2\Phi}{c^2}\right) a^2(t) \left[\mathrm{d}r^2 + S_k^2(r)\mathrm{d}\Omega^2\right]. \tag{2.4}$$

For a photon, $\mathrm{d}s^2 = 0$. If we consider a photon travelling radially, $\mathrm{d}\Omega^2 = 0$ and we have

$$\left(1 + \frac{2\Phi}{c^2}\right) c^2 \mathrm{d}t^2 = \left(1 - \frac{2\Phi}{c^2}\right) a^2(t) \mathrm{d}r^2. \tag{2.5}$$

The proper distance is just $\mathrm{d}x = a(t)\mathrm{d}r$ which allows us to define an effective speed of light in a weak gravitational field as

$$c' = \frac{|\mathrm{d}\boldsymbol{x}|}{\mathrm{d}t} = c \left(\frac{1 + \frac{2\Phi}{c^2}}{1 - \frac{2\Phi}{c^2}}\right)^{\frac{1}{2}} \approx c \left(1 + \frac{2\Phi}{c^2}\right). \tag{2.6}$$

It can be seen from Eqn. 2.6 that the index of refraction in the gravitational field will be

$$n \equiv \frac{c}{c'} \approx \left(1 - \frac{2\Phi}{c^2}\right). \tag{2.7}$$

Since $\Phi \leq 0$, n will be greater than 1 and $c' \leq c$, meaning that *light appears to travel more slowly in a gravitational field*[4], analogously to light being slowed by a medium with refractive index n.

Fermat's principle can be used to find the angle of deflection between an unperturbed ray and a ray perturbed by a gravitational field. If we consider the

[4]Note that light only *appears* to slow down in a gravitational field and actually always travels locally at c, as otherwise both special and general relativity would be violated. Non-relativistically, Fermat's principle is a mathematical shortcut to the observed deflection angle, although it can also be stated relativistically (see [90]). A relativistic derivation of the deflection angle can be found in e.g. [91].

unperturbed path distance, dl, then we can define a curve parameter, λ such that

$$\mathrm{d}l = \left|\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}\lambda}\right| \mathrm{d}\lambda = |\dot{\boldsymbol{x}}|\,\mathrm{d}\lambda. \tag{2.8}$$

Using the Lagrangian for the system, $\mathcal{L}(\boldsymbol{x}, \dot{\boldsymbol{x}}) \equiv \mathrm{n}\left[\boldsymbol{x}(\lambda)\right]|\dot{\boldsymbol{x}}| = \mathrm{n}\left[\boldsymbol{x}(\lambda)\right](\dot{\boldsymbol{x}}^2)^{\frac{1}{2}}$, Fermat's principle can be expressed using the Euler Lagrange equation,

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\frac{\partial\mathcal{L}}{\partial\dot{\boldsymbol{x}}} = \frac{\partial\mathcal{L}}{\partial\boldsymbol{x}}, \tag{2.9}$$

which yields

$$n\dot{\boldsymbol{e}} = \nabla n - \boldsymbol{e}.(\nabla.\boldsymbol{e}), \tag{2.10}$$

with $\boldsymbol{e} \equiv \dot{\boldsymbol{x}}$, the tangent vector to the light ray. Eqn. 2.10 is just the perpendicular gradient of n, meaning

$$\dot{\boldsymbol{e}} = \frac{1}{n}\nabla_\perp n = \nabla_\perp \ln n = \nabla_\perp \ln\left(1 - \frac{2\Phi}{c^2}\right) \approx -\frac{2}{c^2}\nabla_\perp\Phi \tag{2.11}$$

and the **total deflection angle** is given by the integral of $-\dot{\boldsymbol{e}}$ along the light path, that is

$$\hat{\boldsymbol{\alpha}} = \frac{2}{c^2}\int_A^B \nabla_\perp\Phi\,\mathrm{d}\lambda. \tag{2.12}$$

In weak gravitational fields, the deflection angle is very small, allowing us to use the Born approximation and evaluate $\hat{\boldsymbol{\alpha}}$ along the *unperturbed* (straight) path.

The result of Eqn. 2.12 and the Born approximation allow us to find an expression for $\hat{\boldsymbol{\alpha}}$ for deflection by a point mass. Fig. 2.4 illustrates such a situation. The impact parameter, b, is the point of closest approach of the unperturbed ray to the point mass M. The majority of the deflection occurs within $\triangle z \approx \pm b$ of the closest approach. Since the field is weak we can use the Newtonian approximation for the gravitational potential of the lens,

$$\Phi = -\frac{GM}{r} = -\frac{GM}{(x^2+y^2+z^2)^{\frac{1}{2}}} = -\frac{GM}{(b^2+z^2)^{\frac{1}{2}}}. \tag{2.13}$$

Thus we can calculate an expression for $\nabla_\perp\Phi$,

$$\nabla_\perp\Phi = \frac{GM\boldsymbol{b}}{(b^2+z^2)^{\frac{3}{2}}}, \tag{2.14}$$

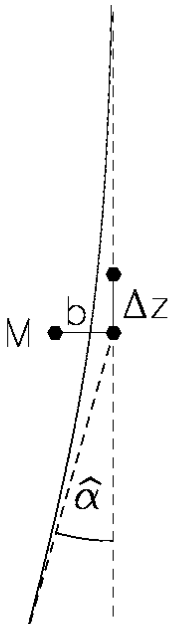

Figure 2.4: The path of a light ray (solid black line) near point mass M, deflected by angle $\hat{\boldsymbol{\alpha}}$. The impact parameter b is the point of closest approach of an unperturbed ray (dashed grey line) [89].

and substitute this into Eqn. 2.12 to find the deflection angle for a point mass, using the Born approximation such that $\mathrm{d}z \approx \mathrm{d}l$:

$$\hat{\boldsymbol{\alpha}} = \frac{2}{c^2} \int \nabla_{\perp} \Phi \, \mathrm{d}z = \frac{4GM}{c^2 b} = \frac{2}{b} R_s \tag{2.15}$$

where R_s is the Schwarzschild radius for a point mass and $b \gg R_s$. Note that the result derived here is exactly twice the deflection expected from Newtonian physics.

2.2.3 Surface Density and the Thin Screen Approximation

A real lens will not be a point mass, but will have some finite mass distribution in space. Since the gravitational field is weak, we follow the approach of [89] and approximate the total deflection produced by a mass distribution as the sum of the deflection angles produced by a series of point masses. We assign the two–dimensional position vector of the light ray, $(\boldsymbol{\xi}, z)$. If each point mass occupies a

small volume element $\mathrm{d}V$ with mass $\mathrm{d}m = \rho(\boldsymbol{r})\mathrm{d}V$ (where ρ is the mass density), then we can assign a position vector $(\boldsymbol{\xi'}, z')$ to each point mass. At a mass element, the light ray has impact parameter $\boldsymbol{\xi} - \boldsymbol{\xi'}$ and the total deflection angle is given by the sum of each deflection,

$$\begin{aligned} \hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) &= \frac{4G}{c^2} \sum \mathrm{d}m(\boldsymbol{\xi'}, z') \, \frac{\boldsymbol{\xi} - \boldsymbol{\xi'}}{|\boldsymbol{\xi} - \boldsymbol{\xi'}|^2} \\ &= \frac{4G}{c^2} \int \mathrm{d}^2\xi' \int \mathrm{d}z' \rho(\boldsymbol{\xi'}, z') \, \frac{\boldsymbol{\xi} - \boldsymbol{\xi'}}{|\boldsymbol{\xi} - \boldsymbol{\xi'}|^2}. \end{aligned} \tag{2.16}$$

Since most of the light deflection occurs within $\triangle z \approx \pm b$ of closest approach and $\triangle z \ll D_d, D_{ds}$ this allows us to use the **thin screen approximation**: the lens is thin compared to the path length of the light ray. This approximation allows us to project the mass distribution of the lens as a thin mass sheet that lies perpendicular to the line of sight. The **surface density** of such a mass sheet is given as

$$\Sigma(\boldsymbol{\xi}) = \int_{-\infty}^{+\infty} \rho(\boldsymbol{\xi}, z) \, \mathrm{d}z. \tag{2.17}$$

This allows us to write the deflection angle as a two–dimensional vector in terms of the surface density,

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \frac{4G}{c^2} \int \mathrm{d}^2\xi' \, \Sigma(\boldsymbol{\xi'}) \, \frac{\boldsymbol{\xi} - \boldsymbol{\xi'}}{|\boldsymbol{\xi} - \boldsymbol{\xi'}|^2} \, . \tag{2.18}$$

2.2.4 Convergence and the Lensing Potential

From Fig 2.3 we know that $\boldsymbol{\xi} = D_d \boldsymbol{\theta}$ and combining this with Eqn. 2.18 it is now possible to write down a new expression for the reduced deflection angle,

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{4G}{c^2} \frac{D_{ds} D_d}{D_s} \int \mathrm{d}^2\theta \, \Sigma(D_d(\boldsymbol{\theta})) \, \frac{\boldsymbol{\theta} - \boldsymbol{\theta'}}{|\boldsymbol{\theta} - \boldsymbol{\theta'}|^2}. \tag{2.19}$$

Defining a critical surface mass density, Σ_{cr} as

$$\Sigma_{cr} = \frac{c^2}{4\pi G} \frac{D_s}{D_{ds} D_d}, \tag{2.20}$$

we can rewrite the reduced deflection angle as

$$\alpha(\theta) = \frac{\Sigma}{\Sigma_{cr}}\,\theta. \tag{2.21}$$

For $\Sigma = \Sigma_{cr}$, $\alpha = \theta = \theta_E$ and we see an Einstein ring if the lens is circularly symmetric. An Einstein ring is formed when an image is deformed into a ring around the lensing mass, with Einstein radius $\theta_E = \sqrt{\frac{4GM}{c^2}\frac{D_{ds}}{D_d Ds}}$.

For a general lens, we define the **convergence** or dimensionless surface mass density κ,

$$\kappa(\boldsymbol{\theta}) = \frac{\Sigma(D_d\boldsymbol{\theta})}{\Sigma_{cr}}. \tag{2.22}$$

The convergence distinguishes between the strong and weak lensing regimes; if $\kappa \geq 1$ then multiple images will be produced for some source positions as expected in cases of strong gravitational lensing. A value of $\kappa \ll 1$ corresponds to the weak lensing regime. It is now possible to relate the reduced deflection angle and the convergence [89],

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{1}{\pi}\int \mathrm{d}^2\theta'\,\kappa(\boldsymbol{\theta'})\,\frac{\boldsymbol{\theta}-\boldsymbol{\theta'}}{|\boldsymbol{\theta}-\boldsymbol{\theta'}|^2}\,. \tag{2.23}$$

We can define the reduced deflection angle as the gradient of a scalar potential, the **effective lensing potential** ψ, which is the two–dimensional analogue of the Newtonian potential Φ of the lens,

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \nabla_\theta\psi(\boldsymbol{\theta}). \tag{2.24}$$

then using the identity $\nabla \ln \boldsymbol{x} = \boldsymbol{x}/|\boldsymbol{x}|^2$ we can write the lensing potential in terms of the convergence,

$$\psi(\boldsymbol{\theta}) = \frac{1}{\pi}\int \mathrm{d}^2\theta'\kappa(\boldsymbol{\theta'})\ln|\boldsymbol{\theta}-\boldsymbol{\theta'}|. \tag{2.25}$$

The inverse two–dimensional Laplacian operator $\delta^{-2} = \int \mathrm{d}^2\theta' \ln|\boldsymbol{\theta}-\boldsymbol{\theta'}|$ gives

$$\boldsymbol{\kappa}(\boldsymbol{\theta}) = \frac{1}{2}\nabla^2_\theta\psi(\boldsymbol{\theta}) \tag{2.26}$$

$$= \frac{1}{2}(\psi_{,11}+\psi_{,22}), \tag{2.27}$$

therefore the convergence is related to the reduced deflection angle by [29]

$$\boldsymbol{\kappa}(\boldsymbol{\theta}) = \frac{1}{2}\nabla_\theta \,.\, \alpha(\boldsymbol{\theta}). \tag{2.28}$$

2.2.5 Magnification and Shear

The lens equation yields the angular positions of the images of a source at $\boldsymbol{\beta}$. The image shapes will be distorted because light bundles are deflected differentially; that is, light bundles emerging from slightly different locations will pass through slightly different potentials. To determine the shape of the image in general one must solve the lens equation for all points within an extended source.

If the angular scale on which the lens properties change is much larger than the angular size of the source, then the local imaging properties of the lens mapping can be described by the lensing Jacobian matrix,

$$\begin{aligned} \mathcal{A} &= \frac{\partial \boldsymbol{\beta}}{\partial \boldsymbol{\theta}} = \frac{\partial}{\partial \theta_j}[\theta_i - \alpha_i(\boldsymbol{\theta})] \\ &= \delta_{ij} - \frac{\partial \alpha_i(\boldsymbol{\theta})}{\partial \theta_j} \\ &= \delta_{ij} - \frac{\partial^2 \psi(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} \end{aligned} \tag{2.29}$$

where δ_{ij} is the kronecker delta. The Jacobian is symmetric and can also be expressed as

$$\mathcal{A}_{ij} = \begin{bmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{bmatrix}. \tag{2.30}$$

Here we have introduced the concept of the **shear**, γ, a measure of the shape distortion of the image. The shear is complex and is given by $\gamma \equiv \gamma_1 + i\gamma_2 = e^{2i\phi}$, where

$$\begin{aligned} \gamma_1 = \frac{1}{2}(\phi_{,11} - \phi_{,22}) &= \gamma\cos(2\phi) \\ \gamma_2 = \phi_{,12} = \phi_{,21} &= \gamma\sin(2\phi). \end{aligned} \tag{2.31}$$

Liouville's theorem combined with the lack of emission or absorption of photons in gravitational light deflection tells us that *lensing conserves surface brightness*. Thus, if gravitational lensing increases the surface area of an image

we will see a **magnification** μ where

$$\mu = \frac{\text{image area}}{\text{source area}} = \left| \frac{\partial \theta_i}{\partial \beta_j} \right|. \tag{2.32}$$

If photons are neither created nor destroyed by gravitational lensing, the net magnitude can increase if photons on previously nearby trajectories are bent towards us by the lensing mass (since the object is dilated while the surface brightness stays constant). Gravitational lensing allows us to observe photons that would have not otherwise reached us in the absence of the lens.

The magnification is related to the Jacobian and hence to the convergence and shear through the determinant,

$$\mu \equiv \frac{1}{\det \mathcal{A}} = \frac{1}{(1-\kappa)^2 - \gamma^2}. \tag{2.33}$$

This tells us that gravitational lensing distorts images in both shape and size. The shape distortion described by γ is due to the presence of a tidal gravitational field, whereas it is a combination of κ and γ that lead to magnification through isotropic and anisotropic focusing by the local matter density respectively. The magnification of images means that in any flux-limited survey, magnification from the effect of weak lensing will increase the number density of galaxy images.

2.3 Observing Weak Lensing

We have derived relationships between the convergence, κ, the lensing potential, ψ, and the shear, γ. In this section, we will examine the relationship between these quantities and those observable from weak lensing measurements, namely the shapes of galaxies. We will discuss some methods to measure galaxy shapes, and introduce some of the important sources of systematic error in lensing analysis that shape measurement techniques must account for.

2.3.1 Ellipticity

In order to measure the shape of a galaxy, we must first define its centre, $I_0 = I(\boldsymbol{\theta} = 0)$, such that it is centred on the origin. Lensing images are altered by a change in size which is characterised by the monopole moment of the galaxy,

and a translation on the sky that is characterised by the dipole moment. In weak lensing the easiest-to-observe quantity is the ellipticity, which may be used to estimate the shear[5], which is related to the quadrupole moment of a galaxy. The quadrupole moment will be

$$Q_{ij} = \frac{1}{A} \int \mathrm{d}^2\theta \, I\left(\boldsymbol{\theta}\right) \Delta\theta_i \Delta\theta_j \tag{2.34}$$

where A is the area, $I\left(\boldsymbol{\theta}\right)$ is the surface brightness of the galaxy and $\Delta\theta$ is the angular separation between two points in the image. From this, the complex ellipticity may be defined as $\epsilon = \epsilon_1 + i\epsilon_2 = |\epsilon| e^{2i\phi}$, where ϵ is can be related to Q_{ij} by

$$\epsilon = \frac{Q_{11} - Q_{22} + 2iQ_{12}}{Q_{11} + Q_{22} + 2\left(Q_{11}Q_{22} - Q_{12}^2\right)^{1/2}}. \tag{2.35}$$

For a circular image, $Q_{11} = Q_{22}$, $Q_{12} = Q_{21} = 0$ and $\epsilon_1 = \epsilon_2 = 0$. We define the reduced shear g to be

$$g(\boldsymbol{\theta}) = \frac{\gamma(\boldsymbol{\theta})}{1 - \kappa(\boldsymbol{\theta})} \tag{2.36}$$

and follow the approach of [29] to write down expressions for the intrinsic source ellipticity,

$$\begin{aligned} \epsilon^s &= \frac{\epsilon - g}{1 - g^*\epsilon} \quad \text{for} \quad |g| \leq 1 \\ &= \frac{1 - g^*\epsilon}{\epsilon^* - g^*} \quad \text{for} \quad |g| > 1. \end{aligned} \tag{2.37}$$

In the weak lensing limit, $|\gamma| \ll 1$ so $|\gamma| \approx |g|$. The inverse of Eqn. 2.37 is then

$$\epsilon \approx \epsilon^s + g. \tag{2.38}$$

Since we cannot separate the intrinsic ellipticity from the shear of individual galaxies, we must treat them statistically and average a large number of galaxy shapes. The measurement of weak lensing relies on the assumption that galaxies are intrinsically randomly orientated on the sky. If this is the case, observing

[5]A third order effect known as **flexion**, which is responsible for skewed and arced galaxy images, is beginning to prove useful. Flexion is routinely observed in strongly lensed images, and has recently been detected in weak lensing images [92, 93]. Additionally, the magnification has been detected through its effect on galaxy number density in flux-limited surveys [94, 95], and may be used in conjunction with cosmic shear to improve cosmological constraints [96].

enough galaxy images allows us to write $\langle \epsilon^s \rangle = 0$, where ϵ^s is the intrinsic ellipticity of a galaxy[6]. Hence, in the weak lensing regime,

$$\langle \epsilon \rangle = \langle g \rangle \,. \tag{2.39}$$

From this, an estimator for the shear is just

$$\gamma \approx g \approx \langle g \rangle = \langle \epsilon \rangle \,. \tag{2.40}$$

2.3.2 Measuring Galaxy Shapes

To measure cosmic shear, we must measure the shapes of galaxies through their ellipticities. Obtaining accurate shape measurements is one of the key challenges in weak lensing analysis, and we discuss in this section the methods used to do this.

The Point Spread Function

The point spread function (PSF) describes the response of the imaging system to a point source, such as a star, which causes the image to be spread out so that it is no longer a point but a finite, two–dimensional shape. For ground based telescopes, aberrations caused by atmospheric seeing are the dominant contributor to the PSF, whereas for space telescopes the PSF comes mainly from diffraction. Atmospheric seeing causes circular smearing and differences in optical path length in the imaging system result in shearing of images. Galaxy images are affected by the PSF in a similar way to stellar images, which alters their observed shapes and contaminates shape measurements. Correcting for this effect is possible by measuring the PSF using the stars in the image field and then deconvolving this from the galaxy image. Multiple stars are needed since each star gives a finite amount of information about the PSF due to noise and pixelation by the optical system. In addition the PSF often varies across the field of view. A minimal number of stars is therefore needed to characterise the PSF and there will be a typical length scale associated with this number of stars. Galaxy correlations on length scales smaller than this may be unreliable if the PSF does not vary

[6]We ignore the effect of any intrinsic galaxy alignments in this section as these complicate our measurement of the shear. Intrinsic alignments are described in §2.3.5.

smoothly, since there is not enough information available about the nature of the PSF on such small scales [97]. Correcting the PSF precisely and reliable in order to measure galaxy shapes is one of the main challenges in contemporary cosmic shear measurements (see §2.3.2).

Shape Measurement Methods

The standard method used to measure galaxy shapes in lensing is the method of Kaiser, Squires and Broadhurst (KSB) [98]. They prescribed a method for measuring galaxy shapes by inverting the effects of the PSF smearing and shearing. This was improved upon by [99] who introduced a seeing correction and [100] who introduced an astrometric distortion correction. KSB has proved useful for correcting the PSF and extracting shear measurements, however [101] found a systematic error of 1-30% in the recovery of the shear signal using this method. This is because the ellipticity parameters should ideally be inferred from the PSF–corrected image, but we instead must find it from the observed, PSF–distorted image which leads to noisy measurements. KSB has been found to be unreliable when applied to CFHTLenS data due to inaccurate assumptions the method makes about the PSF [102].

Due to the inaccuracy in recovering the lensing signal using KSB, other methods of shape measurement have been utilised. These include the method of [103] where the observed galaxy images are stacked and then fitted as PSF–convolved, sheared circular sources. This relies on the assumption that, by stacking the galaxy images, their net intrinsic ellipticity will sum to zero so that the lensing signal will act as a shear on a circular shape. This method is dependent on a well–measured PSF. [104] construct a 'finite resolution shear operator' which gives the response of an observed image to a gravitational shear before smearing by the PSF. The result is valid for any PSF and can be used to calibrate the effect of any shear estimator. [105] propose a method of first convolving images so that the resultant convolved PSF is isotropic and circular. They then decompose the convolved galaxy images into a set of orthogonal, two–dimensional Gaussian–based functions. There is no bias towards particular shapes in this method because adaptive elliptical weights are used. This technique relies on the assumed Gaussianity of the PSF, however, which if invalid could lead to systematic underestimation of cosmological parameters such as σ_8, the matter

power spectrum.

[106] and [107] present a method in which the surface brightness profiles of galaxy images are decomposed into a series of orthogonal basis functions, or '`shapelets`' known as the Gauss–Hermite series. The PSF is first modelled by decomposing stellar images into shapelets and interpolating the resultant coefficients across the image. The galaxy images can then be decomposed into shapelet coefficients, allowing one to construct a linear estimator for the shear. This technique is robust to variations in the size and shape of the PSF, is unbiased, and allows for the shape determination of even complex galaxy profiles. There is, however, a danger of overfitting noise and the reconstruction of flat galaxy profiles is sometimes inaccurate.

A series of shape measurement challenges known as the GRavitational lEnsing Accuracy Testing (GREAT) challenges are ongoing to address the need for accurate, unbiased galaxy shape measurements[7] [108]. GREAT developed out of the Shear TEsting Program (STEP) [101] which was a collaborative effort within the lensing community to develop better shape measurement techniques. GREAT invites experts in other fields such as statistical inference and computer science to develop algorithms that improve upon previous shape measurement techniques. STEP and GREAT arose because of the imminent need for more accurate and reliable shape measurements from the next generation of wide–field lensing surveys (see §2.3.6). In order for competitors to optimise their shape measurement methods, simulations are provided with known shear values to test their methods and calibrate their results. The algorithms can then be applied to 'blind' simulations where the shear in unknown. Each method is then assigned a quality factor based on the error in the results. The method that best recovers the shear from the 'blind' simulations will have the highest quality factor. These challenges have resulted in the emergence of several new shape measurement methods, which, once developed, may further be exploited in the future for weak lensing analysis [101, 108, 109].

[7] `http://www.greatchallenges.info/`.

LENSFIT

LENSFIT[8] is a model–fitting approach to weak lensing shape measurement that uses a Bayesian shape estimation method [110, 111]. Bayesian statistics are chosen over likelihood functions as the input shear is recovered more successfully. This technique consists of fitting model surface brightness profiles to individual galaxies, which typically have six free parameters: central surface brightness, size, ellipticity (e_1 and e_2) and celestial position. One can marginalise over the parameters that do not influence the shear measurements in Fourier space (for isolated galaxies only the ellipticity parameters and the scale length of the galaxy are needed) and then perform a least-squares fit in Fourier space to determine the best–fit shape of a galaxy. The algorithm incorporates an estimate of the PSF into the model fits. LENSFIT has the advantage that it optimises the signal to noise ratio and allows almost unbiased shape estimation. There is some bias in the posterior probability distribution, however, due to the choice of prior; the prior must contain zero shear and is ideally a perfect representation of the underlying intrinsic ellipticity distribution.

The main drawback of the LENSFIT algorithm is that despite utilising fast Fourier transforms to speed up computation times, it takes around one second per galaxy to find the full posterior distribution. This could make LENSFIT unsuitably slow for very large surveys with $\sim 10^8$–10^9 galaxies. Close pairs of galaxies are not treated by the algorithm, but a separate fitting algorithm could be applied to such galaxies. However, since galaxies that are close together on the sky are often sources of intrinsic alignments (see §2.3.5), they are often downweighted and therefore measuring the shapes of isolated galaxies may be more important in lensing analysis. LENSFIT has been tested with STEP and GREAT (and has one of the highest quality factors of all the algorithms submitted) and has been applied successfully to weak lensing measurements of X–ray underluminous galaxy clusters [112]. LENSFIT is the shape measurement tool of choice for CFHTLenS (see section 2.3.6) [113] and shows promise for application to future wide-field lensing surveys.

The complicated issue of measuring accurate galaxy shapes is summarised in Fig. 2.5 (image from the GREAT08 handbook [109]), which shows the series of distortions a typical galaxy image undergoes.

[8] http://www.physics.ox.ac.uk/lensfit/.

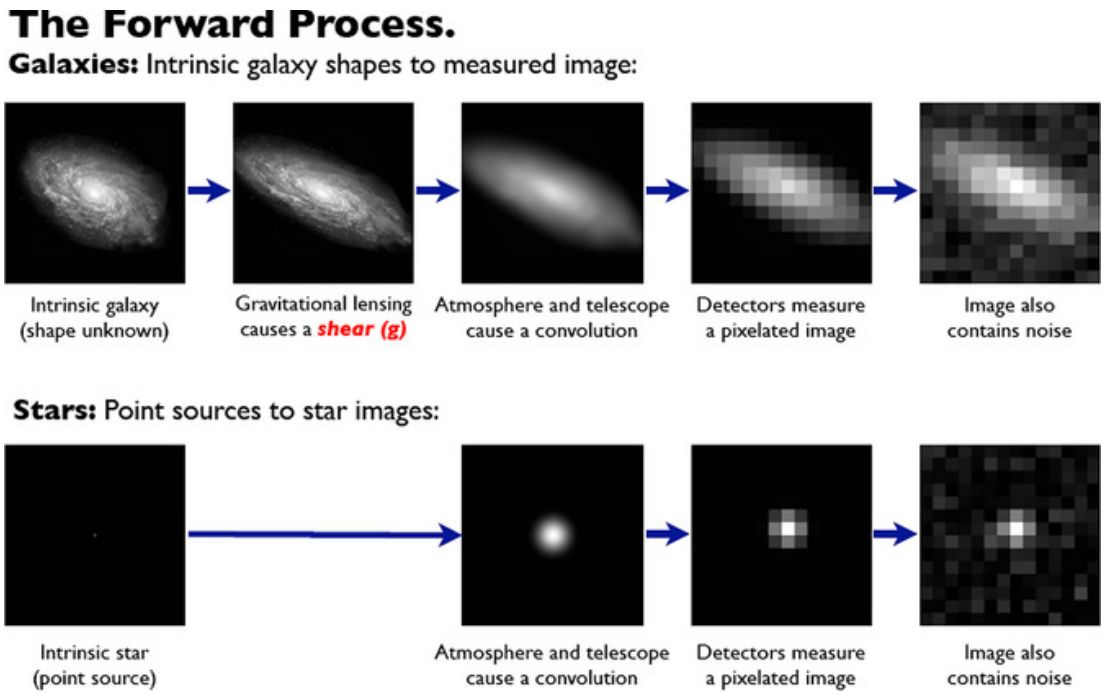

Figure 2.5: *Top*: the sequence of distortions a galaxy image is subjected to. The galaxy image is sheared by the density field it passes through, smeared and further sheared by the atmosphere and telescope optics (this is the PSF) and then pixelised. Finally, a shot noise component due to Poisson noise in the number of photons received and Gaussian noise from detector effects will also be present. *Bottom*: Stellar sources undergo a similar procedure, although they are not sheared by the density field [109].

2.3.3 The Correlation Function and Power Spectrum

We now define two useful observable statistics in weak lensing: the two-point **correlation function** $C(r)$ and the **power spectrum** $P(k)$. Assuming a flat sky[9], the correlation function can be defined for a homogeneous, isotropic and random field $g(\boldsymbol{x})$, such as the matter density field, and is simply [7]

$$C(r) = \langle g(\boldsymbol{x}) g^*(\boldsymbol{x}') \rangle \tag{2.41}$$

where r is the separation between two points in the field such that $r = |\boldsymbol{x} - \boldsymbol{x}'|$. The power spectrum is the Fourier transform of the correlation function and we

[9]The flat sky approximation avoids the need for spherical harmonics, and works well for high wavenumber k.

can derive it by first considering the Fourier transform pair of g in n dimensions,

$$\begin{aligned} \hat{g}(\boldsymbol{k}) &= \int \mathrm{d}^n x\, g(\boldsymbol{x}) e^{i\boldsymbol{x}.\boldsymbol{k}} \\ g(\boldsymbol{x}) &= \frac{1}{(2\pi)^n} \int \mathrm{d}^n k\, \hat{g}(\boldsymbol{k}) e^{-i\boldsymbol{x}.\boldsymbol{k}}. \end{aligned} \tag{2.42}$$

Therefore the correlation function in Fourier space is

$$\langle \hat{g}(\boldsymbol{k}) \hat{g}^*(\boldsymbol{k}') \rangle = \int \mathrm{d}^n x\, e^{i\boldsymbol{x}.\boldsymbol{k}} \int \mathrm{d}^n x'\, e^{-i\boldsymbol{x}'.\boldsymbol{k}'} \langle g(\boldsymbol{x}) g^*(\boldsymbol{x}') \rangle \,. \tag{2.43}$$

Since $\boldsymbol{x}$ and $\boldsymbol{x}'$ are independent, $\mathrm{d}^n x' = \mathrm{d}^n r$ and

$$\langle \hat{g}(\boldsymbol{k}) \hat{g}^*(\boldsymbol{k}') \rangle = \int \mathrm{d}^n x\, e^{i\boldsymbol{x}.(\boldsymbol{k}-\boldsymbol{k}')} \int \mathrm{d}^n r e^{-i\boldsymbol{r}.\boldsymbol{k}'} C(r). \tag{2.44}$$

The first integral in Eqn. 2.44 is the Fourier transform of the Dirac delta function, δ_D; that is, $\int \mathrm{d}^n x\, e^{i\boldsymbol{x}.(\boldsymbol{k}-\boldsymbol{k}')} = (2\pi)^n \delta_D(\boldsymbol{k} - \boldsymbol{k}')$. Thus we can rewrite Eqn. 2.44 as

$$\langle \hat{g}(\boldsymbol{k}) \hat{g}^*(\boldsymbol{k}') \rangle = (2\pi)^n \delta_D(\boldsymbol{k} - \boldsymbol{k}') P(|\boldsymbol{k}'|) \tag{2.45}$$

where we have defined the power spectrum of $g(\boldsymbol{x})$, $P_\delta(k)$, to be [29]

$$P_\delta(k) = \int \mathrm{d}^n r e^{-i\boldsymbol{r}.\boldsymbol{k}'} C(|\,\boldsymbol{r}|). \tag{2.46}$$

This is the 3D power spectrum of mass fluctuations defined in Eqn. 1.45.

P_κ is the **effective convergence power spectrum** and is a measure of the two–point statistics of weak gravitational lensing. It can be found in the same way as $P_\delta(k)$ from the Fourier transform of the two–dimensional effective convergence two–point correlation function, $C_{12} = \langle \bar{\kappa}_{\mathrm{eff},1}(\boldsymbol{\theta}) \bar{\kappa}_{\mathrm{eff},2}(\boldsymbol{\theta}') \rangle$, where $\bar{\kappa}_{\mathrm{eff}}$ is the projection of the density fluctuations δ along the line of sight. It is related to $P_\delta(k)$ by

$$P_\kappa(l) = \frac{9 H_0^4 \Omega_m^2}{4c^4} \int \mathrm{d}w\, \frac{W^2(w)}{a^2(w)} P_\delta\left(\frac{l}{f_K(w)}, w\right) \tag{2.47}$$

where f_K is the comoving angular diameter distance out to radial distance w defined by the curvature of the Universe K, l is the multipole moment and W is a weighting function [114]. l and k are related through the Limber approximation; k and l are the Fourier transforms of the radius and angular scale respectively,

and $k_\perp = l/\chi$, where $k_\perp$ is the transverse component of $\boldsymbol{k}$ and χ is the comoving angular diameter distance (see e.g. [115] for an explanation). The convergence power spectrum may also be written in terms of the shear correlation functions, ξ_+ and ξ_- (see §2.3.4),

$$P_\kappa(l) = \pi \int \mathrm{d}\theta\, \theta \left[\xi_+(\theta) J_0(l\theta) + \xi_-(\theta) J_4(l\theta)\right] \tag{2.48}$$

where J_0 and J_4 are Bessel functions. We refer the reader to [29] for a derivation of Eqns. 2.47 and 2.48.

2.3.4 Shear Correlation

The shear correlation function $\langle\gamma\gamma\rangle_\theta$ can be measured from the ellipticities of galaxy pairs at angular separations θ. It is possible to decompose the shear correlation into tangential and radial components as follows. First, we consider the tangential and radial ellipticity parameters, ϵ_t and ϵ_r, of a galaxy in a coordinate frame defined by the line connecting the centroids of two galaxies,

$$\begin{pmatrix} \epsilon_t \\ \epsilon_r \end{pmatrix} = \begin{pmatrix} \cos 2\phi & \sin\phi \\ -\sin 2\phi & \cos 2\phi \end{pmatrix} \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \end{pmatrix}, \tag{2.49}$$

where ϕ is the angle between the new frame and the original frame in which ϵ_1 and ϵ_2 were defined. From Eqn. 2.40 we know that the ellipticity and the shear of a galaxy are easily interchangeable, so this allows us to re-write Eqn. 2.49 in terms of the tangential shear γ_t and the radial shear γ_r,

$$(\gamma_1 + i\gamma_2)(\cos 2\phi - i \sin 2\phi) = \gamma_t + i\gamma_r = \gamma e^{-2i\phi} \tag{2.50}$$

where γ is the complex shear. The shear correlation functions are thus [116]

$$\xi_\pm = \langle\gamma_\mathrm{t}\gamma_\mathrm{t}\rangle \pm \langle\gamma_\mathrm{r}\gamma_\mathrm{r}\rangle = \frac{1}{2\pi}\int_0^\infty \mathrm{d}l\, l P_\kappa(l) J_{0,4}(l\vartheta). \tag{2.51}$$

The power spectrum relates the observed lensing correlation functions to the cosmological parameters through Eqn. 2.47. Practical estimators for the shear correlation functions are given by [116] as

$$\hat{\xi}_{\pm}(\theta) = \frac{\sum_{ij} w_i w_j (\epsilon_{it}\epsilon_{jt} \pm \epsilon_{ir}\epsilon_{jr})\Delta_{\vartheta}(|\boldsymbol{\vartheta}_i - \boldsymbol{\vartheta}_j|)}{N_p(\theta)}. \tag{2.52}$$

The sum runs over all galaxy pairs at angular positions $\boldsymbol{\vartheta}$. $N_p(\theta)$ is the effective number of pairs in the angular bin based on their weights w_i, which reflect the quality of the galaxy shape measurement. ϵ_{it} and ϵ_{ix} denote the tangential and cross components of the complex ellipticity of a galaxy i with respect to the line joining it to galaxy j. The galaxy weights are normally defined by the quality of the shape measurement. $\Delta_\theta(\phi) = 1$ if the angular separation θ of the galaxy pair is centred on θ, and zero otherwise.

As discussed in §2.3.1, the observable ellipticity of a galaxy, ϵ_i, is a combination of the intrinsic ellipticity ϵ_i^s and the shear at that position in the weak lensing regime where $|\gamma| \ll 1$. If the assumption that galaxy shapes are intrinsically uncorrelated is correct, then $\langle \epsilon_i^s \epsilon_j^{s*} \rangle_\theta = 0$ and $\langle \gamma_i \gamma_j^* \rangle_\theta = \langle \epsilon_i \epsilon_j^* \rangle_\theta$. We can therefore simply correlate galaxy shapes to directly estimate the shear correlation function at a given angular separation.

2.3.5 Systematics

Gravitational lensing is a clean cosmological probe as it does not rely on assumptions about the nature of the underlying matter distribution, in principle making it easy to infer cosmological information from. However, with upcoming telescope surveys providing better statistical constraints on the cosmic shear signal than ever before (see §2.3.6), systematic errors in weak lensing data will dominate the error contribution. It is therefore vitally important to understand and account for these systematics if more accurate and precise cosmological parameter estimates are to be made in the future. In this subsection, we discuss the main sources of systematic error in cosmic shear measurements and how to characterise them.

E and B modes

One way to characterise the measured distortions caused by gravitational lensing is to decompose the shear field into divergence and curl components, known as E– and B–modes. The scalar gravitational potential produces a curl-free shear signal, i.e. only E–modes. B–modes are produced by significant curl components

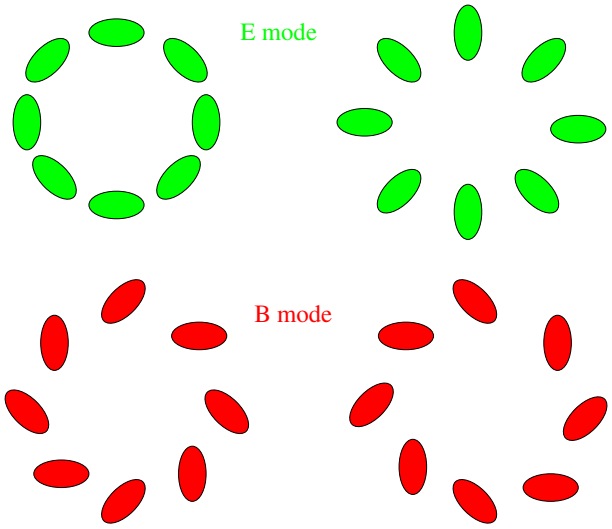

Figure 2.6: E–modes generated by an overdense region (*top left*) and an underdense region (*top right*); below are the B–modes generated by systematics in the same regions [30].

which are not allowed by weak lensing, so their presence is indicative of systematic errors in the data[10]. Note, however, that the absence of B–modes does not imply the absence of systematics, only that any remaining systematics are E-mode only. Fig. 2.6 shows the shear pattern produced by E– and B–modes generated in overdense and underdense regions [85]. Since the B–modes are a 45 ° rotation of the E–modes, they can be detected by rotating the data by 45 ° and repeating the lensing analysis. B–modes will now show up as a signal.

The E– and B– modes of the lensing potential ψ_E and ψ_B are defined by

$$\psi = \psi_E + i\psi_B. \tag{2.53}$$

The E– and B–mode power spectra can be found by decomposing the convergence into E– and B–modes,

$$\kappa(\boldsymbol{\theta}) = \kappa_E(\boldsymbol{\theta}) + i\kappa_B(\boldsymbol{\theta}), \tag{2.54}$$

where κ_E is due to lensing and systematic errors, κ_B is due to errors only and

[10]Weak lensing can generate B–modes due to a breakdown on the Born approximation [117] or through the clustering of source galaxies [118], but they are typically small and only significant on very small angular scales of $\theta \lesssim 1'$.

both κ_E and κ_B are real. In Fourier space with wavenumber l this is

$$\kappa_{\boldsymbol{l}} = \kappa_{E,\boldsymbol{l}} + i\kappa_{B,\boldsymbol{l}}. \tag{2.55}$$

Since κ_E and κ_B are real, $\kappa^*_{E,\boldsymbol{l}} = \kappa_{E,-\boldsymbol{l}}$ and $\kappa^*_{B,\boldsymbol{l}} = \kappa_{B,-\boldsymbol{l}}$, so we can write

$$\kappa^*_{\boldsymbol{l}} = \kappa^*_{E,\boldsymbol{l}} - i\kappa^*_{B,\boldsymbol{l}} = \kappa_{E,-\boldsymbol{l}} - i\kappa_{B,-\boldsymbol{l}} \tag{2.56}$$

$$\kappa^*_{-\boldsymbol{l}} = \kappa_{E,\boldsymbol{l}} - i\kappa_{B,\boldsymbol{l}}. \tag{2.57}$$

The corresponding correlation functions will be

$$\langle\hat{\kappa}(k)\hat{\kappa}^*(k')\rangle = (2\pi)^2\delta_D(\boldsymbol{k}-\boldsymbol{k'})\left[P_E(k)+P_B(k)\right], \tag{2.58}$$

$$\langle\hat{\kappa}(k)\hat{\kappa}(k')\rangle \;= (2\pi)^2\delta_D(\boldsymbol{k}-\boldsymbol{k'})\left[P_E(k)-P_B(k)+2iP_{EB}(k)\right] \tag{2.59}$$

where $P_E(k) = P_\kappa(k)$, the convergence E–mode power spectrum, $P_B(k)$ is the B–mode power spectrum and $P_{EB}(k)$ is the cross power spectrum. According to [118], the B–mode power spectrum and the cross power spectrum can be written in terms of the shear correlation functions, ξ_+ and ξ_- as

$$\begin{aligned} P_B(l) &= \pi\int \mathrm{d}\theta\,\theta\left[\xi_+(\theta)J_0(l\theta) - \xi_-(\theta)J_4(l\theta)\right] \\ P_{EB}(l) &= 2\pi\int \mathrm{d}\theta\,\xi_{\mathrm{x}}(\theta)J_4(l\theta). \end{aligned} \tag{2.60}$$

Data systematics

In weak lensing measurements, the most significant systematics are data-related. Lensing requires us to detect percent-level distortions to galaxy shapes, but spatially- and temporally-varying distortions of the order of $\sim 10\%$ are caused by the atmosphere, the telescope and the detector. The PSF mentioned in §2.3.2 encapsulates these artificial sources of distortion and if unaccounted for will swamp the shear signal by an order of magnitude. Any errors in the PSF modelling or deconvolution will lead directly to an error on the measured ellipticity and hence the shear, which will produce biases in parameter estimates.

Intrinsic Galaxy Alignments

The presence of significant B–modes in the shear field is indicative of a systematic error. One of the main sources of systematic error in weak lensing data is the presence of intrinsic galaxy alignments. Such alignments challenge the assumption that galaxies are orientated randomly on the sky; whilst this assumption is approximately true on large spatial scales, it is known that the shapes of galaxies that are close together on the sky can be correlated. These correlations can mimic the lensing signal.

For the purposes of weak lensing analysis there are two main types of intrinsic alignments. The first, often referred to as the **intrinsic–intrinsic (II) alignment signal** is important for galaxies that are physically close together in space and arises as a consequence of galaxy formation mechanisms. As a galaxy forms, its ellipticity will be determined by the shape of the dark matter halo in which it forms. A halo experiencing gravitational collapse in a tidal field will collapse asymmetrically, acquiring a torque. This torque will cause the halo to become a thin, oblate disc orientated normal to its angular momentum vector. Any galaxies forming within the halo will experience a very similar tidal field and therefore will have a tendency to be aligned with one another. The contaminating effect of II alignments is strongly dependent on the depth of the survey, since for a given angular separation, galaxies that are at high redshift will be more widely separated in real space than close-by galaxies [20, 85].

The second type of intrinsic alignment is the shear–shape correlation or **gravitational–intrinsic (GI) alignment**. This effect is more subtle than the II alignment and was not considered as a potential systematic until as recently as 2004 [119]. The effect arises from galaxies that lie at small angular separations but at very different redshifts. A foreground galaxy will have an orientation that is influenced by the local tidal field that it forms in, and an elliptical galaxy will typically have its major axis orientated radially with respect to an overdensity. The light from a background galaxy along a similar line of sight will be tangentially sheared by the tidal field of the foreground so that the two galaxy shapes are anti-correlated (See Fig. 2.7). This effect dampens the observed lensing shear power spectrum [85].

The effect of the II and GI alignments is to alter the relation $\langle \epsilon_i \epsilon_j^* \rangle_\theta \simeq \langle \gamma_i \gamma_j^* \rangle_\theta$, as the measured ellipticity correlation is no longer a clean measure of the shear

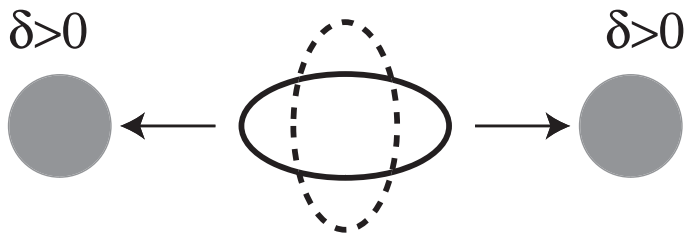

Figure 2.7: The effect of the GI correlation on the shear power spectrum. The overdense regions in the foreground (grey masses) induce a tidal field (arrows) that causes the major axis of the foreground galaxy (solid ellipse) to become aligned with the field. The distant galaxy (dashed ellipse) is gravitationally sheared tangential to these masses, resulting in an anti-correlation between the shapes of galaxies at different redshifts [119].

correlation, but will contain contributions from these systematics,

$$\left\langle \epsilon_i \epsilon_j^* \right\rangle_\theta = \underbrace{\left\langle \gamma_i \gamma_j^* \right\rangle_\theta}_{\text{GG}} + \underbrace{\left\langle \epsilon_i^{\rm s} \epsilon_j^{\rm s*} \right\rangle_\theta}_{\text{II}} + \underbrace{\left\langle \gamma_i \epsilon_j^{\rm s*} \right\rangle + \left\langle \epsilon_i^{\rm s} \gamma_j^* \right\rangle_\theta}_{\text{GI}} . \tag{2.61}$$

Here, $\left\langle \epsilon_i^{\rm s} \epsilon_j^{\rm s*} \right\rangle_\theta$ is the contribution to the lensing signal from the II alignments and $\left\langle \gamma_i \epsilon_j^{\rm s*} \right\rangle + \left\langle \epsilon_i^{\rm s} \gamma_j^* \right\rangle_\theta$ is the contribution from GI alignments. If galaxy j is at a higher redshift than galaxy i, then $\left\langle \gamma_i \epsilon_j^{\rm s*} \right\rangle$ is zero (except in the presence of significant photometric redshift errors as we will see in Chapter 5). The effect of these systematics is strongly dependent on both angular scale and redshift, and the contamination to the lensing signal is not negligible. The effect of IAs on the observed lensing signal, and various methods to take them into account in data analysis are discussed in more detail in Chapter 5.

Photometric Redshifts

The deflection of light by large–scale structure depends on the three-dimensional matter distribution, so in order to accurately measure cosmological parameters from lensing surveys it is important to know the galaxy redshift distribution of a survey. Currently, the uncertainty in the redshift distribution is one of the main sources of error in measuring the amplitude of the matter power spectrum σ_8. For surveys without measurements of the redshift distribution, a model distribution could be fitted to the data, but if the model is incorrect this will

introduce errors which can bias cosmological constraints. For most contemporary and future lensing surveys, the redshift distribution will be characterised by taking photometric redshifts (photo-zs) of all the survey galaxies. Traditionally, redshifts are obtained via spectroscopy from the redshift of emission lines in the spectra of galaxies, but attaining spectra for all survey galaxies is unfeasibly slow at the depths required for weak lensing surveys. Spectra take time to obtain, and the number of galaxies expected in near–future lensing surveys is high ($\sim 10^8$–10^9) meaning we typically only obtain spectra for a small subsample of survey objects [20, 120].

Photometric redshifts are obtained through photometry using several broad–band filters and relies on the detection of strong features (such as the Lyman break at 4000 Å), which can be seen in relatively crude filters. As a result, photometric redshifts are less precise than spectroscopic ones, but they can be obtained to a fainter magnitude and much more quickly than spectroscopic redshifts. This means that it is possible to obtain photometric redshifts for every galaxy in even the largest weak lensing survey. Errors as small as $\Delta z/(1+z) \approx 0.03$ have been obtained with five broad–bands [121]. Greater photometric redshift accuracy allows more accurate and precise measurement of the lensing signal, and allows better determination of IA contaminants. One of the main sources of systematic error on photometric redshift is 'catastrophic outliers', which are galaxies for which the photometric redshift is very wrong. This occurs when, for example, a break in the spectrum at one wavelength is mistaken for a different break at another redshift. Even a small number of these outliers can significantly impact cosmological parameter estimates. It is possible, however, to reduce the effect of these outliers dramatically by calibrating the entire sample of photometric redshifts with a smaller sample of spectroscopic redshifts for some of the survey objects (typically $\sim 10^5$–10^6), which can vastly increase the redshift accuracy [122, 120].

2.3.6 Lensing Surveys

It is only in relatively recent times that weak gravitational lensing has been used as a competitive probe of the cosmology of the Universe. As a result there are a number of contemporary surveys designed to measure weak lensing, including

the the Sloan Digital Sky Survey[11] (SDSS), the Panoramic Survey Telescope And Rapid Response System[12] (Pan-STARRS), Hyper Suprime Cam on the Subaru telescope[13], and several wide-field surveys scheduled for release in the next decade such as the Large Synoptic Survey Telescope[14] (LSST), the Dark Energy Survey[15] (DES), the Kilo Degree Survey[16] (KIDS), the Wide Field InfraRed Survey Telescope[17] (WFIRST) and Euclid[18]. The University of Edinburgh is part of the science consortium for a contemporary lensing survey of relevance to this thesis: the Canada–France–Hawaii Telescope Lensing Survey[19] (CFHTLenS).

CFHTLenS uses data from the 3.6 m Canada–France–Hawaii Telescope Legacy Survey (CFHTLS) in Hawaii collected between 2003 and 2004. It consists of two surveys: the 'Wide' survey covering 154 square degrees, and the 'Deep' survey which covers 4 square degrees, both used to measure weak gravitational lensing. CFHTLenS uses five colour bands (u^*,g',r',i',z') with a limiting magnitude of $g' = 26.6$ in the Wide survey and $g' = 28.9$ in the Deep survey. This corresponds to a mean redshift of ~ 0.81 and ~ 1.10 respectively. Images were taken with the 1 sq. degree field-of-view MegaCAM camera. Previous lensing data from telescopes such as Keck and the Hubble Space Telescope, and the Blanco Cosmology Survey have typically covered several tens of square degrees down to a magnitude of ~ 25, making CFHTLenS the largest and deepest weak lensing survey to date. The use of five filters means that photometric redshifts have been obtained for every object, with a mean scatter of $\Delta z = 0.04(1 + z)$ and a 4% catastrophic outlier rate. Early analysis of a subset of CFHTLenS data has yielded constraints on the matter power spectrum $\sigma_8 = 0.771 \pm 0.029$ and the matter density parameter $\Omega_m = 0.248 \pm 0.019$ when combined with WMAP3 results [50], the dark energy equation of state parameter $-0.10 < 1 + w < 0.06$ from combined CFHTLenS–Wide, CFHTLenS–Supernova Legacy Survey and WMAP5 data [86] and constraints on the neutrino masses of $0.03\text{eV} < \Sigma m_\nu < 0.54$ eV to 95% confidence from a combination of probes [123].

[11] `http://www.sdss.org`.
[12] `http://pan-starrs.ifa.hawaii.edu/public/home.html`.
[13] `http://www.naoj.org/Projects/HSC/index.html`.
[14] `http://www.lsst.org/lsst`.
[15] `https://www.darkenergysurvey.org/the-project`.
[16] `http://www.astro-wise.org/projects/KIDS/`.
[17] `http://wfirst.gsfc.nasa.gov/`.
[18] `http://sci.esa.int/euclid`.
[19] `http://www.cfhtlens.org`.

An early analysis was performed on several tens of square degrees of CFHTLenS data using KSB and preliminary results have been released in [124, 50]. We review the analysis of CFHTLenS and the cosmic shear results achieved in more detail in Chapter 7.

The lessons learnt from CFHTLenS will be applied to upcoming astronomical surveys that hope to constrain cosmology through weak lensing, such as KIDS, DES and Euclid. All these surveys will use multiple astrophysical phenomena to measure cosmological parameters. KIDS will study dark matter and dark energy through weak lensing in addition to observing high redshift quasars, galaxy clusters and studying galaxy evolution. DES will simultaneously use four probes to better measure dark energy parameters - weak lensing, SNe Ia, BAOs and galaxy clustering. These probes are complimentary and a comparison of the results may give us unprecedented information about the dark energy equation of state. Both KIDS ands DES are considerably larger than CFHTLenS and represent the beginning of the next generation of wide-field astronomical surveys that will observe more galaxies for lensing than ever before. Euclid is a space-based European Space Agency mission due to launch at the end of this decade. It will be optimised for weak lensing and BAO observations, but will also measure galaxy clustering, redshift space distortions and the ISW effect, all with a focus on measuring the accelerated expansion of the Universe to high precision. Some of the key facts and figures for these surveys are shown in table 2.1.

2.3.7 Weak Lensing Tomography

Tomography is the process of imaging by sections, and in weak lensing analysis it refers to the binning of data by redshift. Weak lensing tomography is useful because it recovers some of the statistical information contained in galaxy redshifts and can improve error estimation on cosmological parameters. Surveys such as the completed CFHTLenS, and the upcoming KIDS, DES and Euclid are capable of obtaining photometric redshift estimates for all survey galaxies, allowing tomographic redshift binning of the data. Tomographic analyses recover more of the statistical information leading to smaller parameter uncertainties [82, 125, 126, 127, 128, 129, 130]. In addition, the growth of structure in the Universe depends on the expansion history, meaning that a redshift-sensitive analysis can put tighter constraints on the dark energy equation of state in

Table 2.1: Comparison of CFHTLenS and other lensing surveys

	CFHTLenS	KIDS	DES	Euclid
Location	CFHT, Mauna Kea	VST, Paranal	Blanco, CTIO	L2 (space)
Filters	u^*,g',r',i',z'	u',g',r',i' (+IR from VIKING[a])	R,I,Z,Y,J,H	
Survey Area	154 deg^2 (wide)	1500 deg^2	5000 deg^2	15000 deg^2 (wide)
	4 deg^2 (deep)			40 deg^2 (deep)
Limiting Magnitude	$g' = 26.6$ (wide)	$g' = 25.4$	$g' = 24.6$	R,I,Z=24.5
(10σ AB)	28.9 (deep)			
Median seeing	0.7”	0.6”	0.9”	0.1”
Pixel size	0.186”	0.216”	0.27”	0.1”
Field of view	1 deg^2	1 deg^2	2.2 deg^2	0.5 deg^2
Primary mirror	3.6m	2.65m	4m	1.2m
Galaxy density/sq. arcmin	17	12	10	30
Survey start	2003	2011	2012	2019 (expected)

[a]VISTA Kilo-degree INfrared Galaxy survey, `http://www.eso.org/public/teles-instr/surveytelescopes/vista/surveys.html`

particular [49, 131, 132, 133, 134]. Tomography is particularly useful in wide–field imaging where most the lensing information is contained within the two point correlation functions and their corresponding power spectra. This regime is still largely unexplored, although surveys such as Euclid will change this. Weak lensing tomography will be discussed further in Chapters 4 & 5.

2.4 Conclusion

Weak gravitational lensing has recently matured into a powerful cosmological probe in its own right, and measurements of the parameters involved in structure growth (Ω_m, σ_8) from lensing become ever more precise. However there exists a degeneracy between these parameters that may be broken by combining two–dimensional weak lensing results with other probes such as the CMB, baryon acoustic oscillations and SNe Ia. To this end, accurate knowledge of the redshift distribution through photometric redshift measurements is vital. The degeneracies in weak lensing measurements tend to be orthogonal to those of other probes, so combining results can lead to very tight cosmological constraints.

The upcoming generation of wide-field lensing surveys promises to better measure cosmological parameters from lensing than ever before. We will be able to probe the lensing power spectrum on larger angular scales, extending measurements into the linear regime of structure formation. Weak lensing is a very 'clean' probe of cosmology that relies on few assumptions, and as a result the greatest challenge in measuring cosmic shear in the coming years will be to reduce the systematic errors to below the level of the statistical errors. Accounting for catastrophic outliers in the photometric redshift distribution, accurately modelling the PSF, galaxy shape measurement and perhaps most importantly the intrinsic alignments between galaxies remain the main difficulties in shear analysis. Performing tomography by separating data into redshift slices can aid the removal of intrinsic alignments, as well as recovering much of the shear information contained in the redshift.

Chapter 3

The Markov Chain Monte Carlo Algorithm

3.1 Introduction

In this chapter, we discuss a statistical technique based on importance sampling that was used to conduct the analysis presented in this thesis. Importance sampling is the name for the general technique of determining the properties of a distribution by drawing samples from another distribution. On the surface of it, this seems like a strange thing to do — why would we want to sample from a distribution that is different to the one under investigation? The answer is that in many model selection or parameter estimation problems, we do not have a direct way of determining the properties of the distribution of interest. By drawing from a sample distribution which is (over a large enough sample size) representative of the one under investigation, we can directly infer some of its properties.

Importance methods are widely used in statistical analysis because often one does not have a direct way of determining the properties of the distribution of interest. This is often the case in parameter estimation or model selection problems where a Bayesian approach is needed (as is standard practice in the field of cosmology). There are a multitude of techniques for producing sample distributions, and one of the most commonly utilised techniques is the Markov Chain Monte Carlo (MCMC) algorithm. The construction of an MCMC algorithm is an important part of this thesis, and in this chapter we introduce the concept of importance sampling by explaining in detail this sampling technique

in a Bayesian statistical framework. We discuss the advantages of the technique over traditional grid-based analyses, and highlight some of the considerations that must go into constructing an MCMC. We mention some of the limitations of the algorithm and some alternative importance sampling methods that aim to circumvent some of these problems. Finally, we discuss the implementation of our MCMC algorithm in preparation for the analysis presented in Chapters 4-7. The MCMC constructed by the author was used throughout the analysis presented in this thesis.

3.2 Bayes' Theorem

The analysis presented in this thesis is based on Bayesian statistics, which holds Bayes' theorem central. For some data D with parameter(s) π, Bayes' theorem holds that [135, 136]

$$\mathrm{p}(\pi)\,\mathcal{L}(\pi) = E\,\mathcal{P}(\pi). \tag{3.1}$$

Here $\mathrm{p}(\pi)$ is known as the **prior probability**, or prior, representing how we originally distribute the parameters' probability ($\mathrm{p}(\pi)$ denotes the probability of π). The prior should not depend on the data set D being investigated; it characterises our degree of belief in a hypothesis. One must first decide on the range of the parameters π which then defines a 'hypothesis space' over which the prior probabilities must be distributed, with the only constraint that the sum of the probabilities is normalised to unity. The choice of prior in a given analysis is often debated, and is usually influenced by data from earlier experiments or observations. In the absence of any previous data to inform our choice of prior, one often uses a 'flat prior' — to assign an equal probability to every region of parameter space.

$\mathcal{L}(\pi) = \mathrm{p}(D\,|\,\pi)$ is the **likelihood**, or probability of the data D occurring given the parameters π. It is often possible to calculate the data value expected from known values of π, which allows us to calculate the likelihood of the data D for any given point in parameter space. The ability to calculate the likelihood in this way proves extremely useful in importance sampling as we will see in section 3.4.

The **evidence** is given by $E = \mathrm{p}(D) = \int \mathrm{p}(\pi)\mathcal{L}(\pi)d\pi$. This represents how well the priors managed to predict the data, or the average of the likelihood over

the whole parameter space [136]. The evidence is the normalising constant in Bayes' theorem. It is most useful in the problem of model selection; that is, when we have two or more models with different parameter sets. This is because the evidence will be higher for a model if more of its parameter space has a high likelihood, and lower for a model with large regions of its parameter space having low likelihood. The evidence allows one to decide which of two models E_0 and E_1 better fits a given data set (with the assumed priors) through the Bayes factor [137, 138],

$$B_{01} \equiv \frac{E_0}{E_1}. \tag{3.2}$$

The evidence naturally implements Occam's razor, as simpler models with fewer parameters will generally have larger evidence than more complicated models. For the simpler problem of parameter estimation within a model (as is the case with the analysis presented in this thesis), the evidence does not need to be calculated explicitly. This is because often we are not interested in the absolute likelihood of a set of parameters (which E gives us); rather, we are comparing the goodness-of-fit of different combinations of parameter values with each other. In this case, only $\mathcal{L}(\pi)$ for each parameter value is needed.

Finally, the **posterior probability**, or posterior, is given by $\mathcal{P}(\pi) = \mathrm{p}(\pi \mid D)$. The posterior represents the inferred distribution of probability among the models in our parameter space, and it is this distribution that we seek to measure. The posterior tells us the probability of a parameter set being true given the data and can be used to determine which models are favoured over others, or which parameter values are a better fit to the data than others.

These definitions allow us to express the posterior in the presence of any prior information I as

$$\mathrm{p}(\pi \mid DI) = \frac{\mathrm{p}(\pi \mid I)\,\mathrm{p}(D \mid \pi I)}{\mathrm{p}(D \mid I)}. \tag{3.3}$$

The focus of importance sampling is to determine as easily and accurately as possible the properties of the posterior from a representative sample from a second distribution. It is worth noting that with a new data set, the identities of the terms in Bayes' theorem can shift — for example, the posterior derived from one data set can be used as the prior for a new data set.

3.2.1 Using Bayes' Theorem

Sometimes we may not wish to constrain every parameter in a model at once. Marginalisation allows us to reduce the number of parameters we are constraining by ignoring the values of those we are not interested in. For example, if our model depends on two parameters (x, y) and we are only interested in x, the marginalised posterior probability for x is [139]

$$\mathrm{p}(\pi(x) \,|\, DI) = \int_{-\infty}^{\infty} \mathrm{p}(\pi(x, y) \,|\, DI)\, \mathrm{d}y. \tag{3.4}$$

The marginalised likelihood can therefore be expressed using some prior information about y as

$$\mathrm{p}(\pi(x) \,|\, I)\ \mathrm{p}(D \,|\, \pi(x)I) = \int_{-\infty}^{\infty} \mathrm{p}(\pi(x, y)I)\ \mathrm{p}(D \,|\, \pi(x, y)I)\, \mathrm{d}y. \tag{3.5}$$

Marginalisation is useful because it allows us to ignore parameters that are unimportant for a particular analysis. For example, when determining best-fit galaxy ellipticity parameters (ϵ_1, ϵ_2) we can marginalise over galaxy parameters such as brightness, position and size. When performing a cosmological analysis, marginalisation allows us to only fit the cosmological parameters of interest; if our focus is on fitting Ω_m and σ_8 we may wish to marginalise over other cosmological parameters that influence the lensing signal such as the dark energy parameters w_0 and w_a [20].

Of particular use in a model-fitting analysis is the χ^2 statistic. χ^2 is defined as the sum of the squares of the normalised residuals between some data $D(\boldsymbol{x}_i)$ and model $\pi(\boldsymbol{x}_i)$ [140],

$$\chi^2 = \sum_{i=1}^{N} \left(\frac{D(\boldsymbol{x}_i) - \pi(\boldsymbol{x}_i)}{\sigma_i} \right)^2, \tag{3.6}$$

where σ_i are the associated standard deviations assuming the data are independent[1]. The lower the value of χ^2, the better a fit to the data a parameter set is (in the limit of $\chi^2 = 0$ for a perfect fit). Calculating χ^2 for multiple parameter combinations and comparing the results is known as least-squares fitting. If the

[1]If the data are not independent, as is the case for the shear correlation function, χ^2 is modified to include the covariances of the data (see Chapter 4).

errors σ_i are Gaussian, χ^2 has a known distribution, and it is possible to determine if the parameters are a good fit to the data using the rule of thumb that a good fit falls within $\chi^2 = \nu \pm \sqrt{2\nu}$. ν is the number of degrees of freedom, determined by $\nu = p - n_{\rm par}$ where p is the number of data points and $n_{\rm par}$ is the number of model parameters being fitted.

χ^2 is related to Bayes' theorem through the likelihood. In the limit of a large number of observations where the central limit theorem applies[2], the likelihood $\mathcal{L}(\pi) = \mathrm{p}(D \,|\, \pi I)$ is related to the χ^2 distribution by

$$\ln(\mathcal{L}) = -\frac{1}{2}\chi^2. \tag{3.7}$$

Thus, the best fit parameter set (out of those tested) is given by the maximum likelihood $\mathcal{L}_{\rm max}$ and credibility intervals around this peak likelihood can be found. For example, the 68.3% confidence level, corresponding to 1σ errors (for Gaussian distributed independent data) is assigned a specific $\Delta\chi^2$ value dependent on ν. If $\nu = 1$, $\Delta\chi^2 = -2\ln(\mathcal{L}/\mathcal{L}_{\rm max}) = 1$ and for $\nu = 2$, $\Delta\chi^2 = 2.3$ [141].

Often, χ^2 is quoted in terms of the reduced χ^2, $\chi^2_{\rm red} = \chi^2/\nu$, which will be close to unity for a good fit[3] and so makes for an easy rule-of-thumb check of the goodness-of-fit. In a Bayesian framework, however, we are often interested in not just how good a fit to the data a particular parameter set is: rather, we are concerned with comparing the goodness-of-fit of different parameter values with each other. Comparing a single model or parameter value to the data and accepting it as a good fit based on the fact that $\chi^2_{\rm red} \sim 1$ can mean that we erroneously do not consider other values that may be an even better fit. Fitting many parameter values to the data and drawing confidence limits around the best fit ensures that we take into account inherent statistical uncertainty in the data.

The χ^2 statistic has the advantages of being simple, easy to calculate from data and relatable to the posterior distribution through the likelihood. It is the statistic we choose to use in this thesis for sampling the posterior distribution and determining confidence limits throughout. In the next section, we discuss

[2]The central limit theorem states that a sufficiently large number N of measurements of an independent random variable (with finite mean and variance σ^2) will tend to be normally distributed about the expectation value for the mean with variance σ^2/N.

[3]This is not the case for correlated data, where the effective number of degrees of freedom is modified. In the presence of correlated data, absolute χ^2 values are less useful but a Bayesian model/parameter comparison can be employed to obtain constraints.

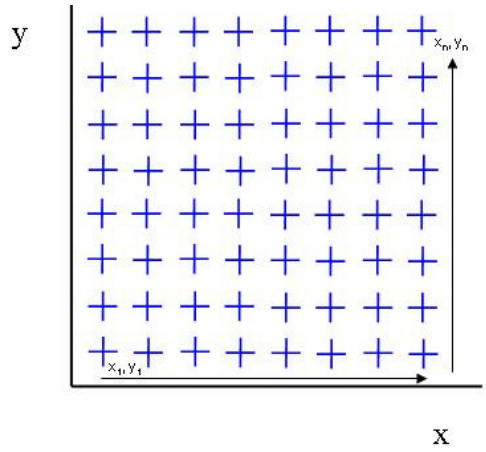

Figure 3.1: Discrete, grid-based sampling of points in parameter space of x and y. The blue crosses represent coordinates in the parameter space at which we perform a least-squares fit to the data.

several algorithms constructed for this purpose.

3.3 Grid-based Analyses

For a given hypothesis or parameter space, the most obvious way to probe the posterior distribution is to use a grid–based approach to calculating the likelihood. Here, one calculates the likelihood of the data at discrete, evenly–spaced points in the parameter space as shown in Figure 3.1 for the two–dimensional parameter space of x and y. Although we have chosen to vary two parameters in this example for ease of visual representation, in principle there is no constraint on the number of dimensions in the analysis. We compare the theoretical prediction of the measurement for a given set of model parameters (x_1, y_1) with the data D to find the likelihood at point (x_1, y_1) and repeat for all points through to (x_n, y_n). This will give an idea of the location of the best fit solution and the shape and position of the likelihood contours.

This is a simple way to estimate the posterior, but it has some serious flaws. The main problem with this method is that the number of grid points needed scales exponentially with the number of dimensions (assuming we sample each dimension the same number of times, i.e. n is constant for all parameters) and computation times quickly become prohibitive. A further issue is the nature of gridding itself — how does one decide where to sample the likelihood, and how

finely spaced should the grid points be? We run the risk of missing peaks in the likelihood or fine detail in the likelihood surface if it lies between grid points.

The gridded likelihood calculation by its nature samples all of the parameter space evenly, meaning much of our computing resources are wasted probing low likelihood regions of the parameter space. Statisticians thus need a tool to help them sample multi–dimensional parameter spaces efficiently while building up a faithful representation of the posterior. Importance sampling provides this tool. In order to understand how importance sampling works, its advantages and its limitations, in the next section we will consider one of the most widely–used examples of the technique: the Metropolis–Hastings MCMC.

3.4 The Metropolis-Hastings MCMC Algorithm

A Metropolis-Hastings MCMC algorithm was constructed by the author for use throughout this thesis. The Metropolis-Hastings MCMC is a simple yet powerful method for importance sampling. In this section we will detail the construction of the Metropolis-Hastings MCMC algorithm, highlighting some of the key issues that need to be considered in its use. We will use 'MCMC' as shorthand for Metropolis-Hastings MCMC throughout.

Returning to our (x, y) parameter space that we considered when discussing a gridded likelihood analysis, we define a starting point for our algorithm. For simplicity, we choose a starting point $\pi_1 = \pi(x_1, y_1)$ randomly anywhere within our (x, y) parameter space, as shown on the left of Figure 3.2.

We calculate the likelihood of this point in parameter space, then step to a new point, π_2. This is done by drawing from a proposal distribution around π_1 and using a random number generator to select a position in the x and y directions for the new point. For simplicity, it is possible to use a tophat distribution defined by lengths $\Delta\pi_x$ and $\Delta\pi_y$ as shown in Figure 3.2 (*middle*). The size of the box is decided by the user, and is something that can (and should) be optimised to make the MCMC converge as quickly as possible. We discuss a technique for finding an efficient proposal distribution in section 3.4.3. For now, we choose a box size that is considerably smaller than our parameter space but not prohibitively small.

We then perform the crucial step in the MCMC – we either accept or reject the new point. This is done by calculating the ratio of the likelihoods of π_1 and π_2,

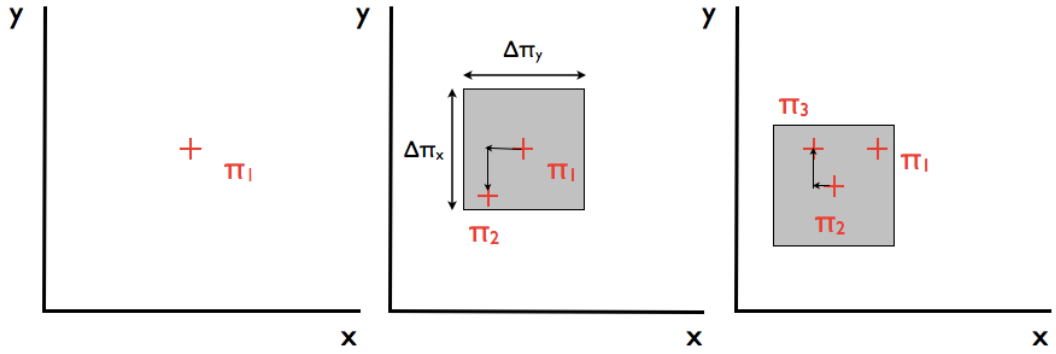

Figure 3.2: *Left*: The red cross defines the starting point π_1 in (x, y) parameter space. *Middle*: Stepping randomly from π_1 to π_2 within the tophat proposal distribution defined by $\Delta\pi_x$ and $\Delta\pi_y$. *Right:* Stepping from π_2 to π_3. The proposal distribution has shifted so that it is centred around the most recently accepted point.

such that if the new point has a higher likelihood than the old it will be accepted with a probability of 1, otherwise the probability of the point being accepted is equal to the ratio of the likelihoods. Formally this means that

$$P(\pi_i, \pi_{i+1}) = \min\left\{1, \frac{\mathcal{L}(\pi_{i+1})q(\pi_{i+1}, \pi_i)}{\mathcal{L}(\pi_i)q(\pi_i, \pi_{i+1})}\right\} \tag{3.8}$$

where $q(\pi_i, \pi_{i+1})$ is the proposal density distribution. For the algorithm to be Markovian we require that the proposal distribution does not change from one iteration to the next (otherwise the output distribution will not accurately sample the posterior) so $q(\pi_i, \pi_{i+1}) = q(\pi_{i+1}, \pi_i)$ and therefore

$$P(\pi_i, \pi_{i+1}) = \min\left\{1, \frac{\mathcal{L}(\pi_{i+1})}{\mathcal{L}(\pi_i)}\right\}. \tag{3.9}$$

It is this selection criterion that gives the Metropolis-Hastings algorithm its name. If π_2 is rejected, then the chain moves back to π_1 and selects another point within the proposal distribution at random, calculates the new likelihood and applies the selection criterion again and again until a point π_2 is accepted. Then the algorithm uses π_2 as its new starting point and selects another point π_3 in the chain from the proposal distribution centred around π_2, as shown in the right hand side of Figure 3.2.

The 'Markov Chain' in MCMC refers to this stepping behaviour; the steps should be discrete and random in addition to the proposal distribution $q(\pi_i)$ being a function of current position π_i only. The ratio of points accepted vs. points tried gives us our acceptance rate. If we are running a chain of n (accepted) points, a higher acceptance rate will take us to our nth point faster. For example, an acceptance rate of 50% means that we will have to perform $2n$ iterations to achieve a total chain of length n. A flowchart summarising the steps taken by the MCMC algorithm is shown in Fig. 3.3.

The MCMC proceeds in this way until we have built up a collection of points in the parameter space that are sufficient to be used in a likelihood analysis (see Figure 3.4 (*left*)). This is possible because of the elegant fact that the density of points in a given region of parameter space is directly proportional to the likelihood of that region, thanks the Metropolis-Hastings selection criterion. This ensures that the MCMC is by nature **ergodic**, meaning that any state (point in parameter space) is eventually reachable from any other with a probability of greater than zero. Ergodicity is important because we need to do more than just find the 'best fit' solution — we need to sample the area around the likelihood peak(s) to accurately represent all of the posterior. Thus in principle all of parameter space can be reached and the distribution of our MCMC points should approximate the posterior distribution we are seeking. There are several methods for drawing likelihood contours from the posterior, including using standard χ^2 values to determine contours, however one of the simplest and most reliable ways is to bin the samples on a grid to produce a 2D histogram, and draw contours around the grid points that contain the top (for example) 68.3% and 95.4% of points to produce the standard 68.3% and 95.4% contours (see Figure 3.4 (*right*)). This method works well assuming that the distribution of points faithfully reproduces the posterior distribution. In order for this to be the case, however, there are some additional checks that must be performed.

3.4.1 Convergence

To ensure that our MCMC chain is both robust and accurate, it must be convergent. This means that the chain has been run long enough to generate a

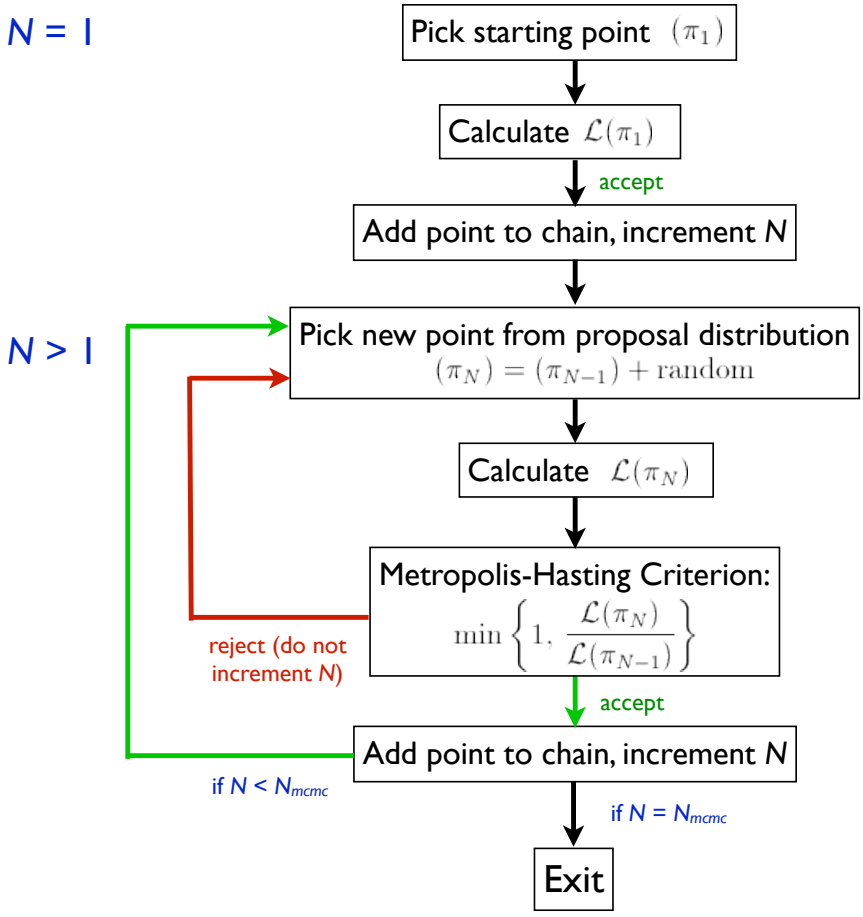

Figure 3.3: Flowchart showing the steps taken by an MCMC algorithm as it steps through the parameter space π. N indicates the current point in the chain, where $N = 1...N_{\text{MCMC}}$ and N_{MCMC} is the total number of desired points in the chain. The first point ($N = 1$) is always accepted.

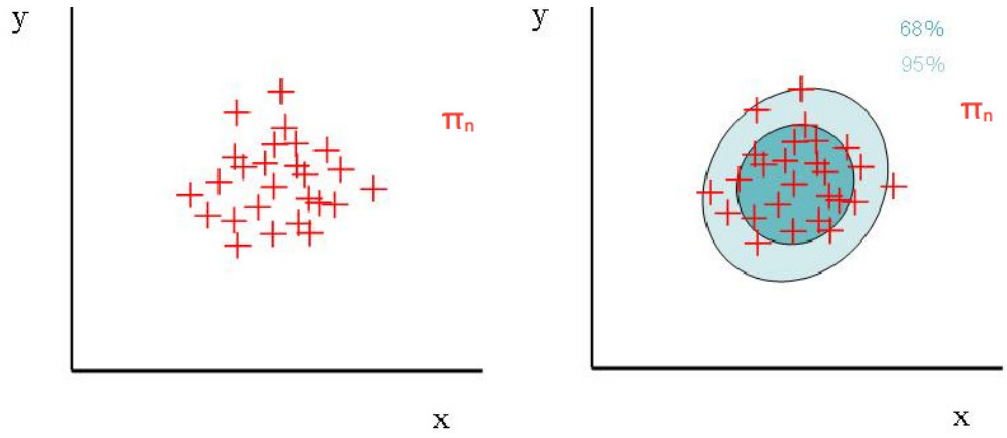

Figure 3.4: *Left*: A chain of n points in parameter space from MCMC. *Right*: Sketch of example likelihood contours from MCMC points.

distribution of independent points that closely match the posterior distribution. There is however no single conclusive test that can be run on the results of an MCMC that will tell us if we have achieved convergence, as it is a feature that is asymptotically reached. One easy test that we can perform is to run multiple chains, each with different random starting points, and compare the results. The variance between the chains should be much smaller than the posterior uncertainty on our measured parameter(s). This idea has been formalised in the form of the *Gelman-Rubin* statistic $R \equiv$ (variance between chains)/(mean variance within the chains). Typically, R should be as close to 1 as possible; preferably $R < 1.03$ [142, 143]. MCMC chains will take some time to 'burn-in' (see next section) so the *Gelman-Rubin* test is commonly performed on the last half of the points in the chain. Alternatively, we can apply a *Gelman-Rubin* test to a single chain by splitting up the chain and applying the statistic to each part. One can also calculate the correlations between points in the chain as a function of chain length. The distance over which the correlations between parameters drops to $1/e$ gives a measure of the correlation length, which should be shorter than the chain length if convergence is to be achieved.

The number of chains needed for stability and convergence is not set in stone. It is possible to have several short chains or one long chain, as long as either method can pass convergence tests. Ideally, we would run both a very long chain and several shorter ones for the same data set and check that both methods give equally reliable results.

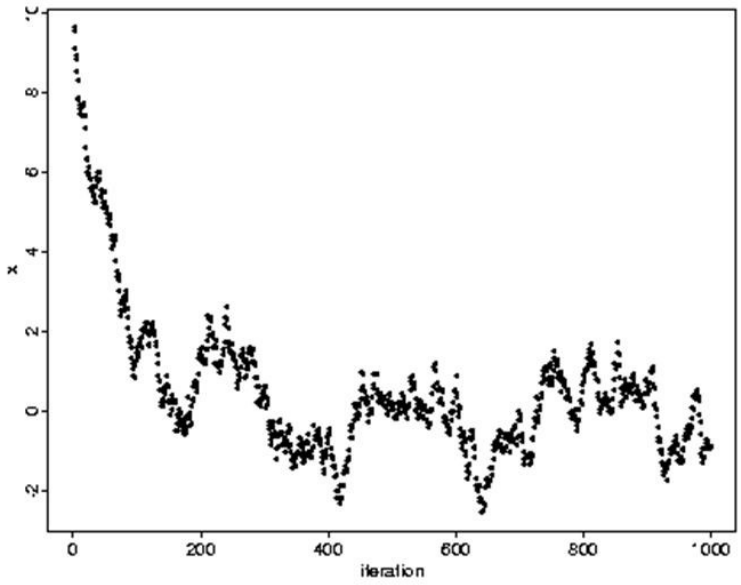

Figure 3.5: Example results from MCMC chain for a parameter x. The burn-in can be clearly seen in the first ~ 200 iterations [144].

3.4.2 Burn-in

MCMC chains typically suffer from a 'burn-in' at the start of the chain during which the points will not trace the posterior distribution well. This is due to the random nature of the starting point; there is a high chance that the chain will start far from the peak of the likelihood and the first n_b points in the chain will be over–representing a region of lower likelihood. Furthermore, these early points will be correlated (and therefore not independent), as the value of one point will strongly affect the value of the next as they converge towards high likelihood values. For this reason, it is common practice to discard these n_b points from our chain(s) to ensure that they do not bias our likelihood analysis. The question remains as to how large n_b should be made to safely ensure no contamination from burn-in. n_b may be as high as half the entire chain length, but often a much shorter burn-in will suffice. Visual inspection of the data can give us strong clues as to how much burn-in needs to be taken into account. For example, in Figure 3.5 the chain begins to obviously converge on a value of x after about 200 iterations. Finding the optimum burn-in length may therefore require some experimentation.

It is also possible to set up an MCMC to have no burn-in. If we know approximately where the peak of the likelihood surface is (either from knowledge of priors from other data sets or from running a short MCMC 'test chain' to determine the location of the peak first) then we can legitimately select a starting point by centering our proposal distribution on the peak and selecting the first point randomly from this [144].

3.4.3 The Proposal Distribution

In our example above we used a simple tophat proposal distribution, and intuitively guessed an appropriate size for it from which to draw our points. However, when testing an MCMC one soon discovers that the acceptance rate of the algorithm is strongly dependent on the choice of proposal distribution and hence there is much room for optimisation. We find that the optimal proposal distribution is one which closely matches the posterior distribution [145, 136]. Why should this be so? Consider a situation in which we are investigating two parameters that happen to be highly correlated. There may be a strong degeneracy along a certain axis in parameter space, such as that between Ω_m and σ_8 as shown in Figure 3.6 (*left*). Drawing a tophat distribution around a point in the chain will make it hard for the MCMC to step away from this point, as the probability that it will attempt to step to an area of lower likelihood is high. If we are able to select a proposal distribution that closely follows the posterior in size, shape and orientation (in this case, a bivariate Gaussian), this makes it much easier to step to a new point with high likelihood, thus increasing the acceptance rate of the algorithm (see Figure 3.6 (*right*)).

The size of our proposal distribution influences both the acceptance rate and the ease with which a chain converges. A very small proposal distribution will only allow the chain to make tiny steps, which means that although the acceptance rate will be high as nearby points have very similar likelihood values, the chain will take many iterations to explore the parameter space. Hence the chain will take a long time to converge. Conversely, a very large proposal distribution will allow the chain to step easily into different regions of the parameter space, but as soon as the chain finds a point of high likelihood stepping away from it will become very difficult. If $\mathcal{L}(\pi_i)$ is relatively high and π_{i+1} is located far from π_i, then $\mathcal{L}(\pi_{i+1})$ will often be much lower than $\mathcal{L}(\pi_i)$ and hence the new point

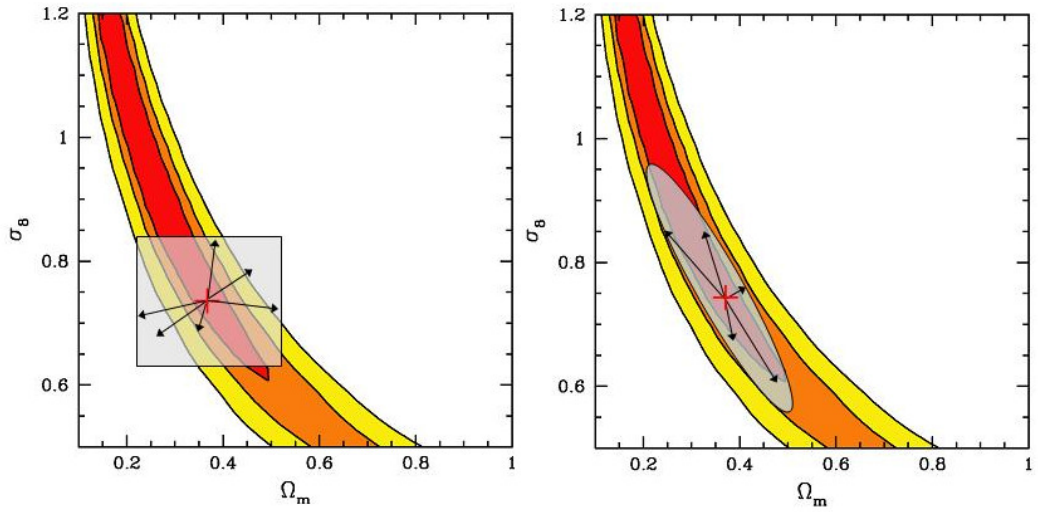

Figure 3.6: *Left*: A tophat proposal distribution gives a poor acceptance rate for the Ω_m–σ_8 degeneracy from weak lensing (Likelihood contours from [124]). *Right*: A bivariate Gaussian distribution results in fewer rejections.

is unlikely to be accepted. Such a chain will therefore also take a long time to run to convergence due to a lowered acceptance ratio. The optimum proposal distribution size is something that must be determined by experimentation. It is worth noting that since the choice of proposal distribution only influences the acceptance rate, it does not bias the posterior in any way as long as the MCMC is convergent.

If we are confident in our choice of priors and know the expected posterior, it is easy to come up with an adequate proposal distribution. Otherwise, we can run a short chain using a tophat distribution and use the posterior from this to produce our improved proposal distribution for a much longer, convergent run. The proposal distribution itself is often optimised in the form of a (multivariate) Gaussian found from principal component analysis (PCA)[4] of the posterior. A Gaussian proposal distribution is often a good first order approximation to the posterior, but is not always appropriate and in some cases exploration of the parameter space may still be very slow.

[4]PCA is the procedure by which a set of correlated observations are converted into a set of linearly uncorrelated variables using an orthonormal transformation. For an explanation of principal component analysis, see for example [146] or §3.6.1.

3.4.4 Priors

As we have thus far taken a Bayesian approach to importance sampling, our investigation of the MCMC algorithm would not be complete without a word about priors. For clarification, the priors chosen on x and y in Figures 3.2-3.4 are simply flat and limited by the edges of the parameter space. As stated earlier, this is a common starting point in parameter estimation if no other information is known. The edges of the parameter space may be constrained by limitations in simulations, or bound a region that encompasses all physical solutions.

Often, we may be attempting to constrain parameters x and y while at the same time marginalising over other 'nuisance' parameters. For example we may be interested only in constraining Ω_m and σ_8, but these depend on the values of other cosmological parameters such as the dimensionless Hubble constant h_0. In this case, there are two possible courses of action. One is to let h_0 vary along with Ω_m and σ_8 and then simply ignore the values of h_0 when plotting our likelihood contours - this is marginalisation as describe in § 3.2.1. The other is to set a prior on the nuisance parameter(s), if possible, from previous data analyses. In this case, we might decide that h_0 has already been well constrained and assume a value of 0.738 (as in [12]). However, there will most likely be very few accepted samples in our chain for which $h_0 = 0.71$ precisely, and it may be better practice to use a narrow prior of $0.730 < h_0 < 0.745$ instead. This is equivalent to assigning a tophat distribution to h_0 over a very small region of parameter space. A more sophisticated approach would be to use the Gaussian error on h_0 of ± 0.024 [12] instead of a top hat. One could use full, non-Gaussian and covariant errors from WMAP on all the parameters being constrained. Constraining nuisance parameters using priors will lead to tighter constraints in our likelihood analysis, however we must be confident in our choice of priors to ensure we are not sacrificing accuracy for false precision.

3.4.5 Caveats

While the number of points needed in a grid-based likelihood analysis scales exponentially with dimension number, chain lengths required for convergence in MCMC scale, at best, linearly with dimension number. This is assuming we have an optimised proposal distribution; in reality the scaling will be somewhat worse.

Despite its advantages over a grid-based approach, MCMC can be very slow to converge for high dimensions. In addition, as we have already stated MCMC is not robust to the choice of proposal distribution and this can further lengthen the time taken to reach convergence.

Another limitation of the MCMC is its difficulty in both probing long tails of a posterior and in dealing with multi-peaked distributions; although the Metropolis-Hastings selection criterion is designed to allow the chain to pass through areas of low likelihood, in practise the MCMC does not sample multiple peaks well. A chain that does not happily sample the full posterior is said to be poorly mixing. One possible solution to a poorly mixing chain is to run multiple chains, each starting in different regions of parameter space in the hope that they will collectively sample all the peaks and regions of the likelihood surface.

Despite its limitations, the Metropolis-Hastings MCMC is a powerful statistical tool if used correctly. For this reason, we chose to construct and implement an MCMC in order to perform the statistical analyses presented in Chapters 4-7. A variety of other methods and algorithms exist, however, in order to overcome some of the limitations of MCMC, and we will briefly mention some of these in the next section.

3.5 Alternative Sampling Methods

In this section we briefly mention some other importance sampling methods that are used in astronomy today, since the MCMC is not the only useful algorithm for finding posterior distributions. We mention some of their strengths, as well as explaining why we choose not to implement them over the MCMC algorithm.

3.5.1 Hamiltonian Monte Carlo

The Hamiltonian Monte Carlo, or Hybrid Monte Carlo (HMC) is a Monte Carlo algorithm that addresses the low efficiency of the MCMC in high dimensions and its low acceptance rate. In this method, each chain position x_i is randomly assigned a momentum u_i, and we define a potential energy

$$U(\boldsymbol{x}) = -\ln\mathcal{P}(\boldsymbol{x}) \tag{3.10}$$

where $\mathcal{P}(\boldsymbol{x})$ is the posterior or target distribution we are attempting to sample from [147]. As before, we approximate the form of $\mathcal{P}(\boldsymbol{x})$ with our proposal distribution, which may be found in advance from, for example, running a short MCMC chain and performing a principal component analysis on the resulting cloud of points as described in section 3.4.3. We can then define the Quantum Mechanical Hamiltonian, $H(\boldsymbol{x}, \boldsymbol{u}) = U(\boldsymbol{x}) + K(\boldsymbol{u})$ where $K(\boldsymbol{u}) = \boldsymbol{u}^T\boldsymbol{u}/2$ is the kinetic energy. This is then used to draw samples from an extended target distribution $\mathcal{P}(\boldsymbol{x}, \boldsymbol{u}) \propto \exp(-H(\boldsymbol{x}, \boldsymbol{u}))$. With our assigned momentum vector, we then follow a trajectory in $(\boldsymbol{x}, \boldsymbol{u})$ phase space, keeping $H(\boldsymbol{x}, \boldsymbol{u})$ constant. The time evolution of the system is governed by the Hamiltonian equations of motion

$$\begin{aligned} \dot{x}_i &= u_i \\ \dot{u}_i &= -\frac{\partial H}{\partial x_i}. \end{aligned} \tag{3.11}$$

In practise the algorithm proceeds by leap–frogging through a series of finite steps in time. One can visualise the likelihood surface of the posterior as a potential well, such that the higher the likelihood value at a given point, the deeper the potential as given by Eqn. (3.10). The time evolution of the algorithm takes it through a region of constant $H(\boldsymbol{x}, \boldsymbol{u})$ within that surface until a new point in the chain is reached. Then a new, random momentum is assigned and the chain proceeds on another path through phase space to ensure the chain does not get trapped in an ellipse of constant $H(\boldsymbol{x}, \boldsymbol{u})$. In essence, the HMC is an MCMC with a different proposal distribution and a phase space of $2n_{\text{par}}$ dimensions instead of n_{par} due to the presence of the momentum term. To obtain the posterior $\mathcal{P}(\boldsymbol{x})$ after a chain has been run, one simply marginalises over the momentum coordinates in $\mathcal{P}(\boldsymbol{x}, \boldsymbol{u})$ to obtain our desired real–space posterior $\mathcal{P}(\boldsymbol{x})$.

The HMC has the advantage that because the total energy of the system $H(\boldsymbol{x}, \boldsymbol{u})$ is kept almost constant then for two points in the chain, the likelihoods $\mathcal{L}(x_i)$ and $\mathcal{L}(x_{i+1})$ will be almost identical and the acceptance ratio will be close to one. The energy is not perfectly conserved from point to point, however, because of the inexact, numerical nature of the leap–frogging behaviour. Conserving $H(\boldsymbol{x}, \boldsymbol{u})$ also depends on having a full knowledge of the posterior, which is what we are trying to measure with an approximation (the prior distribution). Because the HMC can take relatively large step sizes in parameter space with a

high acceptance rate, it can sample the space effectively and without doubling back on itself as the MCMC does due to its random walk nature. Finally, the efficiency of the HMC also scales well with dimension number, so it may be well suited to multi–dimensional analysis. Although HMC is a promising algorithm for importance sampling, the MCMC is simpler to implement and optimise. Additionally, the parameter space we explore is only 3-dimensional and the posterior distribution of the parameters we investigate single-peaked (see §3.6), meaning that MCMC should perform well.

3.5.2 Population Monte Carlo

Population Monte Carlo (PMC) is an adaptive sampling technique that is growing in popularity in the field of cosmology. In PMC, one chooses a sequence of samples from a sequence of importance functions that progressively approximate the posterior [148, 149]. An initial importance function q^1 is chosen, and a sample of points $x_1^1 ... x_n^1$ are generated within this importance function in the parameter space. Next, corresponding importance weights $w_1^1 ... w_n^1$ are assigned to the points based on their relative likelihood values. Finally, the process starts again with the importance function q^2 updated based on the previous weighted sample until some stopping criterion is reached. PMC has the advantage that is is easily parallelisable because many separate importance functions can be generated independently at the same time and their outputs combined. There is a publicly-available version of PMC called CosmoPMC[5] [150] that is in use by the CFHTLenS collaboration. PMC is a very new sampling method that was not widely available when this thesis was started, so we do not implement it. In addition, the choice of initial importance function(s) is critical in producing convergent, well behaved chains, and research is still being conducted to determine the optimal choice of importance functions for different cosmological problems [150].

3.5.3 Nested Sampling

Nested sampling is a relatively new algorithm [136] and proceeds by generating an array of n points $(\pi_1, \pi_2, ..., \pi_n)$ in the parameter space. The likelihood of each

[5]Martin Kilbinger, `http://www2.iap.fr/users/kilbinge/CosmoPMC/`.

point is calculated, and the point with the lowest likelihood $\mathcal{L}(\pi_i)$ is discarded. Then, a new point π_i' is generated by taking a random step from one of the other points and is accepted iff $\mathcal{L}(\pi_i')$ is equal to or greater than the likelihood of the discarded point. This ensures that over repeated iterations, the likelihood contours from the n points move progressively inwards towards the peak of the likelihood (hence the name). One calculates the cumulative evidence after each iteration until some stopping criterion has been reached, then the posterior is estimated by weighting each point according to its likelihood width (the width is determined by the distance from a point's nearest neighbour). Nested sampling is a robust technique that copes well with 'difficult' posteriors, such as multi–peaked or highly correlated distributions. It can be slow, however, especially with high dimensions, so we choose not to implement this method in our analysis [151].

3.5.4 Simulated Annealing

'Simulated Annealing' is so named because of the parallel between the way in which a metal cools and freezes into a minimum energy E crystalline structure (the annealing process) and the search for a minimum in a system such as our parameter space. By analogy, $\mathcal{L} = \exp(-E)$. For a given energy and 'temperature', T, a perturbation is added and the change in energy calculated. If the change in energy is negative, the new configuration is accepted; if it is positive, it is accepted with a probability given by the Boltzmann factor $\exp(-\mathrm{d}E/T)$. This process is repeated for multiple sampling points, then the temperature is reduced and the process repeated until $T = 0$ is reached and the system has the minimum possible energy (and hence the highest possible likelihood). Simulated annealing has the advantage that it is good at avoiding become trapped on local minima (likelihood peaks) due to the selection criterion. However, simulated annealing cannot cope with log-likelihood surfaces that are not concave in shape as the evidence cannot be calculated (an example of a concave surface is a Gaussian) [136], meaning that one must know *a priori* about the shape of the likelihood surface. For this reason, we do not use simulated annealing in our analysis.

3.6 Implementation

A key part of this thesis was the construction of an MCMC algorithm. As we shall see in Chapters 6 & 7, our analysis involved the simultaneous measurement of three parameters at once - Ω_m, σ_8, and the amplitude of a galaxy alignment model used to model the intrinsic alignment signal, A, which will be described in Chapter 5. We choose to constrain the cosmological parameters Ω_m and σ_8 because cosmic shear measurements are able to more precisely constrain this parameter combination than any other, and it is likely that most future cosmic shear analyses will investigate this joint constraint. The intrinsic alignment model parameter is fitted in order to investigate the constraining power a CFHTLenS-like survey will have on the amplitude of the intrinsic alignment signal, which is an important step towards successfully accounting for this systematic. Since it is only weakly dependent on cosmology, the inclusion of A does not significantly degrade the constraints on Ω_m and σ_8. Preliminary analysis of the CFHTLenS-like simulations described in Chapters 4 & 5 showed that there is very little constraining power present on the Hubble parameter h_0 or the dark energy equation of state parameters, w_0 and w_a. Including these parameters in the likelihood analysis would therefore add very little cosmological information at the cost of longer run times to convergence and a loss of constraining power on Ω_m and σ_8. For this reason, we do not include these parameters in our likelihood analysis. To constrain the chosen parameters, an MCMC algorithm was written, implemented and optimised. We perform this analysis on simulated cosmic shear fields binned into different combinations of tomographic and angular bin numbers.

To obtain χ^2 values for different models the input data were compared with theoretical predictions of the shear correlation functions from the publicly-available `NICAEA` code[6]. We apply a tophat prior to limit the parameter space. The dimensions of the parameter space were set by the ranges over which `NICAEA` had been tested and was known to predict the lensing signal accurately. A significant amount of the posterior lies outside the observable parameter space for some parameters, meaning that the MCMC attempts to step outside the parameter space sometimes. This was dealt with by forcing the chain to reject any points that lie outside the parameter space. These rejections were not included in calculations of the acceptance ratio, as they are not failures of the Metropolis-

[6]Martin Kilbinger, `http://www2.iap.fr/users/kilbinge/nicaea/`.

Hastings criterion. Simply rejecting points at the edges of the parameter space in this way had no noticeable effect on the posterior distribution, meaning that this treatment did not bias the analysis.

When analysing simulated data in Chapter 6, we negated the need for a burn-in by selecting the first point in the chain from a proposal distribution centred around the input cosmology of the simulations, which is known exactly.

3.6.1 Proposal Distribution

Our proposal distribution is determined using PCA. First, a short chain of $\sim 10^4$ points was run using a tophat proposal distribution for each tomographic bin combination n_z (in this thesis we use $1 \leq n_z \leq 15$, see Chapter 4). From this, the shape of the posterior is apparent as can be seen in the top part of Fig. 3.7 for the (Ω_m, σ_8) plane. Next, we select all points within $\chi^2 = \chi^2_{\rm min} + \Delta\chi^2$. Since the likelihood contours are often degenerate, particularly in the case of the (Ω_m, σ_8) plane, the value of $\Delta\chi^2$ is chosen such that as many points are selected as possible while the likelihood surface in the region looks approximately Gaussian (close to the peak of the likelihood, the degeneracy is less obvious and an array of points drawn from this region will tend to look more elliptical than those drawn from the whole likelihood surface). From experimentation, for $n_z = 1$ we choose an optimal (unreduced) $\Delta\chi^2$ of 0.25. The choice of $\Delta\chi^2$ is based on the shape of the posterior in the (Ω_m, σ_8) plane because the joint constraint on these parameters is the tightest and also the most degenerate and hence non-Gaussian. Therefore our choice of $\Delta\chi^2$ is limited by the behaviour of these parameters. The bottom part of Fig. 3.7 shows the points from the trial chain that lie above $\chi^2 = \chi^2_{\rm min} + \Delta\chi^2$. Due to the strong degeneracy, even with such a small number of points the likelihood surface is already beginning to look non-Gaussian and we cannot raise the value of $\Delta\chi^2$ further. We focus on the Ω_m-σ_8 contour when optimising the proposal distribution because these parameters are correlated and hence more non-Gaussian than the Ω_m-A or σ_8-A contours, and the optimal $\Delta\chi^2$ value will be most sensitive to this combination.

Performing PCA on the array of points produces a rotation matrix that maps the multivariate Gaussian onto the axes of the parameter space, and a set of variances σ_i^2 that characterise it. This constitutes the Gaussianised proposal distribution used throughout; to implement it, for each dimension i of the

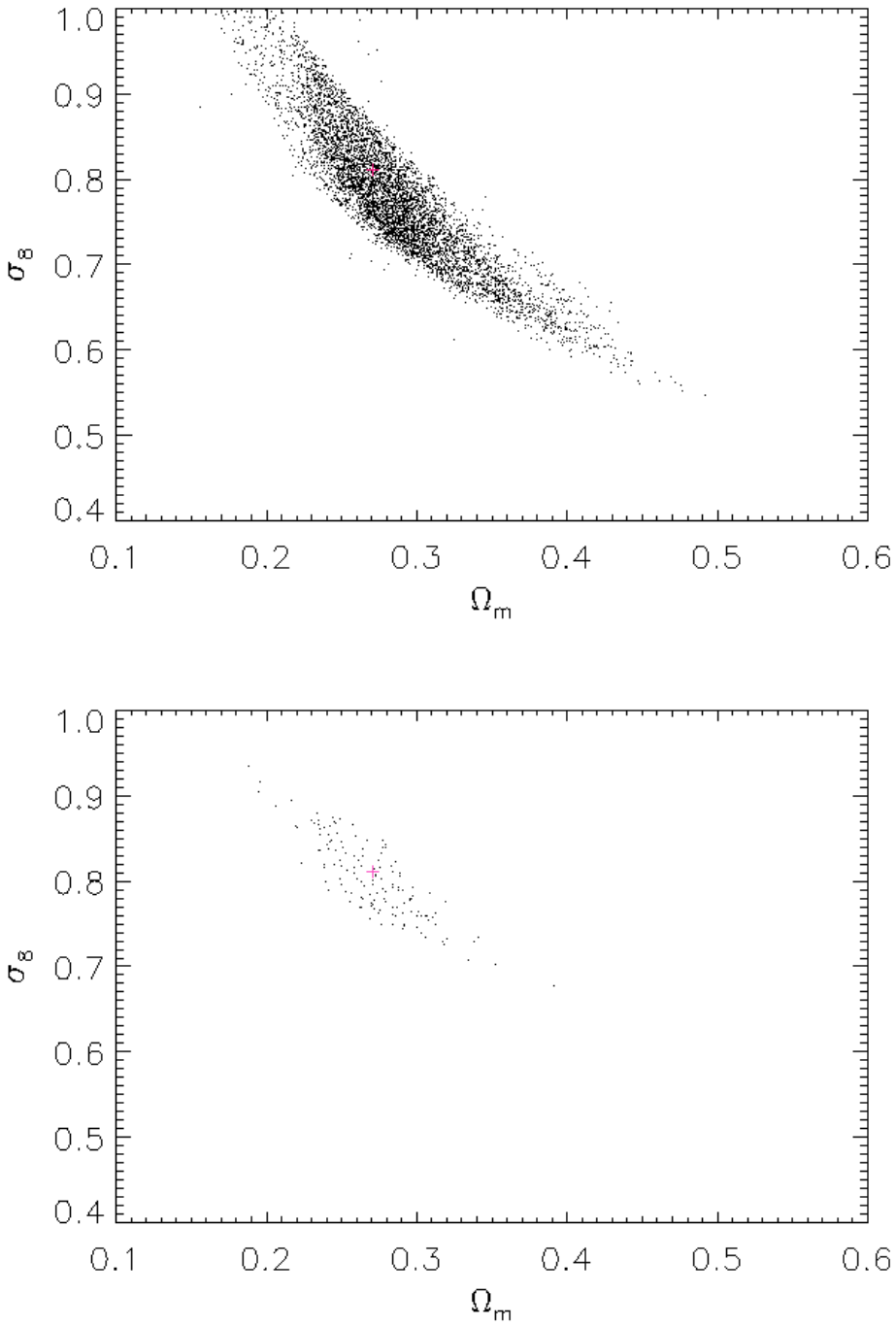

Figure 3.7: *Top*: Positions of accepted points from a trial MCMC chain (with $n_z = 1$ and $n_\theta = 6$) of length $\sim 10^4$ points in the (Ω_m, σ_8) plane. *Bottom*: Approximately 300 points lying within $\chi^2 = \chi^2_{\rm min} + 0.25$. The input cosmology is shown with a pink cross in both plots.

parameter space we select a random number from a Gaussian with variance σ_i^2, and multiply this array of random numbers by the rotation matrix $\mathbf{R}$. The vector $\boldsymbol{\pi}^{\rm map}$ has length i and contains the i-dimensional position of a random point selected from the multivariate Gaussian. Thus for a point n in the chain where $n = 1...n_{\rm chain}$ we can define a new vector $\boldsymbol{\pi}$,

$$[\boldsymbol{\pi}]_n = [\mathbf{R}\,\boldsymbol{\pi}^{\rm map}]_n\,, \tag{3.12}$$

which represents the position of the random point in the true parameter space. In Fig. 3.8 (*top*) we examine the proposal distribution determined from PCA of the trial chain in Fig. 3.7 by plotting $n_{\rm chain} = 5$ x 10^4 random numbers with zero mean and variances σ_i^2. Only two dimensions are shown; (π_1, π_2) maps to (Ω_m, σ_8) once they are rotated by $\mathbf{R}$, as shown in Fig. 3.8 (*bottom*). The rotated array of points is a covariant Gaussian that resembles the distribution of points below $\chi^2 = \chi^2_{\rm min} + 0.25$ in Fig. 3.7 (*bottom*) in both orientation and scale, as expected.

The size of the proposal distribution can be altered by scaling the variances, affecting the acceptance ratio; we find a strong dependence of the acceptance rate on the number of tomographic bins being utilised due to the fact that the credibility intervals shrink for higher bin tomography. This was dealt with by running trial chains for every $n_{\rm z}$ and producing unique proposal distributions and variances for each, ensuring the acceptance rate is approximately constant with $n_{\rm z}$. We choose not to repeat this for every potential combination of angular bins in the analysis as the size of the contours is much less sensitive to changes in the angular bin number than to changes in tomographic bin number, an effect we investigate in Chapter 6.

3.6.2 Comparison with Population Monte Carlo

The MCMC algorithm developed for use in this thesis was tested on simulated data from the CFHTLenS clone simulations[7]. For a 2-bin tomographic analysis spanning $0.0 < z < 2.0$ and ΛCDM cosmology, we perform a likelihood analysis and obtain credibility intervals for Ω_m and σ_8. We compare our result to the result from using a PMC algorithm in Fig. 3.9. Our results show good agreement

[7]The CFHTLenS clone is a set of ray-tracing simulations used by the CFHTLenS collaboration for calibrating lensing analyses which we describe in detail in §4.2.

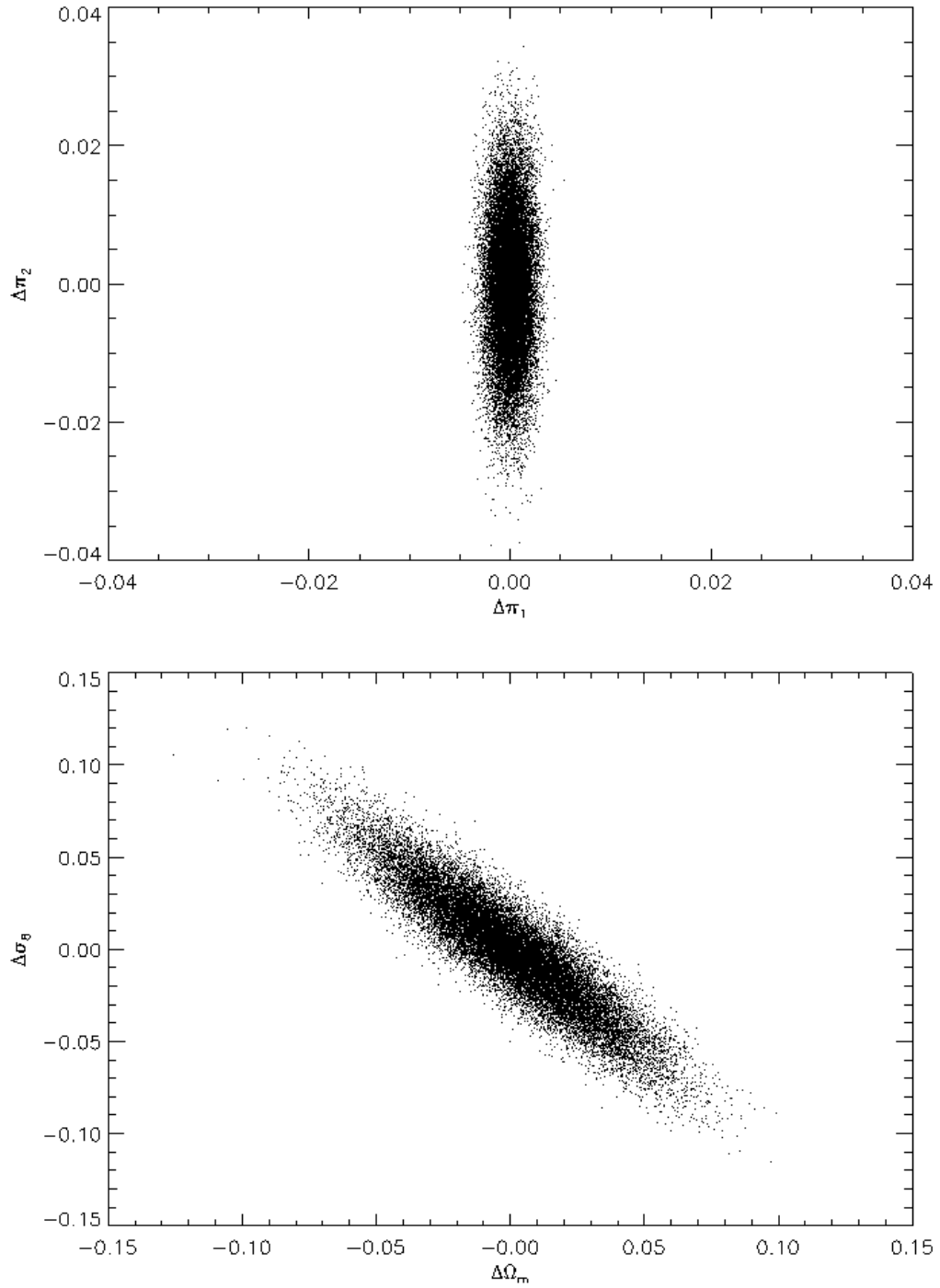

Figure 3.8: *Top*: 5×10^4 random numbers generated with zero mean and variances σ_i^2 determined from PCA performed on the chain from Fig. 3.7. Plotted are the relative positions of the points in 2 dimensions which we denote π_1 and π_2. *Bottom*: The same array of points after each point has been rotated by $\mathbf{R}$. The points now correspond to relative positions in the (Ω_m, σ_8) plane.

with PMC (although the contours do not exactly match due to the difference in the sampling methods used). Both the 68.3% and 95.4% confidence intervals are of a similar width and height from both analyses, although the PMC constraints are slightly broader. This may be due to one or both chains showing a lack of convergence, or differences in the method used to draw the confidence intervals (e.g. the points have been binned on grids of slightly different resolution, see §3.4). The contours from both analyses are centred around the same region, indicating that neither analysis is biased. h_0 was also allowed to vary but we find little constraining power is present in the data set for this parameter, in agreement with the PMC analysis, so we do not show it here.

3.7 Conclusion

In this chapter, we have introduced the concept of importance sampling and illustrated its significance in Bayesian model selection and parameter estimation. Importance sampling methods offer significant advantages over grid–based analyses due primarily to their vastly improved computing times. By describing the individual steps that make up one of the simplest and most widely-used importance sampling methods, the Metropolis-Hastings MCMC, we have highlighted the importance of one's choice of priors and proposal distribution and how it applies to the work presented in this thesis. We have introduced a range of other sampling methods used in cosmology, and explained our choice of the MCMC in light of these. We have detailed the construction of our MCMC algorithm for use in this thesis. We have described the steps taken to optimise the proposal distribution, leading to substantial gains in efficiency and computing time in preparation for our analysis in Chapters 4-7. With a well-tested sampling algorithm in place, we are prepared to analyse both simulated and real cosmic shear data efficiently and accurately.

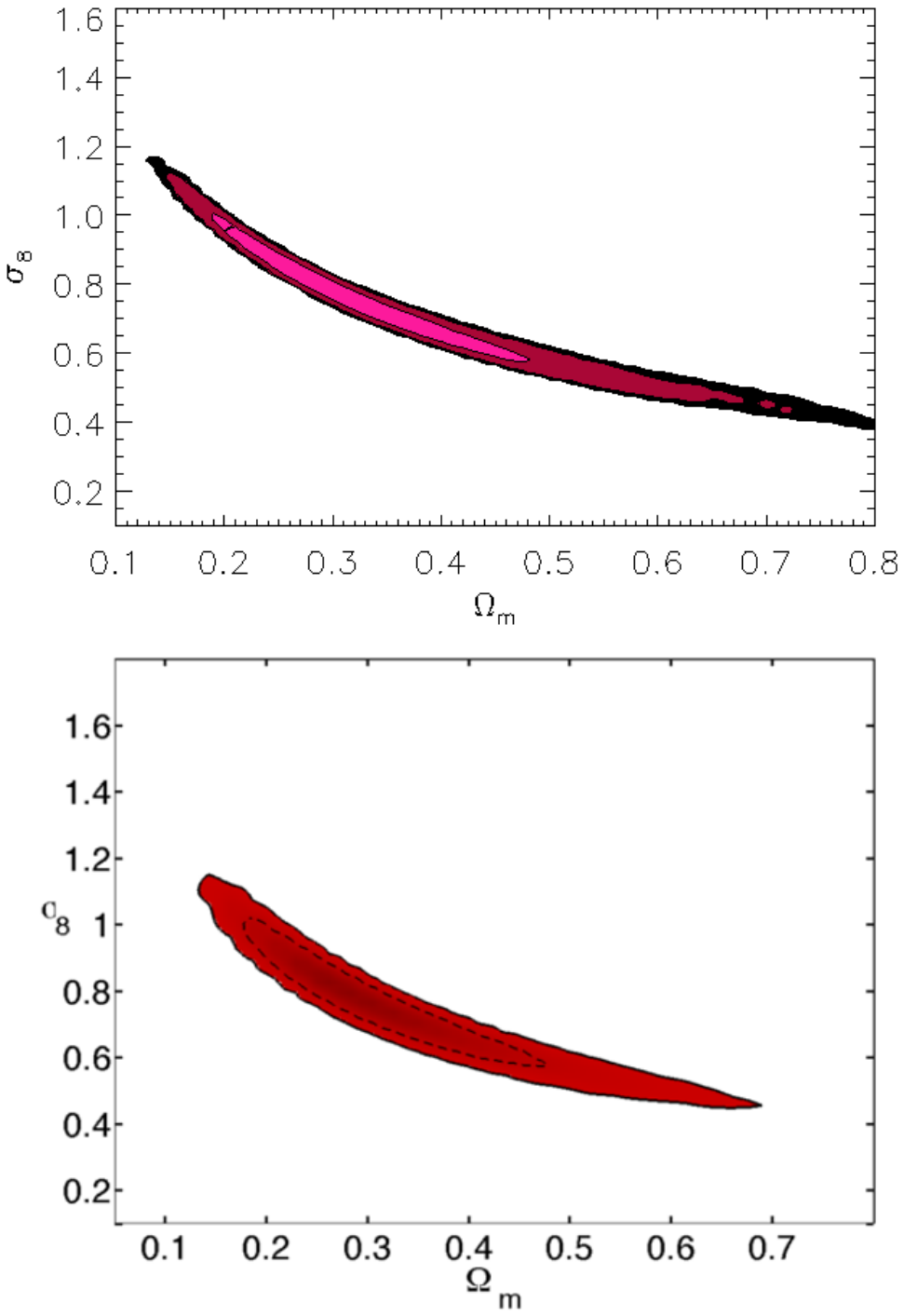

Figure 3.9: *Top*: 68.3%, 95.4% and 99.0% likelihood contours from the MCMC algorithm for a 2-bin tomographic analysis of simulated shear data from the CFHTLenS clone simulations. *Bottom*: 68.3% and 95.4% likelihood contours from PMC (courtesy of Fergus Simpson, IfA, University of Edinburgh).

Chapter 4

Shear Covariance Matrices for Tomography

4.1 Introduction

Weak lensing tomography was first investigated by [126] who utilised power spectra to show that two or three bin tomography can recover most of the cosmological information in a data set. The optimisation of tomographic analysis of cosmic shear was investigated by [152] who found that parameter uncertainties are reduced by a factor of up to 10 by binning the data into 4 redshift bins, with the dark energy density parameter Ω_Λ gaining the most. This is due to the fact that dark energy influences the expansion rate and the growth rate over time, so retaining redshift information allows us to constrain its behaviour at different epochs. Tomographic techniques have been successfully applied to real data utilising up to 6 tomographic bins and have been shown to produce tighter parameter constraints than 2D analyses [153, 154, 155, 156]. We postpone a more in-depth review of the most recent tomographic lensing results until Chapter 5, where we discuss the authors' treatment of intrinsic alignments and photo-zs and compare their analysis techniques to the analysis presented in this thesis.

Different redshift bins will be correlated with each other, as the light from all bins will pass through some common material. Thus to use tomographic information most effectively, it is essential to measure the cross–correlations between redshift bins as well as the shear correlations within each bin [85]. The number and size of redshift bins needed for optimal tomography is limited by

several factors. The first is shot noise — once the number of galaxies in a bin falls to the level of the noise, depending on the statistic used there will be no further gain from subdivision once the shot noise is reached. Additionally, redshift bins cannot be narrower than the typical error on the photometric redshift estimates of the survey objects. [126] found that the use of even just two redshift bins means that much of the statistical information contained in the galaxy redshifts is recovered and the errors on cosmological parameter estimates are improved by up to a factor of twenty, which shows how powerful a tomographic approach can be. This improvement is only possible, however, in the absence of IAs, and [157] found that when IAs are included in tomographic analysis, the number of redshift bins required to recover the full information available must be at least doubled. We discuss and investigate the effect of IAs in Chapters 5 & 6.

Uncertainties in the redshifts of survey galaxies can be a source of systematic error in tomographic lensing analysis. It is not normally possible to obtain highly accurate spectroscopic redshift measurements for all galaxies in large surveys. Instead, easy-to-obtain photometric redshifts are estimated for all survey galaxies, and these may be either calibrated to a training set of spectroscopic redshifts or fitted to a template-fitting algorithm to correct for bias and scatter in the photo-z distribution [158]. Photo-z measurements have some scatter and are also subject to 'catastrophic outliers' whereby the redshift of a galaxy can be seriously misestimated due to degeneracies in the fit to spectroscopic redshifts (for a review, see e.g. [159]). The presence of these types of errors in the photo-z distribution, particularly if poorly accounted for, degrades parameter estimates from weak lensing tomography [160, 161]. These uncertainties may arise because the spectroscopic sample used to calibrate photo-z errors is incomplete and unrepresentative, or because of inaccuracies in photo-z template algorithms.

4.1.1 The covariance matrix of the two-point shear correlation function

Often, the limiting factor that determines the number of tomographic bins possible in a lensing analysis is the behaviour of the covariance of the data. This is because the nonlinear nature of the density field on small scales tends to correlate modes that would otherwise be independent [162, 163]. This effect manifests itself through statistics such as the shear power spectrum becoming

highly correlated, meaning that the amplitude of the signal at one angular scale depends on the amplitude of the signal at all other angular scales. Since lensing data are typically highly correlated over the scales of interest, it is necessary to construct a covariance matrix of the data to take these correlations into account in a likelihood analysis. To perform a likelihood analysis, we require an estimate of the inverse of the covariance matrix. Failure to account for correlations through the off-diagonal terms of the covariance matrix and its inverse means that we will appear to have more constraining power than is actually present, which can lead to spuriously precise parameter constraints and biases in best-fit parameter estimates.

When we split the data into tomographic bins, we introduce additional correlations into the data vector. Not only will the signal at different angular scales be correlated, but the signal at different redshifts will be correlated too. This means we must be very careful when implementing a high number of tomographic bins to ensure we have a well-behaved, non-singular and hence invertible estimate of the covariance matrix with which to perform a likelihood analysis.

[116] derive analytical expressions for the covariance matrices of several weak lensing estimators for contiguous surveys with known mean redshift that can in principle be calculated directly from the data. However, these expressions assume gaussianity of the signal and underestimate the power on small scales. An alternative, easy-to-implement approach is to measure an estimate of the covariance from multiple realisations of simulated data [152, 164]. This approach has the advantages of naturally accounting for cross-bin terms and is independent of survey configuration, and hence more easily applicable to real data. It is also possible to account for the full non-Gaussian covariance on small scales using simulations. Ideally, one generates as many realisations of simulated lensing data as computationally viable, using inputs (such as cosmological parameters and redshift distribution) that are as closely matched to the real data as possible to reduce the chance of biasing the covariance. Using multiple simulated lines of sight results in a more stable estimate of the covariance than using the real data, as we can often simulate a much larger area than the data covers, reducing statistical uncertainties. This is the approach we use in this thesis.

The cosmic shear covariance can be decomposed into two parts: a statistical

shot noise part (due to the uncertainty in galaxy shape measurements), and a sampling variance part (due to the uncertainty in a data set that comes from only sampling a small fraction of the Universe). The sampling variance contribution is significant only on large scales; on small scales the shot noise term dominates. For the shear correlation covariance, the statistical noise contribution comes mainly from the intrinsic galaxy ellipticity distribution and is a consequence of the fact that we cannot separate the cosmic shear and intrinsic ellipticity components of an individual galaxy's observed shape. We investigate the effect of galaxy shape noise on the covariance in Chapter 5; in this chapter we concern ourselves with the sampling variance term of the covariance matrix.

In this chapter, we construct estimators of the shear correlation function covariance from tomographic weak lensing simulations. We investigate the effect of varying the number of realisations of the simulations and the number of tomographic bins on the stability and reliability of the covariance matrix. We determine how many tomographic bins will be obtainable for a CFHTLenS-like survey based on the covariance of the shear signal, and attempt to implement and improve a fitting formula with the aim of increasing the number of viable tomographic bins. This work is important because upcoming lensing surveys are larger than ever before (see Chapter 2), and statistical errors on cosmic shear measurements will be very small. Improper treatment of the covariance matrix can result in imprecise and biased cosmological parameter estimates. We do not want our estimator of the covariance to become a substantial source of systematic error, so we attempt to construct as reliable an estimator as possible. The results of this chapter will have a direct bearing on the analysis presented in Chapter 6, in which we attempt to optimise a tomographic analysis of realistic simulated data.

Several statistics exist for measuring the two-point statistical information contained in the lensing signal. These include the shear correlation functions, the shear dispersion in circular apertures [165], and the aperture mass dispersion [166]. The shear correlation functions $\xi_{\pm}$, discussed in §2.3.4, are the easiest of these two-point statistics to measure directly from data as they are not sensitive to gaps or masked regions [116] and do not require knowledge of shear correlations at small scales that cannot easily be measured from data, hence we utilise these statistics in our analysis. In this thesis, we choose to focus solely on ξ_{+}. This

is motivated by the fact that the ξ_+ signal is stronger than ξ_- (for example, for ΛCDM cosmology and a CFHTLenS-like redshift distribution, ξ_+ is ~ 8 times larger than ξ_- at an angular scale of $2'$), but the noise is the same. Additionally, including ξ_- would double the size of the data vector, and as we will explain below in §4.1.2, this would double the number of realisations of simulations needed to construct an unbiased inverse of the covariance matrix. Since we only have a limited number or realisations of simulations available to us (see §4.2), this is a significant issue.

We denote the covariance matrix estimated directly from the mock data $\hat{\mathbf{C}}$. Under the assumption that the data vector consists of statistically independent components p_i with Gaussian noise[1], the maximum-likelihood estimator for the covariance matrix inverse as described in [167] is

$$\hat{\mathbf{C}}_{ij} = \frac{1}{N-1} \sum_{k=1}^{N} \left(p_i^{(k)} - \mu_i\right) \left(p_j^{(k)} - \mu_j\right) . \tag{4.1}$$

where $\hat{\mathbf{C}}$ runs over i and j (the size of the data vector), $\boldsymbol{\mu}$ is the mean estimated from the data and N is the number of realisations of the data vector. The diagonal of the covariance matrix is simply equal to the variance of the data points and the off-diagonals are the covariances between data points. The more correlated the data vector, the larger in amplitude the off-diagonal terms of the covariance matrix will be.

The size of the covariance matrix does not grow linearly with the number of tomographic bins n_z; the cross correlations between redshift bins are also taken into account. The total number of cross and auto correlation between tomographic bins, n_corr, is given by

$$n_\mathrm{corr} = \frac{n_\mathrm{z}(n_\mathrm{z}+1)}{2}. \tag{4.2}$$

The consequence of this is that as the number of tomographic bins increases, the number of data points increases as a triangular number. For example, for $n_\mathrm{z} = 1$, $n_\mathrm{corr} = 1$, but for $n_\mathrm{z} = 10$, $n_\mathrm{corr} = 55$. The data vector and corresponding covariance matrix for a 10-bin tomographic analysis will be 55 times larger than that of a 1-bin analysis. Therefore we will need 55 times as many realisations of

[1]The assumption of Gaussianity will be seen to be valid in §5.4 due to the inclusion of a dominant shot noise term in the covariance.

the covariance matrix to ensure it is non-singular (see next section).

The structure of the covariance matrix follows the structure of the data vector. This is illustrated in Fig. 4.1 for 2-bin tomography for a data vector measured on 5 angular scales ($n_\theta = 5$). $\hat{\xi}_+(\theta, z)$ is constructed so that the first n_θ values are equal to the values of $\hat{\xi}_+$ arranged in ascending order of angular scale for the lowest redshift bin auto-correlation (the 1-1 correlation). The next n_θ values correspond to the cross correlation between the lowest and highest tomographic bins (1-2). The final n_θ values represent the highest bin auto-correlation (2-2). The total data vector is $n_{\rm corr}$ x $n_\theta = 15$ elements in length. The covariance matrix is then just the two-dimensional extension of this. The highlighted squares represent the terms

$$\begin{aligned} \hat{\mathbf{C}}_{3-2} &= \frac{1}{N-1}\sum_{k=1}^{N}\left(\hat{\xi}^{(k)}_{+,3} - \bar{\xi}_{+,3}\right)\left(\hat{\xi}^{(k)}_{+,2} - \bar{\xi}_{+,2}\right) \\ \hat{\mathbf{C}}_{2-8} &= \frac{1}{N-1}\sum_{k=1}^{N}\left(\hat{\xi}^{(k)}_{+,2} - \bar{\xi}_{+,2}\right)\left(\hat{\xi}^{(k)}_{+,8} - \bar{\xi}_{+,8}\right), \end{aligned} \tag{4.3}$$

where $\bar{\xi}_+$ represents the mean value of $\hat{\xi}_+$ from all N realisations of the simulations. This structure holds for all values of $n_{\rm z}$ and n_θ. The covariance matrix is symmetric about the diagonal such that $\hat{\mathbf{C}}_{ij} = \hat{\mathbf{C}}_{ji}$.

The importance of the covariance matrix in cosmological analysis becomes clear through its central role in likelihood analysis. The covariance matrix enters the likelihood calculation through the χ^2 goodness-of-fit statistic, which can be used to obtain best-fit cosmological parameters $\boldsymbol{\pi}_0$ and confidence regions as described in Chapter 3. For correlated data, the definition of χ^2 for a set of model parameters $\boldsymbol{\pi}$ in Eqn. 3.6 is modified to include the covariance matrix inverse,

$$\chi^2(\pi) = \sum_{ij}\left(\bar{\xi}_+ - \xi_+(\pi)\right)_i \hat{\mathbf{C}}^{-1}_{ij}\left(\bar{\xi}_+ - \xi_+(\pi)\right)_j. \tag{4.4}$$

4.1.2 The Anderson Correction

[168] and [169] (hereafter H07) found that if the length of the data vector, p, exceeds the number of realisations N of simulated data used in calculating $\hat{\mathbf{C}}$ then the covariance matrix becomes singular and the likelihood cannot be calculated. This constraint represents a fundamental limit on the number of tomographic bins

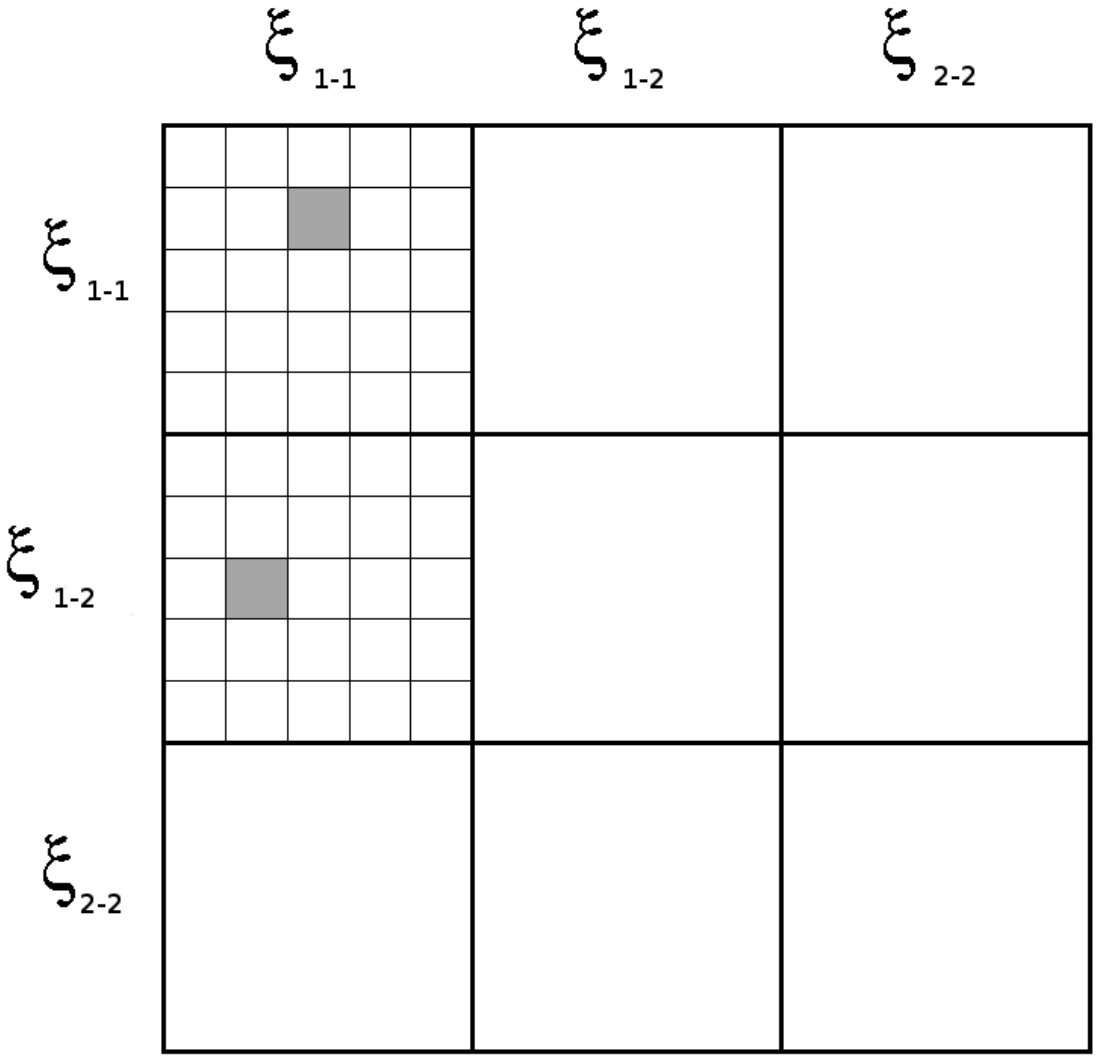

Figure 4.1: A schematic showing the substructure of the covariance matrix for $n_z = 2$, $n_\theta = 5$. Figure modified from Fig. 2 in [152].

one can construct from a given data set and it means that in order to reliably use the covariance matrix, we must ensure that as many realisations of the simulations as possible are employed to construct it. Generating large, realistic simulations is computationally costly, so the number of tomographic bins obtainable may be limited by computing power and our subsequent ability to generate sufficient realisations.

In general, noise in the covariance matrix will cause a bias in the inverted matrix even when the covariance matrix estimator $\hat{\mathbf{C}}$ is unbiased itself. For this reason, one should pseudo-invert the covariance matrix through e.g. singular value decomposition (see §5.4.2) rather than attempt to invert it directly. H07 prove analytically that $\hat{\mathbf{C}}$ is always singular when $p > N$, and that even when using SVD a bias in the inverted matrix is present that is highly dependent on the covariance model chosen, making the inverse unreliable. This is true even for the case of $p < N$, meaning that any likelihood analysis involving the inverse estimated directly from data or simulations must correct for this bias. In this chapter we will refer to three different measures of the inverse of the covariance matrix of ξ_+; we list their definitions here for clarity:

$\boldsymbol{\Sigma}^{-1}$ = 'True' inverse of the covariance matrix of the data, found analytically.

$\hat{\mathbf{C}}_*^{-1}$ = Biased estimator of the inverse of $\hat{\mathbf{C}}$, where $\hat{\mathbf{C}}$ is estimated directly from data or simulations.

$\hat{\mathbf{C}}^{-1}$ = (Relatively) unbiased estimator of the inverse, after applying a correction factor to $\hat{\mathbf{C}}_*^{-1}$.

$\boldsymbol{\Sigma}^{-1}$ often cannot be calculated directly from real surveys, so we must approximate it. In the case of $p < N$, matrix inversion of the covariance matrix estimator (estimated from multiple lines of sight as described in Eqn. 4.1) results in an estimator of $\boldsymbol{\Sigma}^{-1}$,

$$\hat{\mathbf{C}}_*^{-1} \approx \boldsymbol{\Sigma}^{-1}. \tag{4.5}$$

However $\hat{\mathbf{C}}_*^{-1}$ is biased and hence not equal to the unbiased estimator of the inverse $\hat{\mathbf{C}}^{-1}$. An unbiased estimator for $\boldsymbol{\Sigma}^{-1}$ is thus given in Anderson 2003 [170] as

$$\hat{\mathbf{C}}^{-1} = \frac{N-p-2}{N-1}\hat{\mathbf{C}}_*^{-1} \text{ for } p < N-2 \tag{4.6}$$

when the mean is estimated from the data. The ratio $\frac{N-p-2}{N-1}$ will hereafter be referred to as the Anderson correction, and is necessary because the use of a biased estimator $\hat{\mathbf{C}}_*^{-1}$ will produce serious underestimates in the sizes of parameter credibility intervals, particularly in the limit $p/N \to 1$. The Anderson correction produces credibility intervals that more honestly reflect the precision attainable with the data. The correction is derived based on the idealised case of Gaussian shot noise and statistically independent data vectors.

4.2 Cosmic Shear Simulations

We use two types of lensing simulations to measure covariances of $\hat{\xi}_+$: N-body simulations and Gaussian shear fields. We describe the characteristics of the two simulations in this section, explaining why we use them both and how we measure $\hat{\xi}_+$ from them.

4.2.1 N-body simulations

One of the most accurate and reliable ways to simulate lensing data is using N-body simulations [163, 171, 172]. N-body ray-tracing lensing simulations contain collisionless dark matter particles and rely on the assumption that the evolution of large scale structure is mainly dependent on the dark matter distribution [162]. In such a simulation, one calculates photon geodesics through large scale structure in three dimensions, starting at the observer and moving out towards the source positions. The cumulative deformation of the light ray along each trajectory is then calculated to produce pixelated shear fields at the desire redshifts. Once shear fields have been generated, they may be populated with galaxies. Each galaxy will have a complex shear assigned to it (and perhaps other parameters) based on its position in the field.

One of the largest N-body dark matter structure simulation carried out to date is the Millennium simulation [173], which was run in 2005[2]. It has 2160^3 particles

[2]At the time of running, Millennium was the largest simulation of its kind. It has recently been superseded in 2011 by the Horizon simulation, which has box sizes up to $1081.5h^{-1}$ Mpc on a side [174].

in a cubic region $500h^{-1}$ Mpc on a side. When the Millennium simulation was generated, σ_8 was not well constrained from observations and subsequently the simulations were generated with an input $\sigma_8 = 0.9$. We now know σ_8 is closer to ~ 0.8 [26]. Since the covariance matrix of ξ_+ is cosmology-dependent (see §4.5), we need to use simulations that more closely match the real Universe to ensure we do not introduce biases. For this reason we do not use the Millennium simulation in our analysis.

In this thesis, we utilise the N-body simulations of [163], which are in use by the CFHTLenS collaboration as a tool to calibrate various lensing analyses. They consist of 184 independent lines of sight each with an area of 12.38 square degrees (10.80 square degrees once CFHTLenS masks are taken into consideration), with 1024^2 pixels and a pixel size of $0.21'$. We refer to this suite of simulations as the 'clone' simulations, in which the positions of galaxies placed within the clones map onto the galaxy positions measured from the CFHTLenS fields. 16 CFHTLenS subfields (each of approximate area of 0.7 square degrees) are tiled onto each clone field and the overlap discarded until all CFHTLenS subfields have been used. The process is then repeated until all clone fields are populated. This method means that galaxy positions in the CFHTLenS are used multiple times (in different lensing fields), hence the name 'clone simulations'. Due to this method of population, the clustering of galaxies in the clone is not correlated with the underlying density field as it would be in the real Universe (i.e. we would expect to find more galaxies in overdense regions). However, this approach is statistically no worse than placing galaxies randomly in the shear field, and source clustering of this nature is expected be only a percent-level effect [175] and hence should not bias the recovered shear signal significantly. The simulations have a particle count of 1024^3 and a box size of between $147.0h^{-1}$ Mpc and $231.1h^{-1}$Mpc on side depending on the redshift of the lens. The clones have a flat ΛCDM cosmology with input parameters described in Table 4.1. For comparison, the most recent cosmological constraints from WMAP7 are also shown. The clone cosmology matches to WMAP7 to within 2σ. They incorporate secondary effects such as source clustering, the effect of which on lensing measurements is not yet fully understood.

The clone simulations have a small volume and high particle density, and are thus well-positioned to probe small scales where nonlinear effects are important.

Table 4.1: Key cosmological parameters from WMAP and the clone simulations.

Parameter	WMAP 7yr mean	Clone input
Ω_m	0.266 ± 0.029	0.279
h_0	0.704 ± 0.025	0.701
σ_8	$0.811^{+0.030}_{-0.031}$	0.817
w_0^*	-0.93 ± 0.12	-1.0
w_1^*	$-0.38^{+0.66}_{-0.65}$	0.0
Ω_b	0.0455 ± 0.0028	0.0460
n_s	0.967 ± 0.014	0.960

WMAP constraints assume flat ΛCDM. * indicates parameters for which the combined results from multiple cosmological probes are shown (BAO+H_0+$D_{\Delta t}$+SN, see Table 4 of [26]) instead of just the WMAP 7 year mean result.

They can accurately reproduce the lensing signal at sub-arcminute scales, which is essential if we are to match the resolution of modern surveys and simulate realistic estimators of $\hat{\xi}_+$ and $\hat{\mathbf{C}}$. However, dense simulations like this are computationally expensive to create, so the number of lines of sight available to us is limited. For this reason, we choose to cut each clone simulation into 4 on a 2 x 2 grid; this gives us 736 quasi-independent lines of sight each with an area of 2.70 square degrees (1.64 degrees on a side). This enables us to effectively quadruple the number of measurements of $\hat{\xi}_+$ we are able to make and will allow us to use more tomographic bins in a likelihood analysis. The drawback of this is that on large angular scales $\hat{\xi}_+$ will begin to become unreliable as edge effects come into play much sooner. However, most of the cosmological information contained within $\hat{\xi}_+$ for a CFHTLenS-like survey is on small scales where the signal is strongest ($\lesssim 10'$) so this approach is unlikely to effect our analysis. An important caveat to our use of the clones is that they do not incorporate a model for intrinsic galaxy alignments, whose effects must be taken into account when analysing data as discussed in Chapter 2.

4.2.2 Gaussian simulations

N-body simulations are complex and time-consuming to produce. The lack of intrinsic alignments and limited number of lines of sight available from the clones are large drawbacks to their use in a realistic cosmic shear likelihood analysis.

For this reason, we also consider a second set of simulations that addresses these issues. An alternative to using full nonlinear N-body simulations is to simulate Gaussian lensing fields which do not account for the non-Gaussianity of the lensing signal on small scales.

In this thesis we use the Gaussian simulation code of Brown & Battye [176]. These simulations contain a model for IAs which will be discussed in Chapter 5. To construct shear fields corresponding to a given cosmology, the simulation code requires knowledge of the lensing (and IA) power spectra for every redshift auto- and cross-correlation prediction for that cosmology. The power spectra are used to generate flat, pixelated shear maps. More detail on the construction of shear and IA fields from these simulations is given in §5.2.1.

Estimating ξ_+ from Gaussian shear fields will cause us to underestimate the error on small scales ($\lesssim 10'$), as nonlinear evolution of the density field in the real Universe leads to stronger correlations between modes and hence higher variances and covariances [162]. This is one of the main limitations of using Gaussian simulations, and one we attempt to address in §4.4. However, generating Gaussian fields is quick and straightforward, allowing us to generate many more realisations of the shear field. This is important for reducing statistical errors and producing a stable covariance matrix, and may allow us to use a higher number of tomographic bins in cosmological analysis.

We modify the simulation code so that Gaussian shear fields may be generated with any desired input cosmology and redshift distribution; we use `NICAEA` (see §3.6) to determine input power spectra that match the clone simulations in both these aspects. `NICAEA` is well-tested and requires the user to define the input cosmological parameters and redshift distribution. It then calculates theoretical shear correlation functions and power spectra for these inputs, for any number of tomographic bins. All the shear fields for a given tomographic bin combination must be generated simultaneously, incorporating the cross power spectra, to ensure the tomographic fields are correlated correctly. We simulate 800 semi-independent lines of sight of 18.20 square degrees. These are constructed from 50 truly independent simulations of 291.3 square degrees split into a 4x4 grid. We utilise only 736 of these lines of sight in order to match the number of realisations that are available from the clone simulations, however we are able to generate more realisations as required. We set the resolution of the simulations to the

maximum value to reduce the effect of resolution issues, giving us fields with 2048^2 pixels and a pixel size of $0.5'$ and $l_{\rm max} = 3000$.

4.2.3 Constructing tomographic galaxy catalogues

Input tomographic redshift distribution

We describe how we produce mock galaxy catalogues for a number of tomographic bins in the range $1 \leq n_{\rm z} \leq 15$ from both the N-body and Gaussian simulations. The N-body simulations consist of galaxy catalogues that mimic the CFHTLenS in position and redshift distribution; we must apply tomographic cuts to these catalogues to obtain galaxy catalogues corresponding to different redshift bins. We must also determine the redshift distribution for each tomographic bin so that we can produce Gaussian simulations that match the clones in input as closely as possible.

The redshift distribution of the clone is established from the galaxies in all CFHTLenS fields with a photo-z cut of $0.1 < z_p < 1.3$. [177] argue that this is a safe range over which photometric redshift estimates are reliable. Each galaxy in the catalogue has a probability distribution function (PDF) for its true redshift associated with it found from the Bayesian Photometric Redshift (BPZ) Estimation code of [178]. We used a catalogue manipulation tool called `ldac`[3] to extract the relevant columns from the entire CFHTLenS galaxy catalogues, namely the photometric redshift estimate and PDF of each galaxy, stored in histogram form from $0 < z < 3$.

The sum of the PDFs was used to produce a redshift distribution that is thought to closely match the true redshift distribution of the sources. To produce realistic tomographic redshift bins from the clone catalogues, we cut the master galaxy catalogue by z_p into $n_{\rm z}$ different bins, ensuring that there are the same number of galaxies in each bin. We then find the sum of the PDFs for each $n_{\rm z}$ bin. This can be done for any number of tomographic bins and some examples are shown in Fig 4.2. Importantly, the distributions from the sum of the PDFs overlap, as expected if there is contamination between bins due to photo-z uncertainties. This means that the II and GI alignment signals (introduced

[3]Leiden Data Analysis Centre, see `http://marvinweb.astro.uni-bonn.de/data_products/THELIWWW/index.html`.

in Chapter 2) will be non-negligible in the cross and auto correlations between redshift bins respectively, an effect that will be discussed in Chapter 5.

Next, input power spectra were generated using the clone simulation cosmology and the derived PDF redshift distributions for $1 \leq n_{\rm z} \leq 15$. Gaussian shear fields were then produced from these power spectra, matching the clones (and CFHTLenS) in input cosmology and redshift distribution.

Input galaxy distribution

We now have tomographic galaxy catalogues for the clone simulations that contain a list of galaxy positions and ellipticities, with the positions determined from the CFHTLenS data catalogues and the ellipticities determined from the simulated 3D shear field of the clone. We need to populate the Gaussian shear fields with galaxies to produce a similar set of galaxy catalogues.

We must determine how many galaxies to populate the Gaussian fields with, and also how they will be distributed. The N-body galaxy catalogues were constructed to match the galaxy density of the CFHTLenS. The effective galaxy density was determined directly from the CFHTLenS catalogues using each galaxy's assigned `Lensfit` weight. The galaxy density can be calculated by simply summing the number of survey galaxies and dividing by total survey area, but by using the `Lensfit` weights w_i we can get an estimate of the more representative *effective* galaxy number density, $n_g(\text{eff})$,

$$n_g(\text{eff}) = \frac{\left(\sum_{i=1}^{N_{\rm gal}} w_i\right)^2}{\sum_{i=1}^{N_{\rm gal}} w_i^2}. \tag{4.7}$$

From all CFHTLenS fields, we find the mean effective galaxy density to be 13.34 per sq. arcmin (total over all tomographic slices). We therefore populate the Gaussian shear fields with this galaxy density.

We assign galaxy positions in the Gaussian fields randomly. An improvement to this approach would be to populate the fields with the CFHTLenS positions, as in the clones, or to add some degree of source clustering as this would better mimic real galaxy behaviour. However, intrinsic source clustering is only a percent-

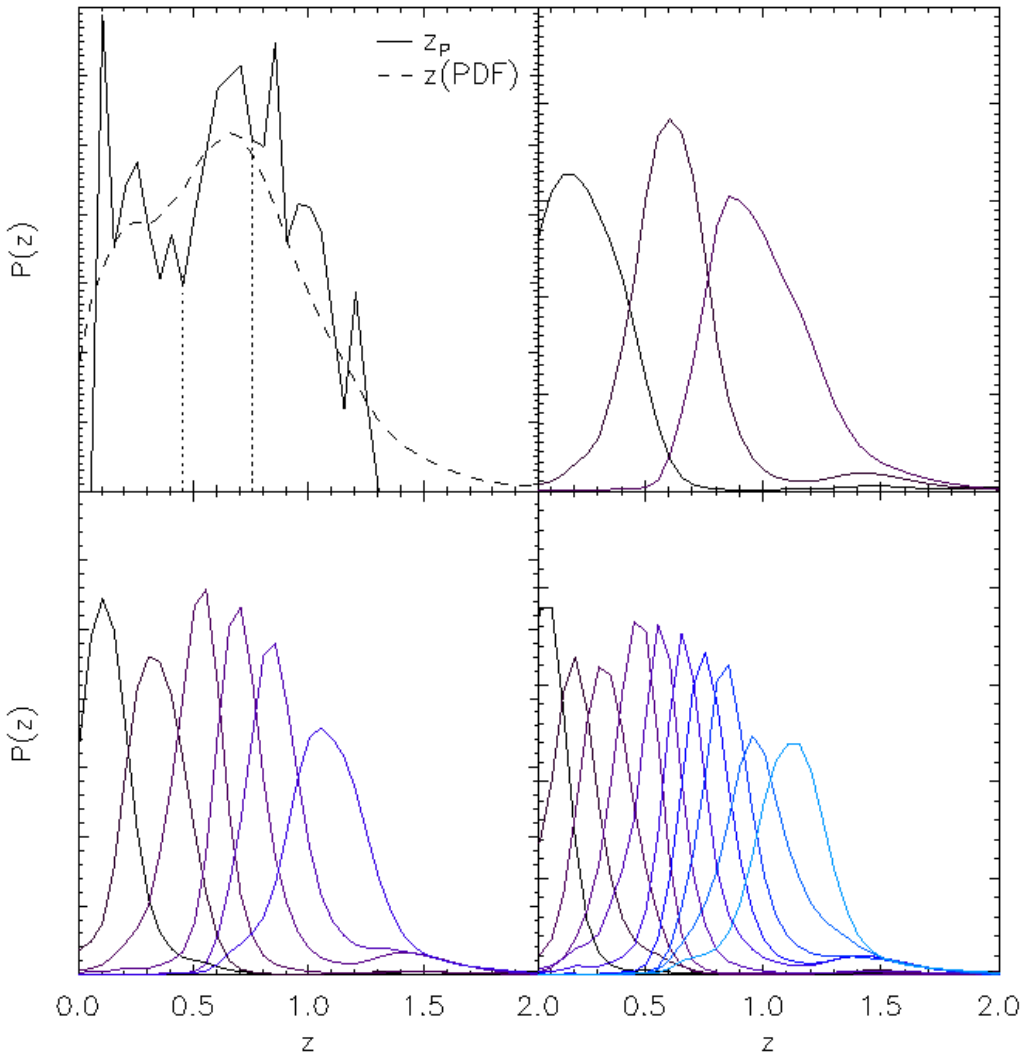

Figure 4.2: *Top left*: The total redshift probability distribution from CFHTLenS galaxies for the raw photo-zs ($0.1 < z_p < 1.3$, solid line) and from the sum of the PDFs (z(PDF), dashed line) (redshift information from [177]). Vertical dotted lines show where cuts in z_p were performed to produce tomographic redshift distributions for the case of $n_z = 3$. *Top right*: The resulting z(PDF) distributions for the $n_z = 3$ cut. *Bottom*: z(PDF) for $n_z = 6$ and 10. $P(z)$ is scaled to sum to 1, as only the relative size and shape of the redshift bins within a tomographic combination affects the shear correlation function. The y-axes have been rescaled to better show the shape of the redshift distributions in each plot.

level effect [175] and the signal is dominated by other effects such as pixelisation (particularly on small scales; see §4.3).

We assign ellipticities to the galaxies in the Gaussian simulations by interpolating between the shear values at the surrounding pixel positions in the shear fields. This ensures that the galaxy ellipticities are accurate tracers of the underlying shear field only and the effect of pixelation on galaxy ellipticity is minimised. We do not include galaxy shape noise in any of the simulated galaxy catalogues at this stage as we are only interested in the behaviour of the pure shear covariance matrix in this chapter. Incorporating a realistic noise contribution in simulations is necessary for a full cosmological likelihood analysis to be useful and applicable to real data, and for this reason we measure and apply a shape noise component in Chapter 5.

4.3 Estimating ξ_+

We follow the same procedure to measure $\hat{\xi}_+$ from the clone and Gaussian simulations. The publicly-available `ATHENA` code[4] is used to calculate the shear correlation functions from the mock galaxy catalogues. This is a tree code to calculate second-order correlation functions over a range of angular scales. `ATHENA` has been robustly tested for CFHTLenS and shown to be sufficient for the level of accuracy required for this thesis. `ATHENA` correlates nodes of galaxies (a node is defined as all galaxies closer than some open angle threshold, or OATH) with each other over a range of angular scales, multiplying the mean, weighted ellipticites of both nodes for a given binned barycentre distance. We set OATH=0.04 radians; maximum node size is equal to the OATH multiplied by the separation between nodes. This means that OATH=0.0 is equivalent to brute-force correlation of every galaxy with every other galaxy, which is more accurate but slow. We calculate the shear correlation function over a large number (~ 1000) of angular scales between $\theta_{min} = 1'$ and $\theta_{max} = 85'$, which we then re-bin more coarsely to get n_θ different angular bins. We set $3 \leq n_\theta \leq 15$ and place the angular bins evenly in logarithmic space to give a more uniform signal-to-noise ratio, which is higher on small scales. The range for n_θ is limited by the resulting size of the covariance matrix; as the limiting case of $p/N = 1$ is reached at progressively

[4]Martin Kilbinger, `http://www2.iap.fr/users/kilbinge/athena/`.

lower numbers of tomographic bins, for $n_\theta \gtrsim 15$ we can use a maximum of only 9 tomographic bins (see Table 4.2).

The redshift of individual galaxies is not explicitly needed (and indeed is not known for the Gaussian catalogues as the redshift distribution is already incorporated into the power spectra used to generate the simulations). The auto correlation functions are those found from galaxies within the same redshift bin and the cross correlation functions are found from correlating galaxies in different bins.

$\hat{\xi}_+$ is averaged over all 736 lines of sight for each tomographic and angular bin combination for both suites of simulated data; we denote these estimators $\hat{\xi}_{+,\mathrm{NG}}$ and $\hat{\xi}_{+,\mathrm{G}}$ as determined from the clone and Gaussian simulations respectively. In principle, calculating $\hat{\xi}_+$ over a large number of fine bins will recover more of the cosmological information and produce tighter parameter constraints than more coarse sampling, but the data points are highly correlated and increasing n_θ above some optimum may have the unwanted effect of producing a covariance matrix with strong off-diagonal terms that requires many lines of sight to estimate with sufficient accuracy. It also increases the size of our data vector and hence the number of lines of sight needed to calculate the covariance matrix.

Figs. 4.3 and 4.4 show the correlation functions obtained from the N-body and Gaussian simulations respectively for a representative sample of $n_\mathrm{z} = 1$, 2, 3 & 6, and $n_\theta = 10$. The input theoretical shear correlation functions were obtained using `NICAEA`, which calculates the expected value of the shear correlation function for given angular scales based on the input cosmology and redshift distribution. $\hat{\xi}_{+,\mathrm{NG}}$ begins to fall below the expected value on large scales at all redshifts, starting at $\sim 30'$. This is due to the finite box size of the N-body simulations. Since N-body simulations necessarily simulate a finite volume, density perturbations larger than the box length will be suppressed as they cannot be simulated, leading to a suppression in the recovered shear correlation function on scales considerably smaller than the box length [179].

The input cosmology of the Gaussian simulations are reproduced well by the mock galaxy catalogues except at small angular scales of $\theta \lesssim 5'$. This bias on small scales is present in all n_z combinations and is most likely due to the finite resolution of the simulations (we do not see this effect in the clone simulations because of their higher resolution). This issue affects our likelihood analysis in

Chapter 6; we describe therein how it is dealt with. In general, $\hat{\xi}_+(\theta)$ is behaving as expected at all n_z; rising with the mean redshift of the slices being correlated. This occurs because there is a greater volume of lensing material along our line of sight to the most distant redshift bins. On all but the smallest scales ($\theta \lesssim 5'$), the Gaussian galaxy catalogues are a better match to the theory than the clone.

4.4 Estimating the covariance matrix of $\hat{\xi}_+$

We denote the covariance matrix estimator for the clone catalogues as $\hat{\mathbf{C}}_{\rm NG}$ and for the Gaussian catalogues as $\hat{\mathbf{C}}_{\rm G}$. In this section we produce estimators of the covariance according to Eqn. 4.1 for the entire desired range of combinations $1 \leq n_z \leq 15$ and $3 \leq n_\theta \leq 15$, and compare $\hat{\mathbf{C}}_{\rm NG}$ and $\hat{\mathbf{C}}_{\rm G}$, investigating the optimal treatment of the covariance matrix to ensure our estimator is as unbiased and reliable as possible for cosmological analysis.

4.4.1 Area Scaling

We wish to construct a covariance matrix that is of use in the analysis of CFHTLenS data which spans a total survey area of 154 square degrees. However, the clone and Gaussian fields we use to construct our covariance estimators are much smaller: 2.7 and 18.2 square degrees respectively. Clearly, the raw covariance estimators $\hat{\mathbf{C}}_{\rm NG}$ and $\hat{\mathbf{C}}_{\rm G}$ cannot be directly compared since the larger Gaussian fields will contain more galaxies, meaning our estimate of $\bar{\xi}_{+,\rm G}$ is more reliable and the error on the measurement will be smaller.

Eqns. 32-34 in [116] show how the size of every component of the covariance matrix is inversely proportional to survey area. It is therefore possible to scale the entire covariance matrix to match the expected uncertainties in surveys of different sizes. We therefore scale all covariance matrix elements by the ratio of the area of the simulated fields to that of the CFHTLenS survey, that is by a factor of $2.7/154 = 0.017$ for $\hat{\mathbf{C}}_{\rm NG}$ and $18.2/154 = 0.12$ for $\hat{\mathbf{C}}_{\rm G}$. This produces uncertainties in $\hat{\xi}_+$ and hence parameter constraints that are in line with those expected from CFHTLenS and allowing us to directly compare the covariance matrices from the non-Gaussian and Gaussian simulations.

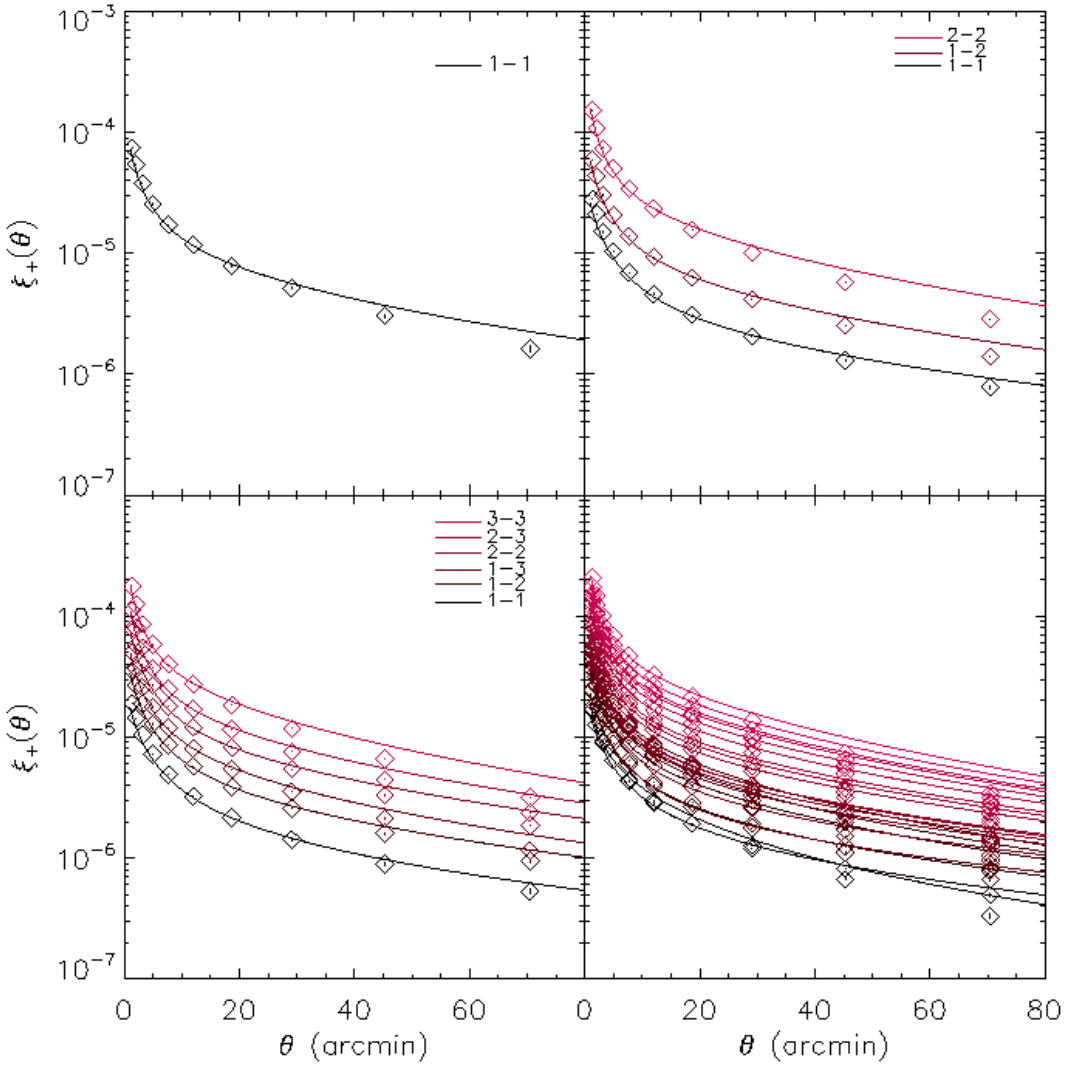

Figure 4.3: $\hat{\xi}_{+,\mathrm{NG}}(\theta)$ estimated from the CFHTLenS clone simulations for $n_\mathrm{z} = 1$ (*top left*), 2 (*top right*), 3 (*bottom left*) and 6 (*bottom right*) with $n_\theta = 10$ (points), plotted with the input theoretical correlation function $\xi_+^\mathrm{th}(\theta)$ (lines). Error bars correspond to the standard deviation of all 736 lines of sight, scaled to the area of the CFHTLenS. The legend in the top right of each plot corresponds to the redshift bins being correlated. We omit the legend for $n_\mathrm{z} = 6$ to avoid crowding the image.

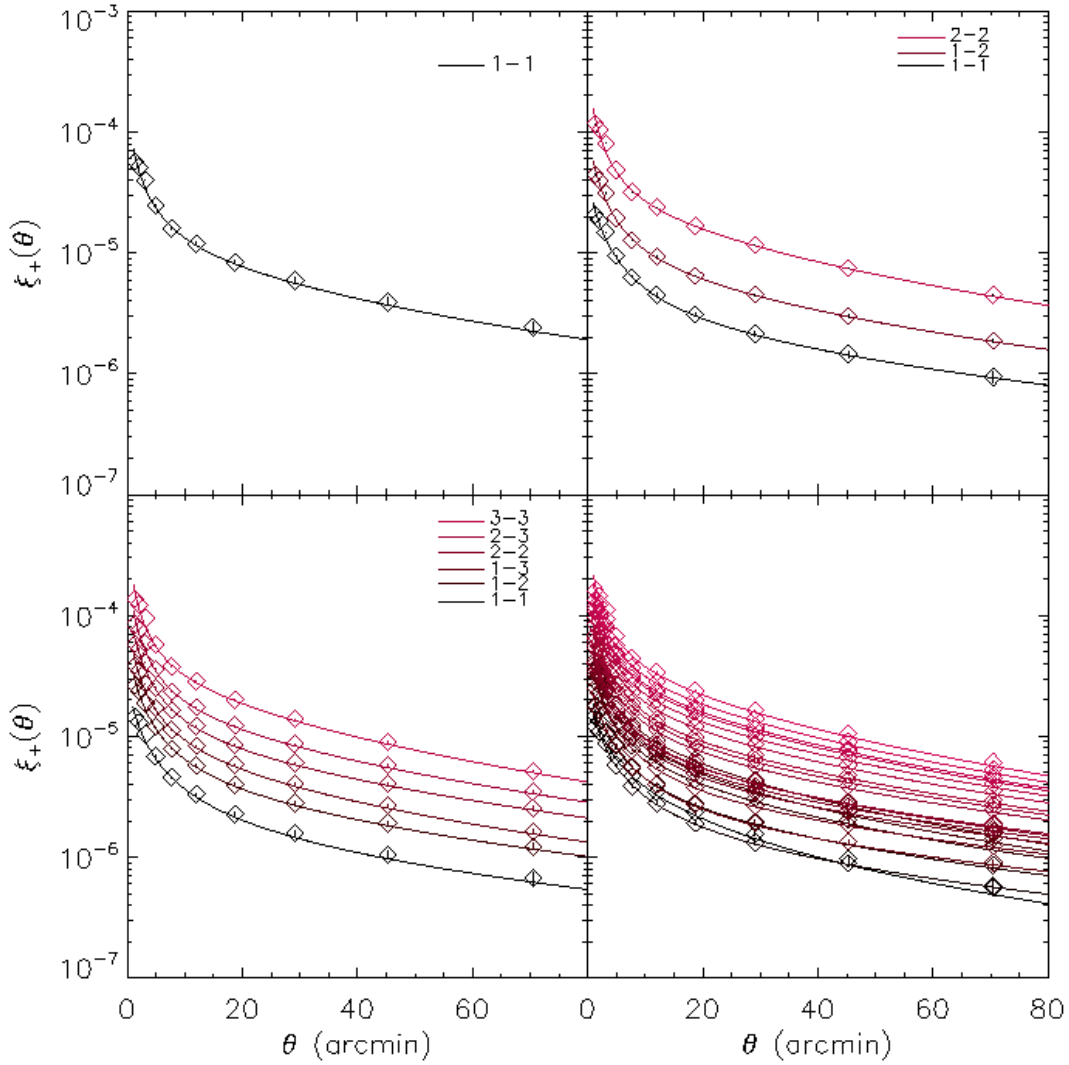

Figure 4.4: As in Fig. 4.3 but with $\hat{\xi}_{+,\mathrm{G}}(\theta)$ calculated from the Gaussian simulations.

We plot the relative amplitude of the terms in $\hat{\mathbf{C}}_{\rm NG}$ and $\hat{\mathbf{C}}_{\rm G}$ for $n_{\rm z} = 2$, $n_\theta = 10$ in Fig. 4.5 and Fig. 4.6 respectively, to illustrate the general behaviour of the covariance matrix. As expected, the small scale terms are higher in amplitude than the large scale terms since there are fewer galaxy pairs to correlate at small θ. The covariance of high redshift correlations is larger than for low redshift slices, because at high z, a given angular scale represents a larger real-space distance and the sampling variance variance will be higher. In general, $\hat{\mathbf{C}}_{\rm NG}$ is higher in amplitude than $\hat{\mathbf{C}}_{\rm G}$, due to the contribution to the covariance from nonlinearity (particularly at small scales). These plots illustrate the strong degree of off-diagonality present in the covariance matrix and therefore the importance of taking the covariance matrix into account in cosmological analyses.

4.4.2 Stability of the Covariance Matrix

A large number of lines of sight (and corresponding measurements of $\hat{\xi}_+$) are needed to ensure our covariance matrix estimator is a reliable representation of the 'true' covariance matrix $\mathbf{\Sigma}$. If too few realisations of the simulations are used, $\hat{\mathbf{C}}$ may be noisy or biased. We investigate the behaviour of the covariance matrix found from an increasing number of lines of sight for $n_{\rm z} = 1$ and $n_\theta = 10$. We construct 1000 estimates of $\hat{\mathbf{C}}_{\rm NG}$ for each N and $n_{\rm z}$ by bootstrapping from all 736 lines of sight[5]. We then plot the mean of the trace of $\hat{\mathbf{C}}_{\rm NG}$ as a function of N in Fig. 4.7. We also plot the standard deviation from the 1000 estimates. We plot only the results for $n_{\rm z} = 1$ and $n_{\rm z} = 10$ as we find that the mean and variance is tomography-invariant. The mean and variance of the trace drop by over an order of magnitude by $N \sim 100$, after which there is comparatively little improvement in the stability of the covariance with increasing N. By $N \sim 650$, the trace of the covariance is within $\sim 10\%$ of the value for $N = 736$. This plot shows that the number of lines of sight needed to construct an accurate estimate of the covariance is solely a function of the sample variance (and hence survey area and galaxy density) and is independent of how the data are binned, although there is no clear cut-off above which we can be sure we have enough

[5]Bootstrapping is a process used to generate multiple equivalent realisations from a single set of data. In this case, we select N random lines of sight from the full 736, and measure their covariance. We repeat this to get 1000 measurements of $\hat{\mathbf{C}}_{\rm NG}$ for a given N. These 1000 measurements will not be independent as a given line of sight will contribute to multiple measures of the covariance, so one must choose carefully when it is appropriate to bootstrap.

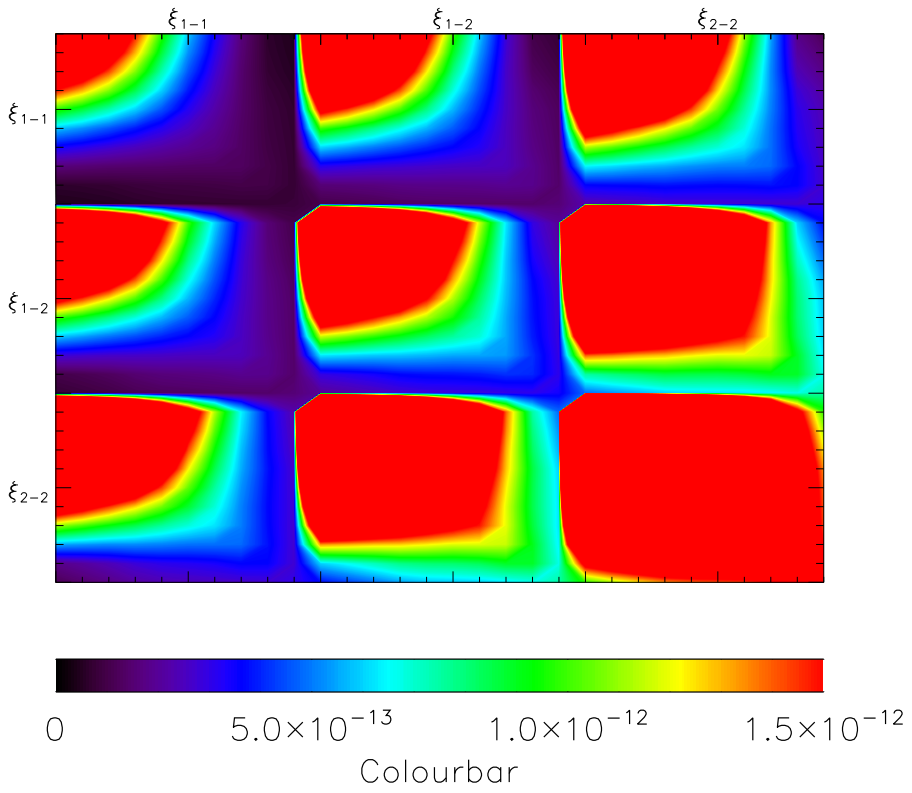

Figure 4.5: Contour plot of $\hat{\mathbf{C}}_{\rm NG}$ with $n_{\rm z} = 2$, $n_\theta = 10$. Contour levels indicated by the colour bar are arbitrary and chosen to best highlight the relative amplitude of different terms in the covariance matrix. Each subsection corresponds to a redshift cross or auto correlation and θ increases from top left to bottom right in each subsection (as in Fig. 4.1). Red = highest amplitude terms, black = lowest amplitude.

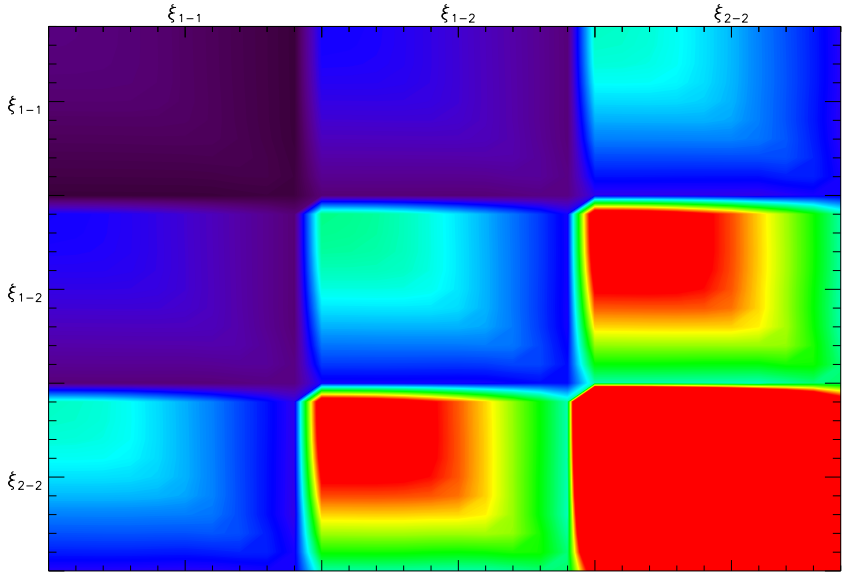

Figure 4.6: As above but for $\hat{\mathbf{C}}_{\rm G}$ with $n_{\rm z} = 2$, $n_\theta = 10$. Contour levels match those used in Fig. 4.5.

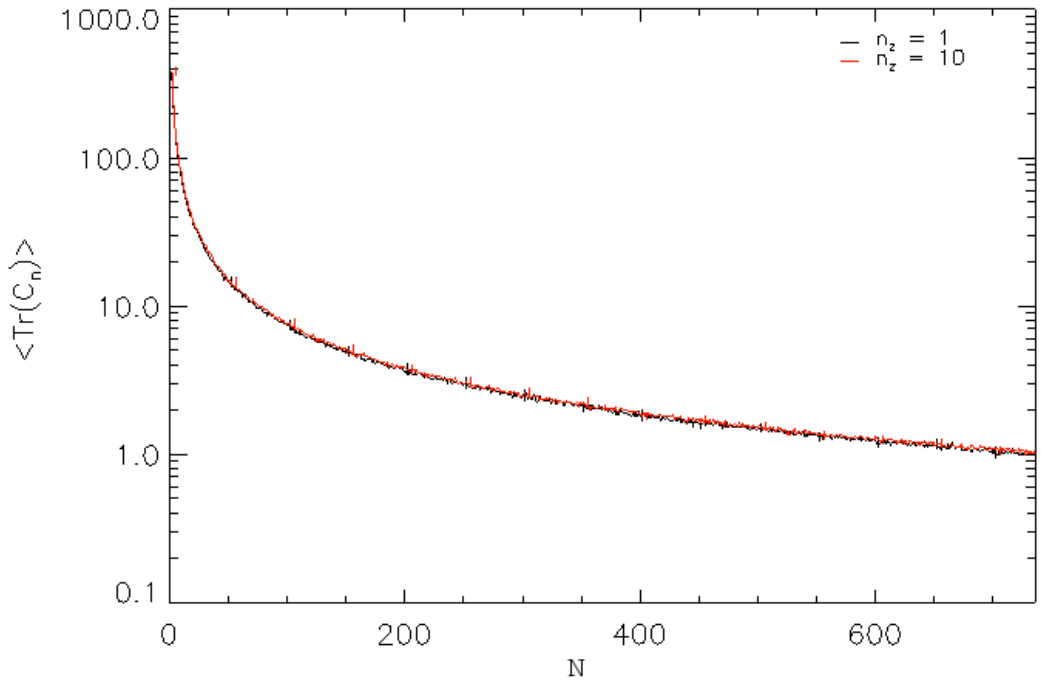

Figure 4.7: The mean trace of 1000 bootstrapped $\hat{\mathbf{C}}_{\rm NG}$, constructed from N lines of sight, as a fraction of the trace of the bootstrapped covariance matrix from 368 lines of sight for $n_{\rm z} = 1$, 10 and $n_\theta = 15$. Error bars are equivalent to the standard error on the mean from all 1000 bootstraps, plotted for multiples of $n = 50$. The errors from $n_{\rm z} = 10$ have been offset horizontally for an easier comparison to the $n_{\rm z} = 1$ errors.

lines of sight to ensure stability. For this reason, further tests are required to ensure the covariance matrix estimate is reliable, an issue which we address in §5.4.2.

4.4.3 Testing The Anderson Correction

As any bias in parameter contours can impact directly on parameter estimation, it is important to determine the validity of the Anderson correction in Eqn. 4.6 for this analysis. To do this, a large number ($N = 3200$) of Gaussian shear fields of 18.2 square degrees were simulated and the covariance matrix of the shear correlation functions of this set was estimated using Eqn. 4.1. Each simulated sky patch consists of three arbitrary tomographic slices and ξ_+ was measured at

six angular scales ranging from $1' \leq \theta \leq 85'$, meaning the data vector p contained 36 elements in total. Since $N \gg p$ in this case, we can assume that

$$\hat{\mathbf{C}}^{-1}_{*,N=3200} = \hat{\mathbf{C}}^{-1}. \tag{4.8}$$

New covariance matrices for different values of N were constructed by using subsets of the simulated Gaussian fields. As many unique covariance matrices as possible were generated for each N from the full 3200 fields. This ensures that when N is small and the covariance matrix may not be stable and unbiased, there are multiple covariance matrices to average over and improve the statistics. In this analysis, the smallest value of N was set to 40, which is close to the limiting case of $N = p$. This allowed us to compute 80 separate covariance matrices from the 3200 fields. The Anderson correction was then applied to each covariance matrix. The average of the trace of the covariance matrix for different N before and after the Anderson correction is shown in Fig. 4.8.

This plot closely matches Fig. 1 in H07. We find that application of the Anderson correction factor reliably removes this bias to within a few percent as long as $\frac{p}{N} \lesssim 0.8$. In the region $0.8 < \frac{p}{N} < 1$ the corrected matrix inverse $\hat{\mathbf{C}}^{-1}$ is still somewhat (unpredictably) biased, so it is prudent to ensure that the number of realisations of the simulations used in our analysis outweighs the length of the data vector by at least this factor.

The Anderson correction is especially pertinent in tomographic analyses, as without it the addition of redshift bins into the analysis has the effect of spuriously reducing the size of the credibility intervals on parameter estimates because the inverse covariance is poorly estimated. There seems to be more constraining power than there actually is. Without the Anderson correction, the contours are spuriously small, thus its use allows more accurate but less precise determination of cosmological parameters. It also reduces the precision that can apparently be gained by cutting the data up into higher numbers of tomographic bins; adding more bins will always increase the precision of the parameter estimates, but now at a diminishing rate.

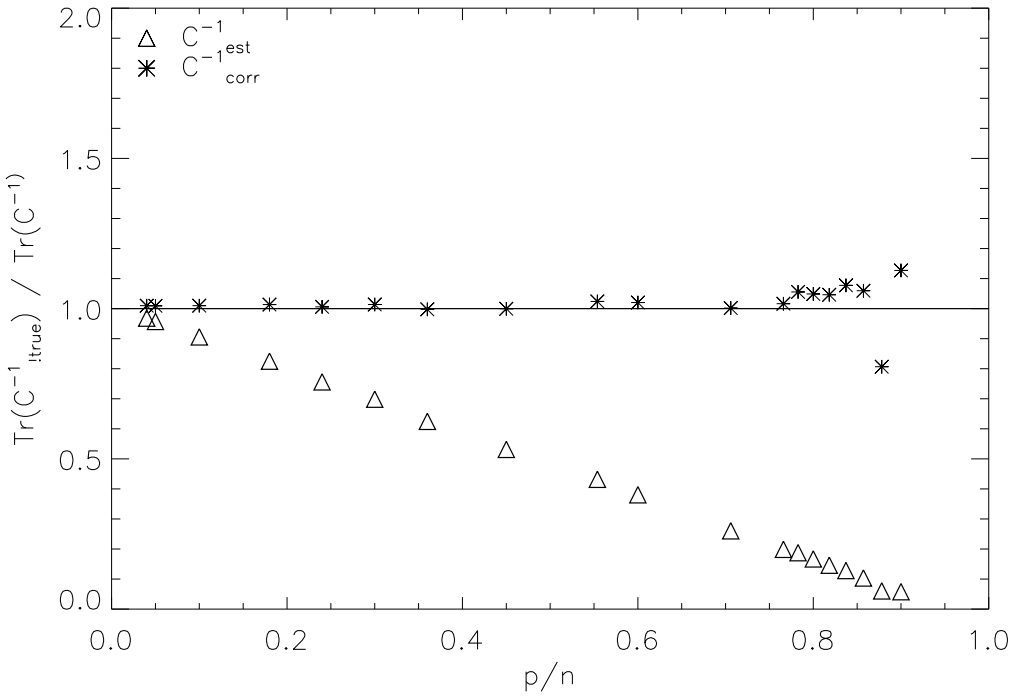

Figure 4.8: The ratio of the traces of the 'true' covariance matrix inverse ($\mathbf{C}^{-1}_{\rm true} = \mathbf{C}^{-1}_{*,\,\rm N=3200}$) to the covariance matrix inverse estimated from N simulations ($\mathbf{C}^{-1}_{\rm est} = \hat{\mathbf{C}}^{-1}_{*}$) and the Hartlap-corrected matrix inverse ($\mathbf{C}^{-1}_{\rm corr} = \hat{\mathbf{C}}^{-1}$) for different ratios of p/N measured from Gaussian simulations with shape noise. $n_{\rm z} = 3$, $n_\theta = 6$. Only the trace of the matrix is considered as the Anderson correction scales every element in the covariance matrix by the same amount.

4.4.4 Correcting for Non Gaussianity

The Sato Correction

The cosmic shear signal-to-noise ratio is highest at small scales, which are affected by nonlinear clustering [162]. As discussed earlier, $\hat{\mathbf{C}}_{\rm G}$ is fully Gaussian on all scales and hence does does not contain any nonlinear contributions and will be lower in amplitude than $\hat{\mathbf{C}}_{\rm NG}$ on scales where the nonlinear clustering is significant. This is the main disadvantage in using Gaussian simulations to analyse cosmic shear. If we can reliably correct for non-Gaussianity, then we can use the Gaussian simulations instead of the clones in our cosmological shear analysis (presented in Chapter 6). Since we can quickly generate Gaussian shear fields, this means we could produce a more stable estimate of the covariance matrix for use with higher-bin tomography.

To attempt to address the discrepancy between the two covariance matrices, we follow the approach of Sato et al. 2011 [180] and apply a fitting formula to $\hat{\mathbf{C}}_{\rm G}$ to estimate the contribution that would be present from nonlinear effects. This fitting formula (Eqn. 6 in [180]) defines a ratio $F(\theta, \theta', z_m)$ between the Gaussian and the non-Gaussian covariance matrices for median source redshift z_m and the correlation functions at two separations θ and θ':

$$F(\theta, \theta', z_m) \equiv \frac{C_{\rm NG}[\xi_+(\theta)\xi_+(\theta'); z_m]}{C_{\rm G}[\xi_+(\theta)\xi_+(\theta'); z_m]}. \tag{4.9}$$

[180] use 1000 ray-tracing simulations to compute $\hat{\mathbf{C}}_{\rm NG}$ and $\hat{\mathbf{C}}_{\rm G}$, then they derive the calibration function $F(\theta, \theta', z_m)$ by parametrising it with the form

$$F(\theta, \theta', z_m) = \left(a(z_m) + \frac{b(z_m)}{(\theta\theta')^{c(z_m)}}\right) d(z_m)^{|\theta-\theta'|}. \tag{4.10}$$

The parameters a, b, c and d are then estimated from simulations with 6 different source redshifts, z_m to be

$$\begin{aligned} a(z_m) &= -z_m^{a_1} \exp(a_2 z_m) + a_3 \\ b(z_m) &= b_1 z_m^{b_2} + b_3 \\ c(z_m) &= c_1 z_m^{c_2} + c_3 \\ d(z_m) &= d_1 z_m^{d_2} + d_3, \end{aligned} \tag{4.11}$$

where $(a_1, a_2, a_3) = (-3.7683, 0.9752, 1.4048)$, $(b_1, b_2, b_3) = (10.7926, -2.0284, -0.2266)$, $(c_1, c_2, c_3) = (-0.3664, -0.5733, 0.6863)$ and $(d_1, d_2, d_3) = (0.2450, 0.1218, 0.7076)$.

There are two problems with applying this correction to our analysis. First, [180] assume a delta-function redshift distribution for each redshift slice, rather than a realistic distribution such as that measured from CFHTLenS. Second, the Sato correction was derived using non-tomographic simulations, so it may not be applicable to the auto-correlations of tomographic bins, and even less so to cross-correlations, which we require. Despite these limitations, we investigate whether the Sato correction can perform well enough on a realistic tomographic analysis to be of use in this thesis.

Since our source redshifts encompass a range of values, z_m was taken to be the median redshift of each tomographic bin (for cross correlations between bins, we average the medians of the contributing bins). [180] only apply their fitting formula below $10'$ as the covariance matrix is largely Gaussian above this scale. We apply the Sato correction to both the cross- and auto-correlations of the covariance matrix. Figs. 4.9– 4.10 shows the effect of the fitting formula on the discrepancy between $\hat{\mathbf{C}}_{\rm NG}$ and $\hat{\mathbf{C}}_{\rm G}$ for several tomographic combinations. We apply the fitting formula to every element of our Gaussian matrix regardless of angular scale (with the condition that $F(\theta, \theta', z_m) \geq 1$), but find that correction only has a significant effect below $\sim 10'$, in good agreement with [180].

We find that on nonlinear scales below $\sim 10'$ for $z_m \gtrsim 0.34$, the fitting formula works well to correct the matrix for non-Gaussianity to a similar accuracy of [180] of around 25% or better. At $\sim 10'$, the fit ceases very quickly to have an effect and the Gaussian and non Gaussian covariance matrices match well. It is important to correct for nonlinearity properly on these scales, as systematic errors are smaller in this range than at small scales of a few arcminutes or less. Thus any discrepancy between the Sato correction and the N-body prediction will act as a new source of systematic error, which could have implications for a cosmological analysis.

Additionally we find that the Sato correction breaks down entirely for some of the redshift slices and is unable to approximate the clone covariance from the Gaussian. This is because at redshift $z_m \lesssim 0.34$, the $c(z_m)$ coefficient becomes negative and $F(\theta, \theta', z_m)$ begins to increase with angular scale, instead of decreasing. [180] use simulations with $z_m \geq 0.6$ to derive their fit, so they do not

probe this low-redshift regime and encounter this problem. The effect becomes significant when we include three or more tomographic bins in the covariance matrix, since for our input redshift distribution we will always have at least one redshift bin with $z_m \lesssim 0.34$. Even above this critical redshift, we see the fit performing poorly at $z_m \sim 0.4$ (for example in the 1-3 bin in Fig. 4.10) and for many of the correlations involving the lowest redshift bin for $n_z \geq 3$ due to this effect.

Improving the Sato Correction

Since the redshift distribution we use is that of the CFHTLenS galaxies, it is clear that despite its success at correcting a 2D, high-redshift covariance, the Sato correction in its present form is inadequate for correcting for non-Gaussianity in tomographic simulations with realistic redshift distributions and narrow bins. This is not surprising as we are pushing the method beyond what it was designed for. For this reason, we attempt to modify the coefficients in the Sato fit to better correct for non-Gaussianity over the redshift range spanned by the tomographic bins in our analysis.

We use the MCMC algorithm described in Chapter 3 and the expression for χ^2 in Eqn. 4.4 to attempt to find better-fitting values of the coefficients in Eqn. 4.10. We set the diagonal of $\hat{\mathbf{C}}_{\rm NG}$ as the 'data' and the diagonal of $\hat{\mathbf{C}}_{\rm G}$ plus the Sato correction as the 'model'. We obtain an estimate of the error on $\hat{\mathbf{C}}_{\rm NG}$ by constructing the covariance of the covariance matrix (which we denote the co-covariance). This is done by constructing 8 estimators of the clone covariance from 92 realisations each (since the covariance estimated from $\lesssim 100$ realisations begins to become unstable, see Fig. 4.7). Since the Sato fit coefficients in Eqn. 4.11 are independent of angular scale and redshift, we do not explicitly perform a best fit to the off-diagonals of the Gaussian covariance matrix.

We restrict the range of angular scales to which we apply the likelihood fit to $\theta < 12'$ since the Sato correction needed diminishes quickly above $\sim 10'$. We select $n_\theta = p = 6$, such that $p/N = 0.8$, the highest value for which the inverse estimator of the covariance is stable. We must also ensure that there is at least 1 degree of freedom in χ^2 (ν, equal to the number of data points minus the number of fitted parameters, see Chapter 3). This means that we cannot fit all twelve Sato coefficients, and instead choose to fit just c_1, c_2 and c_3 since the Sato fit first

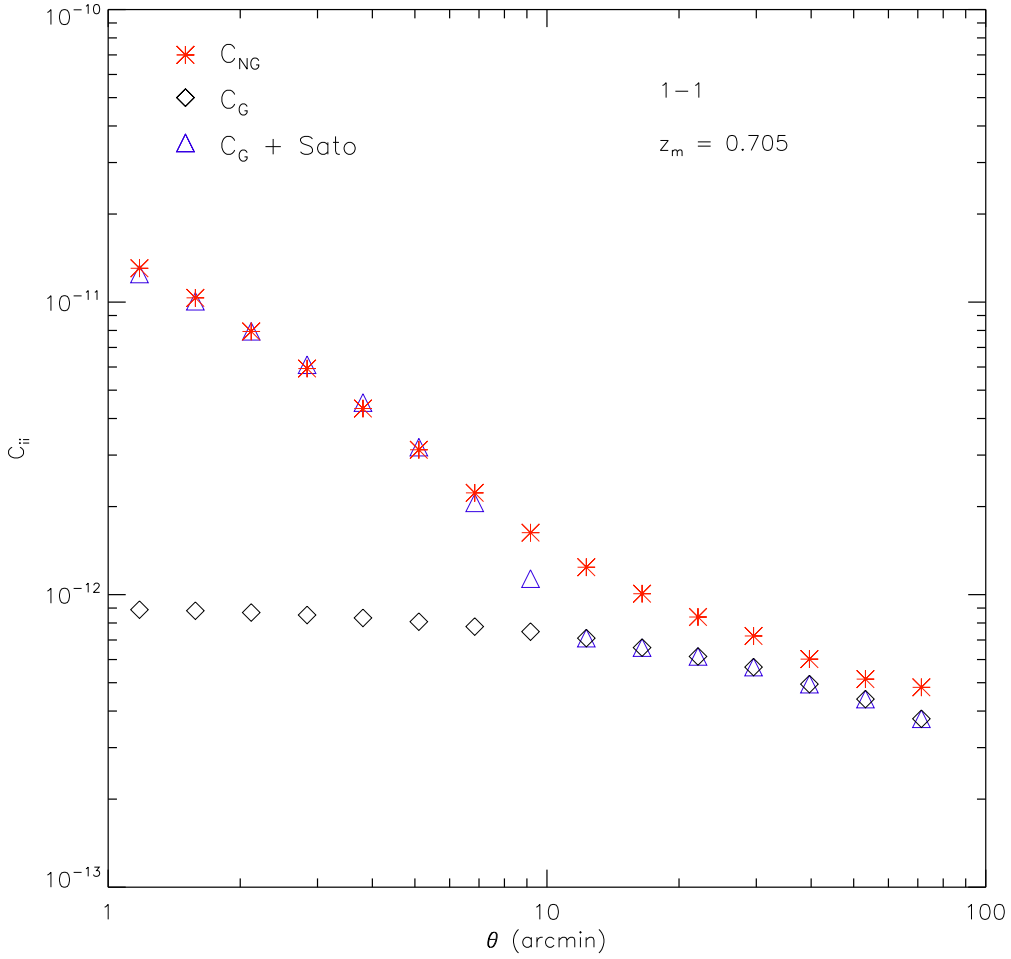

Figure 4.9: The diagonal of the covariance matrix from the Gaussian simulations with (blue triangles) and without (black diamonds) the Sato correction applied, for $n_z = 1$ and $n_\theta = 15$. Also plotted is the clone covariance diagonal (red asterisks) for comparison. All covariance matrix terms are scaled to match the sampling expected from a CFHTLenS-sized survey. The median redshift and bins contributing to the plotted redshift correlation are shown in the top right of the plot.

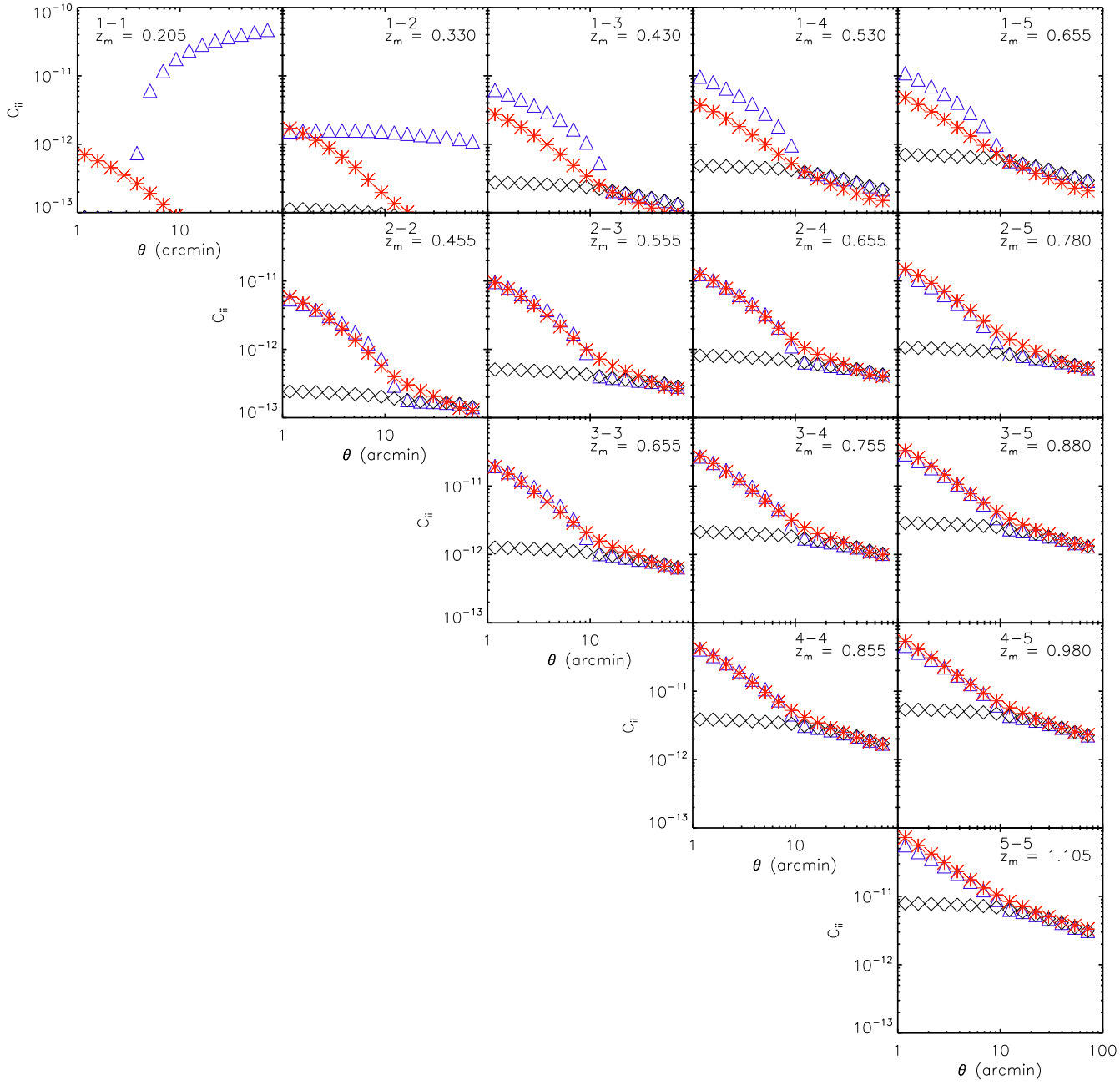

Figure 4.10: Clone, Gaussian and Sato-corrected Gaussian covariance diagonal for $n_z = 5$, $n_\theta = 15$. Symbols are the same as in Fig. 4.9.

fails when $c(z_m)$ becomes negative. We fix all other coefficients to the fiducial values in [180].

We must fit just one redshift correlation at a time when $n_z > 1$ to ensure $p = 6$. We perform a χ^2 fit on every auto- and cross-correlation for $n_z = 1$ and 2, and we also investigate correlations where either the total $z_m \leq 0.34$ and the fit fails completely, or correlations where one of the contributing redshift slices has $z_m \leq 0.34$ and the fit does poorly for $n_z = 5$. This corresponds to all correlations with the lowest redshift bin in them. Due to the way $c(z_m)$ is parameterised, we find that the best-fit coefficients c_1, c_2 and c_3 are highly degenerate. We therefore restrict the range of parameter space we probe fairly narrowly to the ranges shown in Figs. 4.11-4.12. We plot just the best fit results from $n_z = 1$ and the 1-1 auto-correlation from $n_z = 5$. The results for redshift correlations with total $z_m \geq 0.34$ are very consistent and resemble the results for $n_z = 1$; the mean coefficients from almost all fits for $n_z = 1$, 2 and 5 are: $c_1 = -0.341$, $c_2 = -0.470$ and $c_3 = 0.619$ which are very similar to the fiducial values. Fig. 4.12, however, shows that when the total median redshift is lower than the threshold value of 0.34, the best-fit constraints are quite different. For the 1-1 and 1-2 correlations for $n_z = 5$, the best-fit values are: $c_1 = -0.351$, $c_2 = -0.488$ and $c_3 = 0.840$. The main alteration is to the value of c_3. The fact that the best-fit coefficients behave very differently on either side of the limiting redshift tells us that we cannot use just one set of coefficients to characterise non-Gaussianity over all redshift slices.

We re-plot Figs. 4.9 and 4.10 with our alternative coefficients to illustrate the effect of these best-fit values. We use the best-fit values of $c_1 = -0.341$, $c_2 = -0.470$ and $c_3 = 0.619$ for the majority of bins where $z_m > 0.34$ and the alternative best-fit values of $c_1 = -0.351$, $c_2 = -0.488$ and $c_3 = 0.840$ where $z_m > 0.34$. This allows us to retain the relatively good agreement we find at high z_m, while attempting to better account for non-Gaussianity at low z_m. The results are shown in Figs. 4.13-4.14. The best-fit coefficients for $z_m \geq 0.34$ perform well, as expected. However, the best-fit coefficients for $z_m \leq 0.34$ still perform very poorly and almost no non-Gaussian correction is being applied. This is most likely due to the need for optimisation over all the coefficients in Eqn. 4.11, not just $c(z_m)$. Additionally, the Sato correction was not calibrated to a tomographic analysis. This means that there are cases where the fit does not perform as expected even at high z_m. For example, the 1-5 and 2-4 correlations in Fig. 4.14

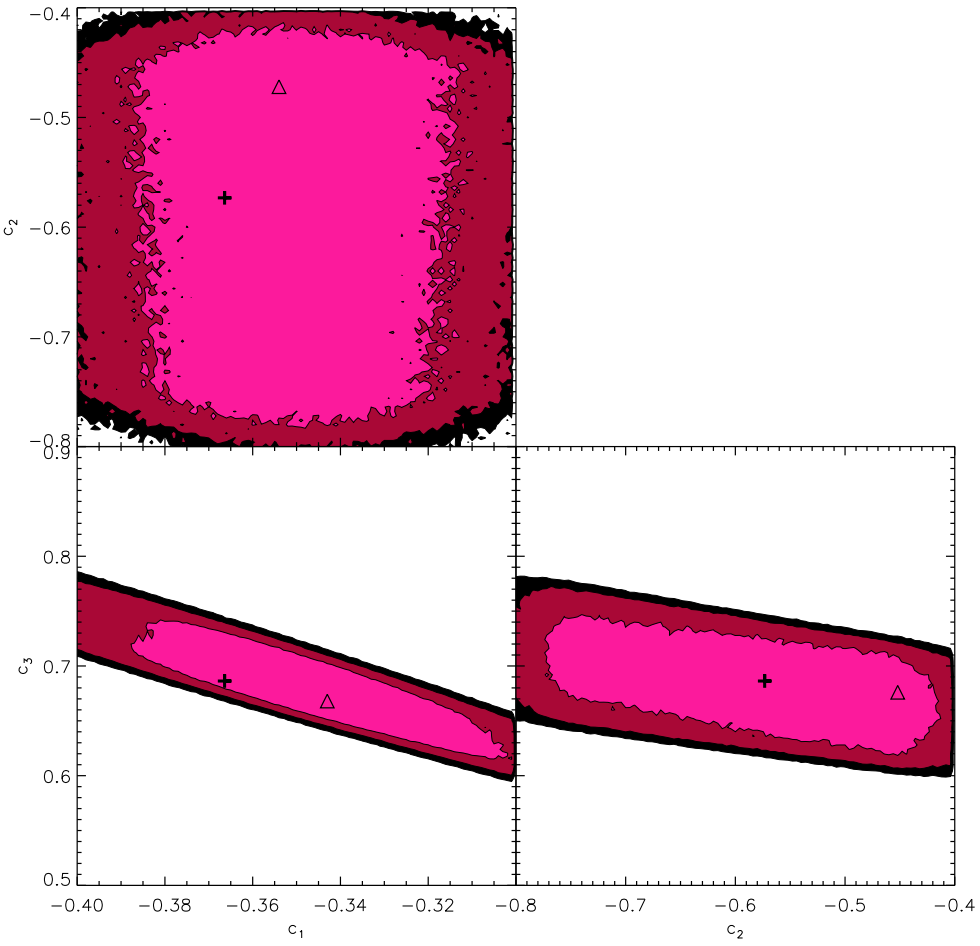

Figure 4.11: The 68.3%, 95.4% and 99.0% likelihood contours for c_1, c_2 and c_3 in Eqn. 4.11 for $n_z = 1$, $n_\theta = 6$. Crosses represent the fiducial coefficient values from [180], the triangles mark the best-fit values from this analysis.

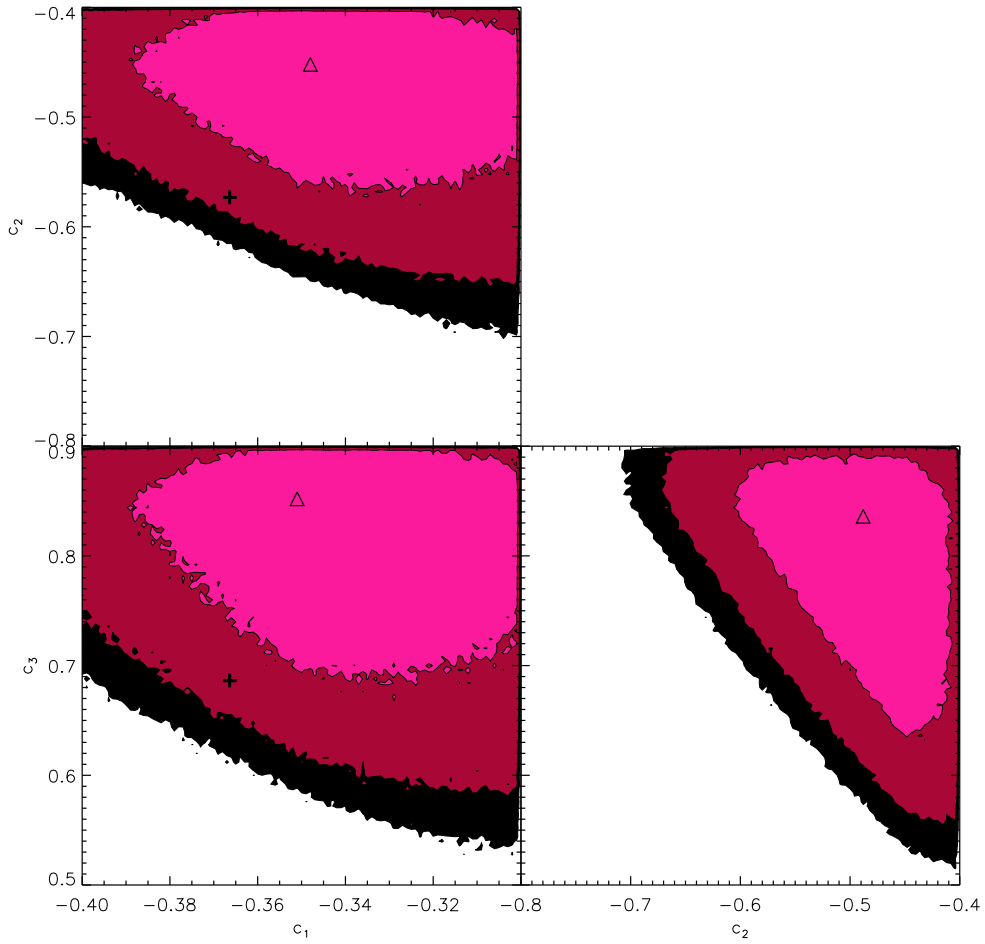

Figure 4.12: The 68.3%, 95.4% and 99.0% likelihood contours for c_1, c_2 and c_3 in Eqn. 4.11 for the 1-1 redshift correlation of $n_z = 5$, $n_\theta = 6$. Crosses represent the fiducial coefficient values from [180], the triangles mark the best-fit values from this analysis.

have the same median redshift and hence the same Sato correction. This is clearly not appropriate, since the non-Gaussian contribution measured from the clone for the 1-5 correlation is lower than for the 2-4 correlation. Thus, to accurately correct for non-Gaussianity in a tomographic analysis, the Sato correction is not appropriate in its current form and would have to be reparametrised to incorporate information on both redshift slices in a correlation.

4.5 Discussion

We have constructed estimators of the covariance matrix of the shear correlation function measured from 736 lines of sight from Gaussian and N-body simulations. Figs. 4.5-4.6 show that the covariance matrices from both simulations is highly off-diagonal as $\hat{\xi}_+$ are very correlated. Within each redshift correlation, similar angular scales are strongly correlated, and as the redshift increases, more widely-separated angular scales become more correlated. This is because at high redshift, a given angular scale corresponds to a smaller real-space separation than at low redshift. These plots highlight the significance of incorporating the covariance matrix into a likelihood analysis - we cannot assume that the variance of the shear correlation function contains the majority of the error. It is therefore critical to use the correct expression for χ^2, which we express in Eqn. 4.4, in any likelihood analysis.

The area scaling derived by [116] that we apply to our covariance matrices works well to scale the errors to those expected from a CFHTLenS-like survey of 154 square degrees. This is evidenced in Figs. 4.9-4.10, as the clone and Gaussian covariances (derived from simulations with different areas) match well at large scales as expected. This large-scale ($\gtrsim 10'$) agreement, as well as the success of both simulations in recovering the shear correlation function (see Figs. 4.3-4.4), is good evidence that both sets of simulations are performing well on the expected scales. The covariance matrix becomes more stable and a better estimator of the true covariance with increasing numbers of realisations, as shown in Fig. 4.7. After around $N \gtrsim 100$, the covariance matrix is relatively stable. Since we have 736 lines of sight from the clone simulations, we expect that the covariance matrix is stable enough to perform a likelihood analysis. We find that the stability of the covariance matrix is unaffected by the number of tomographic bins used. This is

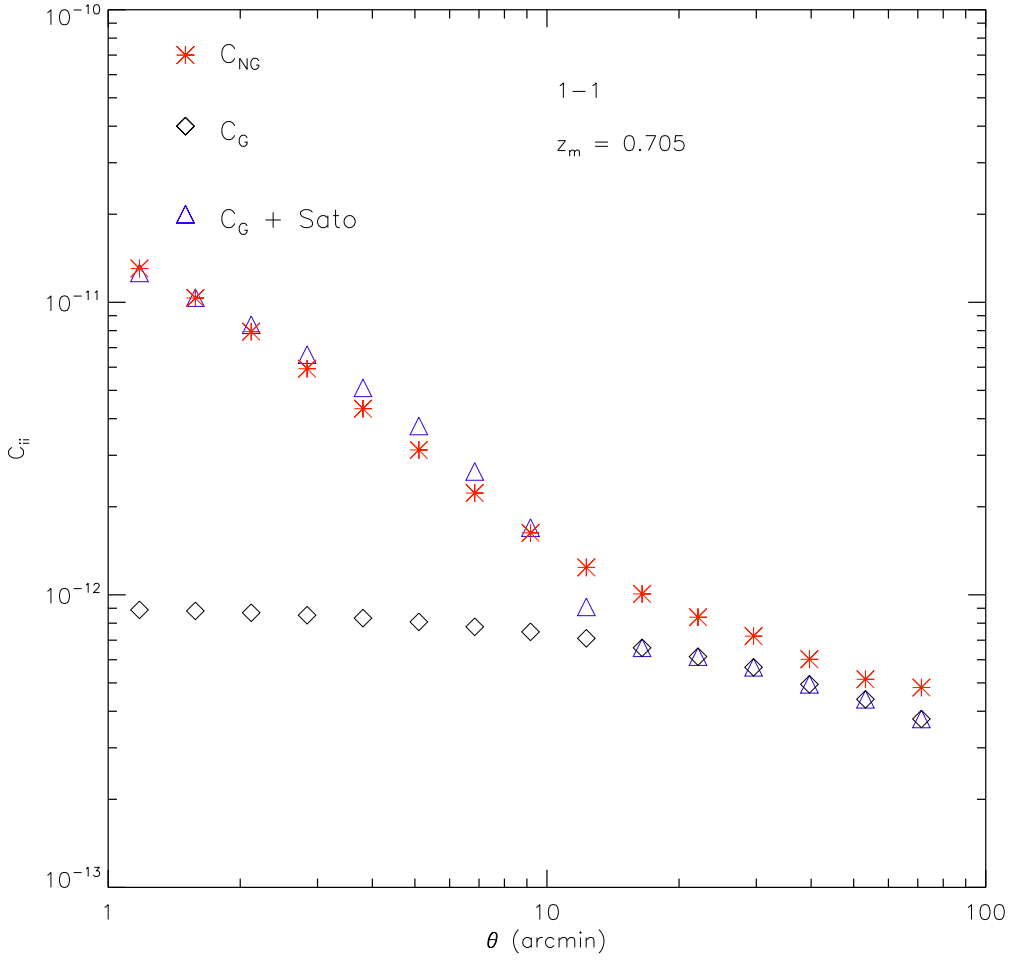

Figure 4.13: Clone, Gaussian and Sato-corrected Gaussian covariance diagonal using $c(z_m)$ coefficients determined by an MCMC analysis ($c_1 = -0.341$, $c_2 = -0.470$ and $c_3 = 0.619$), for $n_z = 1$, $n_\theta = 15$.

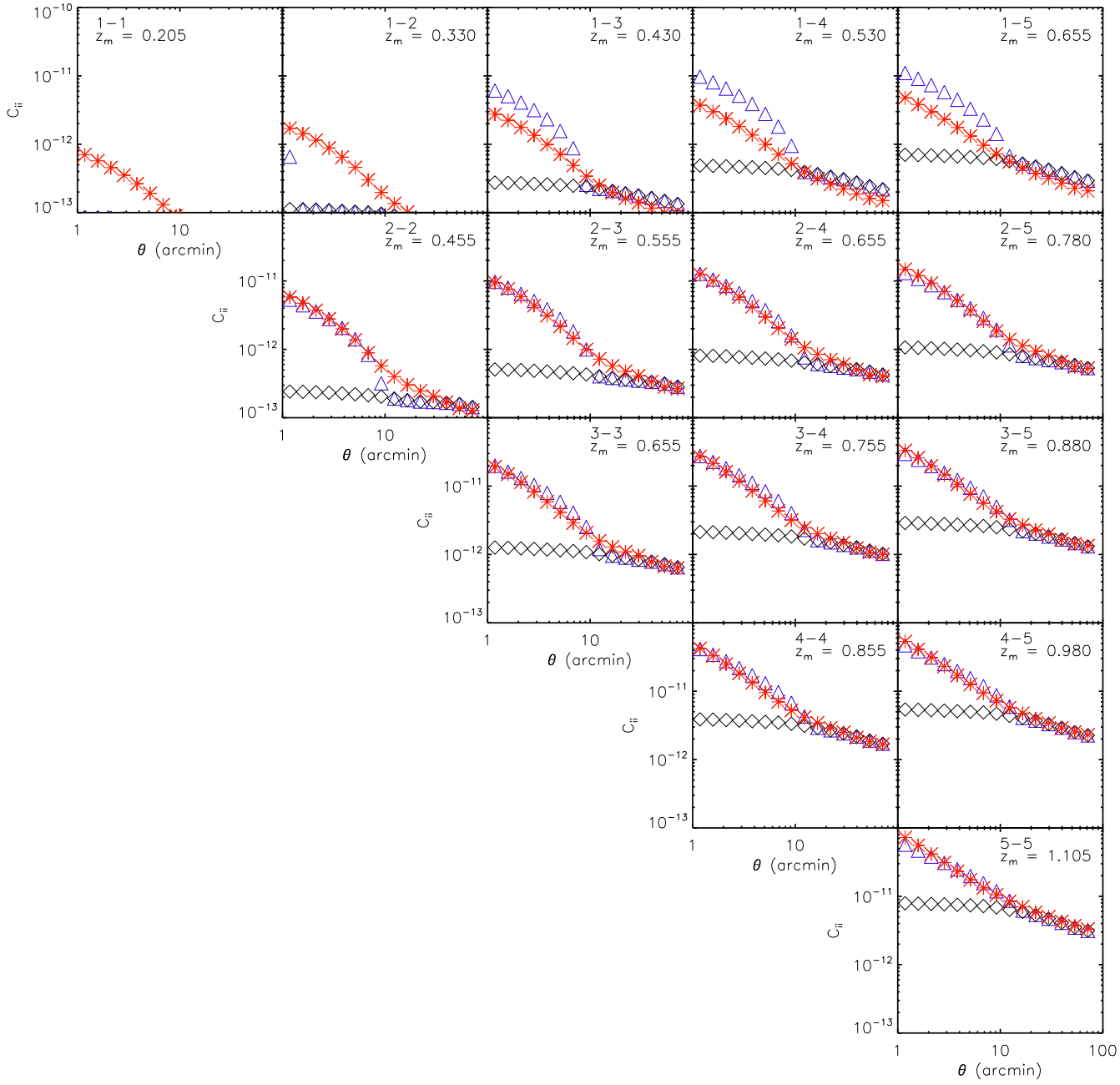

Figure 4.14: Clone, Gaussian and Sato-corrected Gaussian covariance diagonal using $c(z_m)$ coefficients determined by an MCMC analysis ($c_1 = -0.341$, $c_2 = -0.470$ and $c_3 = 0.619$ for $z_m \geq 0.34$; $c_1 = -0.351$, $c_2 = -0.488$ and $c_3 = 0.840$ for $z_m < 0.34$), for $n_z = 5$, $n_\theta = 15$.

most likely because our covariance matrix is sampling variance limited, which is a function of survey area, configuration and galaxy density and is hence constant with n_z. The stability of the covariance matrix with n_z may vary, however, once other effects such as galaxy shape noise are included in its estimate. We investigate this effect in the next chapter.

We investigate whether the non-Gaussianity of the shear covariance can be accounted for by fitting the empirically-derived correction in [180] to the Gaussian covariance. We show in Figs. 4.9-4.10 that the correction appears to work well when the median redshifts of both bins contributing to a given redshift correlation are above a critical redshift of $z_m \gtrsim 0.34$. If the median redshift of one of the bins is below this, the fit tends to over-correct for non-Gaussianity and the resulting covariance is too high. If the median redshift of both slices is below 0.34, the fit fails completely as the $c(z_m)$ function introduced in Eqn. 4.9 and defined in Eqn. 4.11 becomes negative. We attempt to correct this unsuccessfully. We conclude that for a shear analysis that is both tomographic and involves low redshift slices, the Sato fit is unable to correct for non-Gaussianity to a high enough degree of accuracy to justify using Gaussian simulations.

Our test of the Anderson correction has shown that ideally we need $p/N < 0.8$ to enable us to produce a stable estimator of the inverse of the covariance matrix, as shown in Fig. 4.8. We must keep this limit in mind when deciding on the maximum number of tomographic and angular bins we wish to constrain in a likelihood analysis. The total size of the data vector p is determined by the number of correlations between redshift bins multiplied by the number of angular scales used. Therefore we can somewhat increase the number of tomographic bins obtainable by decreasing the number of angular scales we measure $\hat{\xi}_+$ on. However, if we set n_θ too low, we will have too few data points over which to sample the shape of $\hat{\xi}_+$ and we will lose information and hence constraining power. We calculate the minimum number of realisations of the covariance matrix, $N_{\rm min}$, needed to satisfy the Anderson limit of $p/N < 0.8$ for all combinations of $1 \leq n_z \leq 15$, $3 \leq n_\theta \leq 5$ in table 4.2.

Table 4.2 shows the number of lines of sight needed for different numbers of tomographic bins. Bin combinations that result in $0.8 \leq p/N \leq 1.0$ from the 736 clones lines of sight are coloured in blue, and combinations for which

Table 4.2: Minimum no. of realisations $N_{\rm min}$ needed for various $n_{\rm z}$ and n_θ to ensure $p/N < 0.8$.

$\rightarrow n_\theta$ / $\downarrow n_{\rm z}$	3	4	5	6	7	8	9	10	11	12	13	14	15
1	4	5	7	8	9	10	12	13	14	15	17	18	19
2	12	15	19	23	27	30	34	38	42	45	49	53	57
3	23	30	38	45	53	60	68	75	83	90	98	105	113
4	38	50	63	75	88	100	113	125	138	150	163	175	188
5	57	75	94	113	132	150	169	188	207	225	244	263	282
6	79	105	132	158	184	210	237	263	289	315	342	368	394
7	105	140	175	210	245	280	315	350	385	420	455	490	525
8	135	180	225	432	315	360	405	450	495	540	585	630	675
9	169	225	282	338	394	450	507	563	619	675	732	788	844
10	206	275	344	413	482	550	619	688	757	825	894	963	1032
11	248	330	413	495	578	660	743	825	908	990	1073	1155	1238
12	293	390	488	585	683	780	878	975	1073	1170	1268	1365	1463
13	342	455	569	683	797	910	1024	1138	1252	1365	1479	1593	1707
14	394	525	657	788	919	1050	1182	1313	1444	1575	1707	1838	1969
15	450	600	750	900	1050	1200	1350	1500	1650	1800	1950	2100	2250

$p/N > 1.0$ are coloured in red. We find that to be able to perform a 10-bin tomographic likelihood analysis, one must use a minimum of 206 realisations of the data vector to calculate the covariance matrix to ensure $p/N < 0.8$. Using fewer realisations than $N_{\rm min}$ for a given bin combination will result in a matrix that is not stable even under pseudo-inversion and, in the case of $p/N > 1$, a χ^2 calculation that breaks down. With the clone simulations, 10-bin tomography is statistically feasible since we divide the clones into 2 x 2 subfields to produce 736 lines of sight in this analysis and are below the limiting value of p/N for $n_\theta \leq 13$. 15-bin tomography is just on the edge of feasibility, although performing a 15-bin tomographic analysis above $n_\theta = 4$ would take us into the unstable regime where $p/N > 0.8$, and our (pseudo) inversion of the covariance matrix may be unreliable. This table shows the fundamental limit from using only 736 lines of sight which must be taken into account in a likelihood analysis.

If one is to use the clone simulations to perform a cosmological likelihood analysis, it is statistically feasible to use up to $\sim$ 15 tomographic redshift bins. However, the values derived in table 4.2 only tell us what the statistical limitations on a tomographic analysis are. They do not take into account

other effects that may limit the number of tomographic bins it is feasible to analyse, such as photometric redshift errors, galaxy shot noise, the scale over which the lensing signal changes or the strength of the correlations caused by strongly overlapping redshift bins (such as the ones in this analysis, see Fig. 4.2). Additionally, systematic errors due to IAs may limit the accuracy and precision of a tomographic analysis, and prohibitive computing times for high n_z combinations may introduce a practical limitation to such an analysis. In Chapter 6, we incorporate these effects into a likelihood analysis to investigate what tomographic combinations are feasible and which is optimal for a CFHTLenS-like survey.

Gaussian statistics lead to an underestimate of the cosmic shear covariance on small scales, and an ad hoc correction like the Sato correction in Eqn. 4.10 is not easily applicable to all surveys as it must be calibrated to the cosmology of a specific simulation and was not designed for tomography. A potential method for producing covariance matrices that better approximate the N-body results in the nonlinear regime would be the use of log-normal shear fields. With a log-normal approximation, the logarithm of the shear values in the simulated shear fields appears Gaussian distributed, instead of the shear values themselves. Log-normal simulations produce more realistic covariance matrices and parameter credibility intervals than Gaussian simulations do at a slightly increased computational cost [181]. A future extension to this work would be to investigate the covariance matrix generated from log-normal simulations compared to that from N-body simulations, and their ability to recreate non-Gaussianty at small scales.

Most cosmic shear analyses assume that the covariance matrix of the shear is constant with cosmology, however [182] find that this is not the case. They calculate analytically covariance matrices for 2500 different cosmologies using the approach described in [116], and find that likelihood intervals for Ω_m and σ_8 can vary significantly in size and orientation with the choice of underlying cosmology. The magnitude of the effect increases with the statistical precision of the survey, meaning that this is an effect that future large-scale lensing surveys may have to account for carefully.

In data analysis of real surveys where the full covariance matrix is estimated from simulations, one would ideally calculate a covariance matrix for every point in parameter space that we wish to probe. However, this would entail generating a set of simulations from which to measure $\hat{\mathbf{C}}$ for every cosmology, which

would require a prohibitive amount of computational power. [182] recommend determining approximate best-fit parameter values (perhaps through the use of priors or running a trial MCMC chain as described in Chapter 3) and estimating the covariance matrix of this point only. Since most of the likelihood is close to the best-fit value(s), the covariance matrix estimate will hopefully not vary too much over the regions of interest. A compromise between the two approaches might be to use the results of such a trial MCMC chain to get an idea of the location and size of the likelihood contours. Then, covariance matrices corresponding to several cosmologies within e.g. the 99% likelihood contour could be estimated, and one could interpolate between covariance matrices to get a better estimate of the covariance corresponding to any point in the parameter space. This would mean generating several sets of realisations of simulations (instead of just one) using different input cosmologies, but one can use as many or as few sets as computational and time resources allow to improve the accuracy of the interpolation.

4.6 Conclusions

In this chapter we have described two types of shear simulations: N-body and Gaussian. The clones include nonlinear structure and produce realistic error estimates, but are expensive to compute and we only have 736 realisations of 2.70 square degrees. The Gaussian simulations are easy-to-generate, populated Gaussian shear fields spanning 18.20 square degrees each that recover the mean correlation function well but underestimate errors on small scales. We have constructed and measured estimators for the shear correlation function and its covariance from both suites of simulations. To perform a cosmological likelihood analysis on cosmic shear data, as we will do in Chapter 6, one must have an honest, reliable estimate of the covariance of the data from which to calculate parameter constraints. We have investigated the number of realisations of simulations that must go into constructing this estimator, and found that we have enough lines of sight from both simulations to ensure stability. We investigate whether we can reliably correct for nonlinearity that is missing from the Gaussian simulations, and find that an empirical correction employed in [180] is insufficient. This means that if we are to use a covariance matrix that best approximates the true covariance

from real data, we must use the clone simulations to estimate it and we are limited to 736 lines of sight.

We investigated how many tomographic bins one can use in a cosmological likelihood analysis using the covariance matrix. We find that the Anderson correction of [169] works well in the limit of $p/N < 0.8$; that is, when the number of realisations from which we measure the data vector outweigh the number of elements in the data vector by a factor of 25%. This limit means as long as the clone simulations are gridded up to produce the full 736 lines of sight, we can (in theory) comfortably achieve 10-bin tomography and (uncomfortably) reach 15-bin tomography in a cosmic shear likelihood analysis. The number of tomographic bins that are optimal and even feasible may well be less than this, and we constrain this in Chapter 6.

Chapter 5

Intrinsic Alignments and Shape Noise Covariance

5.1 Introduction

In the previous chapter, we estimated the shear correlation function and its covariance matrix from N-body and Gaussian simulations, and concluded that we could not account for non-Gaussianity in the covariance matrix without the use of N-body simulations. The covariance matrix in Chapter 4, however, is a measure of only the sample variance. Real lensing data also contains shot noise due to the uncertainty inherent in galaxy shape measurements. Additionally, galaxy ellipticities in the real Universe are intrinsically correlated with each other and the density field, meaning that lensing measurements are contaminated by the II and GI alignment signals introduced in §2.3.5 which contribute to the measured shear correlation function and its covariance. Any realistic tomographic likelihood analysis must therefore take into account the effect of both shot noise and IAs on the shear correlation function and the covariance matrix.

There have been three key tomographic analyses of cosmic shear data to date. A 2-bin analysis of CFHTLS -Deep data was carried out by [154]. Photo-zs were obtained from survey galaxies, but the effects of catastrophic outliers are not accounted for and could be significant. The fitting function of [183] is used to approximate the true redshift distribution, which may not be accurate. The authors ignore the effect of intrinsic alignments, but since the data are split into only two broad tomographic bins, this approach is unlikely to introduce

significant bias in the cosmological constraints. The constraining power of the Deep survey is limited by its small area of $\sim$ 4 square degrees, however [154] constrain σ_8 to 0.86 ± 0.05 (68% CL, $\Omega_m = 0.3$, using the halo model of [28]), showing that tomography has the power to produce useful cosmological constraints. The authors produce confidence limits from the analytical covariance matrix prescription of [116], which fails to take into account non-Gaussianity and cross-bin correlations.

A 3-bin tomographic analysis of HST COSMOS data was later performed by [155]. COSMOS has an effective survey area similar to the CFHTLS-Deep of ~ 1.64 sq. degrees. [155] also neglect IAs and use photometric redshifts (obtained from the ground). The authors deal with catastrophic outliers by excluding from the analysis galaxies which have significantly double-peaked posteriors or any probability in their PDFs above $z = 1$, since in this regime catastrophic outliers could not be reliably identified. This approach, however, results in information loss. The authors use a covariance matrix that includes a shot noise and a sampling variance term to estimate the errors, however they do not take into account the cross term(s) between the two. Despite the presence of a B-mode on small scales, [155] achieve constraints on Ω_m and σ_8 that are consistent with WMAP 3-year results [184], and they find that tomography can produce a factor of three improvement in credibility intervals. A subsequent analysis of the same data set by [156] improved on these results by performing 6-bin tomography. Improved photo-z estimates from 30 bands were used, leading to B-modes that are consistent with zero. The authors partially account for the presence of IAs following the method of [185] by excluding auto-correlations between all narrow redshift bins. This approach removes much of the II signal but at the cost of heavy information loss. Luminous red galaxies (LRGs) are excluded in an attempt to mitigate the GI signal, since these galaxies are known to carry the strongest alignment signal [186, 187]. The covariance is estimated from the Millennium Simulation, which naturally accounts for nonlinearity although it may be inaccurate as σ_8 is higher than the WMAP value (see §4.2). The authors apply the Anderson correction discussed in Chapter 4 to the covariance. Their cosmological constraints are consistent with those of [155] at the 1σ level, and they find a similar improvement to credibility intervals as a result of tomography.

In addition to tomography, it is also possible to perform a full 3D lensing

analysis. [188] perform a proof-of-concept 3D analysis of 0.78 sq. degrees of COMBO-17 data. This involves using photometric redshift information to estimate the shear field via a spherical harmonic expansion [128]. The COMBO-17 dataset has good photometry and hence reliable photo-z estimates. A Fisher matrix prediction for the errors is used, and the authors produce cosmological constraints that are consistent with WMAP 3 results [184]. The authors neglect IAs. 3D lensing has the advantage over tomography that it can straightforwardly exclude nonlinear modes which are difficult to model, although tomography is conceptually simpler and better for direct investigation of the redshift dependence of the signal [189]. For these reasons, 3D lensing and tomography are complimentary and future lensing analyses may benefit from using both techniques.

Along with shape measurement errors and the characterisation of the point spread function, IAs are known to be one of the main sources of systematic error in weak lensing analysis [190, 191, 186, 157, 192]. Analysis of Sloan SDSS data by [186] found that IAs can bias the lensing signal by $\sim$10%, hence it is extremely important to model the effect of GI and II alignments as accurately as possible in future weak lensing analyses. Our understanding of IAs is hindered by the fact that galaxy formation is complicated by gas dynamics and the nonlinear evolution of the matter density field on small scales [176]. We do not currently have a robust, well-tested model of galaxy alignments that can be used to accurately account for their effect on lensing measurements. Various methods for correcting IA contamination have, however, been proposed. [193] and [191] developed optimisation methods for downweighting close galaxy pairs to reduce the II alignment contribution, which for a data set with realistic photometric redshift errors involves assigning weights of zero to galaxy pairs closer than some redshift difference α. The value of α depends on angular scale and survey specification and is chosen to minimise the shear correlation function error. [194] applied this method to COMBO-17 data, reducing a systematic II error of $\sim$10% to a statistical error of $\sim$1%.

[195] employ a 'nulling' technique to reduce the IA signal, assigning nulling weights to the shear signal in each low-redshift bin contributing to the shear cross correlation function. Nulling can reduce the GI signal by a factor of $\sim$100 with respect to the lensing (GG) signal, and conveniently removes the II signal

as well. As a result, parameter estimates are less biased, but since nulling is redshift-dependent this comes at the price of a considerable loss in constraining power amounting to an increase in the 1σ dark energy contour area by a factor of two to three. Importantly, the nulling weights assigned are cosmology dependent, and if the weights are incorrect, this may make parameter biases from IAs worse.

[157] (hereafter BK07) model IAs using a simplification of the linear alignment model of [196] and [119], and extend this into the nonlinear regime by including the full nonlinear matter power spectrum. This 'nonlinear alignment' (NLA) model matches to superCOSMOS data slightly better than other models [197]. BK07 find that neglecting IAs can bias σ_8 by up to $\sim$5% and the dark energy equation of state by $\sim$50% if these are the only parameters being fitted; these figures rise if there are more free cosmological parameters. Additionally, they find that if IAs are accounted for, two or more times as many redshift bins are required to recover 80% of the lensing information, degrading parameter estimates. Current methods for dealing with IAs either result in information loss or else rely on IA models that have been observationally constrained only at low redshift ($z \lesssim 0.3$). It is important to test and develop accurate IA models, particularly when accounting for the GI signal, in order to avoid significant parameter constraint degradation or bias.

The effects of IAs are pertinent when using galaxy redshift estimates in a tomographic analysis. Photo-z errors cause leakage of the IA signal between tomographic bins - there will be a non-negligible II (GI) signal in cross- (auto-) correlations between redshift bins which must be incorporated into any realistic tomographic analysis. Such leakage may cause inaccuracies or biases in lensing measurements unless we are able to accurately model the IA signal both within and between bins. BK07 find that the IA signal is much more sensitive to information loss in the presence of photo-z uncertainties than the GG signal, thus the combined effect of photo-z errors and an inaccurate IA model may considerably degrade parameter constraints.

Previous lensing analyses have been able to make simple assumptions about IAs and photo-z errors because the surveys were small and the statistical errors tended to dominate the analyses. In the future, however, large surveys with billions of galaxies and progressively decreasing statistical errors will become the norm. CFHTLenS has a survey area approximately two orders of magnitude

larger than fields such as HST COSMOS and CFHTLS-Deep used in previous tomographic analyses, and future surveys such as Pan-STARRS, KIDS, DES and Euclid will have even larger deep imaging survey areas on the order of thousands to tens of thousands of square degrees (see §2.3.6). In such large data sets, systematic errors will be the limiting factor in parameter estimation if unaccounted for, as high galaxy counts ensure the statistical errors will be small. Tomography will be an invaluable technique in analysing such surveys as it is particularly sensitive to dark energy parameters. However, IAs are also most prominent in a tomographic analysis and care must be taken to account for their effects [198]. The effect of uncertainties in the photometric redshift errors must also be well-known in order to accurately predict parameters and avoid biases in cosmological parameter constraints. The interplay between errors in lensing analysis requires a more sophisticated treatment that includes a realistic photo-z error distribution.

To produce a full and realistic covariance matrix that includes non-Gaussianity and contributions from IAs and shot noise for every redshift correlation (with realistic photo-z errors), we must estimate it from tomographic shear simulations. Additionally, IAs are most significant on small scales where galaxies are close on the sky. Most N-body simulations do not have high enough resolution to incorporate an IA model on the scales of interest. Those that do are slow to generate because of the higher particle density, meaning the number of lines of sight available to us is low and our estimators may be noisy, hence high-bin tomography will be difficult since the limiting ratio of $p/N = 1$ is reached at lower n_z (see §4.1.2). For these reasons, to investigate the effect of IAs on the shear correlation function and the covariance we turn to the Gaussian simulations of [176], which incorporate the NLA model. As with the shear from these simulations, the errors on the IA signal will be Gaussian and hence almost certainly underestimated. However, unlike the error on the shear, which is well predicted by N-body simulations, we do not know how the error on the IA signal behaves. Any modification to Gaussian errors would therefore be arbitrary, so we use Gaussian covariances for the II and GI alignments in this analysis.

In this chapter, we alter the shear correlation function estimated from the clone simulations in Chapter 4 to include a realistic contribution from shot noise. We also use the Gaussian simulations to incorporate the expected II and GI

alignment signals predicted from the NLA model. We investigate the relative contribution to the total shear correlation function[1], $\hat{\xi}^{\rm TOT}$, of each of these components. We construct covariance estimators for each correlation function component and for the total covariance in preparation for the likelihood analysis in Chapter 6. We also examine the contribution of each component to the correlation function and the covariance. Finally, we investigate the invertibility of the total covariance estimator, $\hat{\mathbf{C}}_{\rm TOT}$ with tomography to determine whether it is possible to use as many tomographic bins in our analysis as we concluded in Chapter 4.

5.2 The Non Linear Alignment Model

In this analysis we utilise the nonlinear alignment model for IAs proposed in [119], developed by BK07 and tested in [199]. This is a simplified version of the linear alignment model proposed in [196]. The model is motivated by the assumption that the intrinsic shear of a galaxy is proportional to the Newtonian potential that it experiences when it forms, since halo ellipticities are perturbed by the local tidal field produced by large scale structure. On large scales, correlations within the intrinsic field are determined by large-scale potential fluctuations; if these potential fluctuations are small then the intrinsic field will be a linear function of the potential and higher-order terms will be negligible. BK07 use only the first term of Eqn. 16 in [119], including the erratum in [200], to represent the II power spectrum,

$$P^{\rm lin}_{\tilde{\gamma}^I}(k) = \frac{C_1^2\bar{\rho}^2}{\bar{D}^2}a^4 P^{\rm lin}_{\delta}(k), \tag{5.1}$$

as this term is approximately ten times larger than the second term, which is ignored. Here, $P^{\rm lin}_{\tilde{\gamma}^I}(k)$ is the II power spectrum, C_1 is a normalisation constant, $\bar{\rho}$ is the mean matter density of the Universe as a function of redshift, $\bar{D} \equiv (1+z)D(z)$ where $D(z)$ is the growth factor of the Universe normalised to the present day and $P^{\rm lin}_{\delta}(k)$ is the linear theory matter power spectrum. The NLA model replaces $P^{\rm lin}_{\delta}(k)$ with the full nonlinear matter power spectrum $P_{\delta}(k)$ in Eqn. 5.1,

$$P^{\rm nl}_{\tilde{\gamma}^I}(k) = \frac{C_1^2\bar{\rho}^2}{\bar{D}^2}a^4 P_{\delta}(k). \tag{5.2}$$

[1]As in Chapter 4 we will only refer to ξ_+ in this chapter onwards, so for notational simplicity we drop the '+' subscript.

Eqns. 5.1 & 5.2 show that the II power spectrum has a very strong dependence on redshift; since $a \propto (1+z)^{-1}$ and $\bar{D} \propto (1+z)$, then $P^{\rm nl}_{\tilde{\gamma}^I}(k) \propto (1+z)^{-6}$. This surprisingly strong redshift-dependence means that the II signal declines very rapidly at high redshift. Equivalently for the GI power spectrum, [119] predict

$$P^{\rm lin}_{\delta\tilde{\gamma}^I}(k) = -\frac{C_1\bar{\rho}}{\bar{D}}a^2 P^{\rm lin}_{\delta}(k) \tag{5.3}$$

which in the NLA model becomes

$$P^{\rm nl}_{\delta\tilde{\gamma}^I}(k) = -\frac{C_1\bar{\rho}}{\bar{D}}a^2 P_{\delta}(k). \tag{5.4}$$

Here, the redshift dependence is $P^{\rm nl}_{\delta\tilde{\gamma}^I}(k) \propto (1+z)^{-3}$, which is the square root of the II dependence. This is to be expected since it is the intrinsic field that carries the strong redshift dependence, and the GI power spectrum comes from correlating the shear and intrinsic fields, whereas the II power spectrum is the result of correlating the intrinsic field with itself.

BK07 find that the NLA model matches better than the linear alignment model to the model for II alignments in [190] and to the GI contamination found in [186] and [187]. Importantly, the NLA model takes into account nonlinearity in the power spectrum at small scales which are expected to arise if galaxies are aligned within their own dark matter haloes (even in the absence of II alignments). It assumes that the II and GI power spectra are independent of each other and that the correlation of galaxy shapes using the component of the shear aligned along the line connecting galaxy shapes is the same as the correlation at 45 degrees to the line joining the galaxies.

BK07 estimate a fiducial value for $C_1 = 5 \text{ x } 10^{-14}$ $(h^2 \mathrm{M}_\odot/\mathrm{Mpc}^{-3})^{-2}$ from superCOSMOS data [119, 197]. We adopt the parametrisation of [199], who fix C_1 to its fiducial value and introduce a new dimensionless amplitude parameter, A, so that

$$\begin{aligned} P^{\rm nl}_{\tilde{\gamma}^I}(k) &= A\frac{C_1^2\bar{\rho}^2}{\bar{D}^2}a^4 P_{\delta}(k) \\ P^{\rm nl}_{\delta\tilde{\gamma}^I}(k) &= -A\frac{C_1\bar{\rho}}{\bar{D}}a^2 P_{\delta}(k). \end{aligned} \tag{5.5}$$

A is equivalent to

$$A = \frac{C_1 \rho_c}{0.0134} \tag{5.6}$$

since $C_1 \rho_c \approx 0.0134$ for the fiducial value of C_1 in BK07. ρ_c is the critical density defined in Eqn. 1.17. Hence $A = 1$ for the fiducial value of C_1, and A has the advantage of being dimensionless. [199] considered different galaxy types in MegaZ and SDSS and found that they are consistent with each other under a three-parameter NLA model, however the best-fit amplitude of the normalisation constant A varies with galaxy type, a result also found by [187] in 2SLAQ and SDSS data. As a result, the authors fit the NLA model to the different galaxy samples allowing for redshift and luminosity dependence. They find no redshift evolution beyond that of the original NLA model and a trend towards increasing IA amplitude with galaxy luminosity. Fig. 11 in [199] recovers the amplitude of A for different galaxy types. For Luminous Red Galaxies (LRGs), the best-fit value of A is as high as $16.09^{+2.75}_{-2.76}$ (1σ), however for non-LRG galaxies the best-fit value of A is as low as $1.20^{+0.90}_{-0.88}$ (1σ) and hence consistent with the results from BK07 and with zero at the 2σ level. LRGs are bright and large but rare, and do not contribute significantly to the galaxy population of CFHTLenS. [201] and [186] find that the IA signal from low-to intermediate redshift galaxies is consistent with zero. It is possible that by recent times most galaxies have lost their alignments due to mergers and other interactions, meaning the IA signal from galaxies at high-z may be more substantial. We generate simulations incorporating the fiducial value of $A = 1$ in the NLA model. If there is significant tension in the likelihood contours for A when fitting to CFHTLenS data, this may indicate that there has been significant evolution of galaxy alignments between those measured at low redshift and those of the CFHTLenS median redshift.

5.2.1 Incorporating the NLA model in the Gaussian simulations

We wish to measure the NLA II and GI alignment signals for a CFHTLenS-like survey, in order to combine them with the shear signal. In this section, we describe how the simulations of [176] are used to generate these signals. To construct shear and IA fields corresponding to a given cosmology, the simulation code requires knowledge of the lensing and IA power spectra for every redshift

auto- and cross-correlation prediction for that cosmology. `NICAEA` (see §3.6) was modified to produce 2D power spectra and correlation function predictions for the NLA model II and GI alignment signal (see §3.6). Power spectra were calculated using `NICAEA` over a range of multipole moments, l, from $l = 1$ (equivalent to the all-sky mode) up to a maximum of $l = 3000$. Setting l equal to the maximum value of 3000 ensures better accuracy on small scales, but is more computationally expensive. The input power spectra are then used to construct a global power spectrum matrix for every multipole moment for redshift bins 1 to n_z with the following structure:

$$\mathbf{C}_l = \begin{pmatrix} G_1G_1 & \cdots & G_1G_{n_\mathrm{z}} & G_1I_1 & \cdots & G_1I_{n_\mathrm{z}} \\ \vdots & \ddots & & \vdots & \ddots & \\ G_{n_\mathrm{z}}G_1 & & G_{n_\mathrm{z}}G_{n_\mathrm{z}} & G_{n_\mathrm{z}}I_1 & & G_{n_\mathrm{z}}I_{n_\mathrm{z}} \\ I_1G_1 & \cdots & I_1G_{n_\mathrm{z}} & I_1I_1 & \cdots & I_1I_{n_\mathrm{z}} \\ \vdots & \ddots & & \vdots & \ddots & \\ I_{n_\mathrm{z}}G_1 & & I_{n_\mathrm{z}}G_{n_\mathrm{z}} & I_{n_\mathrm{z}}I_1 & & I_{n_\mathrm{z}}I_{n_\mathrm{z}}, \end{pmatrix} \tag{5.7}$$

where G represents the cosmic shear and I the intrinsic ellipticity.

In the NLA model, the G and I fields are 100% anti-correlated, which means the 3D power spectra fulfil the condition that $P_{\delta\tilde{\gamma}^I}(k) = -\sqrt{P_{\tilde{\gamma}^I}(k)P_\delta(k)}$ [176]. In reality, this is unlikely to be the case due to the complicated nature of galaxy formation processes, and this high degree of correlation can cause $\mathbf{C}_l$ to be non positive-definite. If, at any multipole, $\mathbf{C}_l$ is not positive-definite, it is replaced with the 'closest' matrix that is positive-definite to remove singular or close-to-singular modes. This is done by replacing all below-machine-precision Eigenvalues with ones equal to machine precision (see §5.4.2 for an explanation). $\mathbf{C}_l$ is then Cholesky decomposed. This is the decomposition of a positive-definite matrix into triangular matrices which can be used to solve a system of linear equations and pseudo-invert a matrix [202]. We note that a non positive-definite $\mathbf{C}_l$ may be physically permissible, but such a matrix can not be reliably inverted or Cholesky decomposed, making the production of shear fields from such a matrix extremely difficult. Decomposing $\mathbf{C}_l$ gives

$$\mathbf{C}_l^{xy} = \sum_z L_l^{xz} L_l^{yz} \tag{5.8}$$

where x, y are the indices of $\mathbf{C}_l$. The multipoles L_l^{xy} are used to generate random realisations of the spin-2 spherical harmonic coefficients of each field from a set of random Gaussian deviates G_{lm}^x,

$$\begin{aligned} a_{l0}^x &= \sum_y L_l^{xy} G_{l0}^y, \\ a_{lm}^x &= \sqrt{\frac{1}{2}} \sum_y L_l^{xy} G_{lm}^y. \end{aligned} \tag{5.9}$$

These are then transformed to real space via a spin-2 transform using the `HEALPIX` package[2], such that the fields lie in a strip spanning ± 9 degrees around the equator. The code then projects the simulation onto flat, pixelated shear maps. The simulation output consists of one shear (G) and one intrinsic (I) field for each redshift bin. The spin-2 transform ensures that the G and I fields are correlated.

In addition to the constraint on the 3D power spectrum due to the 100% anti-correlation between G and I fields, [176] find that for the 2D projected power spectra to be positive definite, the further constraint $\left|\mathbf{C}_l^{GI}\right| \leq \sqrt{\mathbf{C}_l^{GG}\mathbf{C}_l^{II}}$ must be met for every value of l. For certain redshift distributions this constraint is not met, resulting in non positive-definite power spectrum matrices which cannot be Cholesky decomposed. [176] introduce a 3D correlation coefficient to allow one to decrease the degree of anti-correlation between I and G fields. We utilise a 2D correlation coefficient $0 \leq B \leq 1$, analogous to this (since we use the 2D power spectra produced by `NICAEA` as input):

$$\left|B\,\mathbf{C}_l^{GI}\right| \leq \sqrt{\mathbf{C}_l^{GG}\mathbf{C}_l^{II}}. \tag{5.10}$$

In the unaltered NLA model, $B = 1$, so we want B to be as close to this value as possible to match the NLA model as closely as possible. We calculate the maximum value of $|B|$ that satisfies this relation for all tomographic redshift distributions in this analysis to be 0.5. Therefore the G and I fields will be 50% anti-correlated, meaning that the GI signal will be decreased by 50% relative to the standard NLA model. Failure to do this would mean that $\mathbf{C}_l$ would be replaced with the closest-positive-definite matrix for at least some modes and

[2]Hierarchical Equal Area isoLatitude Pixelization of a sphere, a package that produces subdivisions of a spherical surface and can be used to produce (non-flat) sky maps. See `http://healpix.jpl.nasa.gov/`.

the power spectra used to generate the simulations would differ from the input, which may cause us to misestimate both the resulting correlation function and its covariance. This 50% anti correlation is used throughout the rest of this thesis, and is used to calculate the theoretical amplitude of the GI correlation function that we fit CFHTLenS data to in Chapter 7.

We create a suite of simulations using this technique and analyse the GG, II and GI correlation functions by populating the simulations with galaxies, which we describe in §5.3.3. Unfortunately we found that for the redshift distribution used, there was a numerical limit to the accuracy of the code such that the resulting GG and II correlation functions are slightly higher than the input theory for that cosmology by around 10^{-7} at all angular scales. This effect is particularly significant for the II signal as the amplitude of the II correlation function is of the order 10^{-7} on the scales of interest. This problem occurs because when the GI power spectra (and hence the GI components of $\mathbf{C}_l$) are non-zero, the Cholesky decomposition of $\mathbf{C}_l$ becomes unreliable. The power spectrum matrix contains very small terms, and including a GI component means the matrix also becomes strongly off-diagonal. The precision with which we are able to pseudo-invert the matrix is not high enough (limited by machine precision), resulting in unreliable values of the shear and IA amplitude in the pixelated fields. We find, however, that the GI correlation function extracted from such fields is unaffected. This is because the modification to the G fields by this effect is uncorrelated from that of the I fields, so that when the fields are correlated together it does not affect our measurements. Since the GG and II correlation functions are well behaved in the absence of a GI signal in $\mathbf{C}_l$ (as we show in §5.3.4), we produce two runs of the Gaussian simulations - one without and one with a GI power spectrum, to enable us to measure the II and GI correlation functions respectively from the separate runs.

5.3 The total shear correlation function

In lensing data, the presence of IAs will alter our measurement of the shear correlation function and its covariance. Additionally, uncertainties in galaxy shape measurements will increase many of the terms in the covariance matrix. We describe in this section how we construct an estimate of the total correlation

function from simulations that is equivalent to one measured from real data.

5.3.1 Constructing a realistic correlation function estimator

To estimate the total shear correlation function, $\hat{\xi}^{\rm TOT}$, we must understand the different signals that contribute to it. In the weak lensing limit, the observed ellipticity of a galaxy, ϵ_i will be a combination of the shear γ_i and the total intrinsic ellipticity, $\epsilon^s_{{\rm TOT},i}$,

$$\epsilon_i = \gamma_i + \epsilon^s_{{\rm TOT},i}\,. \tag{5.11}$$

The total intrinsic ellipticity consists of two parts: an intrinsic ellipticity component that is correlated with the density field, $\epsilon^s_{\delta,i}$, and an uncorrelated ellipticity component that incorporates the shape measurement error, n_i. So the observed ellipticity is

$$\epsilon_i = \gamma_i + \epsilon^s_{\delta,i} + n_i. \tag{5.12}$$

Of these terms, n_i is the largest (of the order 10^{-1}), the shear is of the order 10^{-3}, or $\sim 1\%$ of this, and $\epsilon^s_{\delta,i}$ is smaller still on the order of 10^{-4}.

The observed ellipticity correlation will be

$$\begin{aligned}\langle\epsilon_i\epsilon^*_j\rangle_\theta &= \langle(\epsilon^s_{\delta,i}+\gamma_i+n_i)(\epsilon^{s*}_{\delta,j}+\gamma^*_j+n^*_j)\rangle_\theta \\ &= \langle\gamma_i\gamma^*_j\rangle_\theta + \langle\epsilon^s_{\delta,i}\epsilon^{s*}_{\delta,j}\rangle_\theta + \langle\gamma_i\epsilon^{s*}_{\delta,j}\rangle_\theta + \langle\epsilon^s_{\delta,i}\gamma^*_j\rangle_\theta + \langle n_i n^*_j\rangle_\theta \\ &\quad +\langle\gamma_i n^*_j\rangle_\theta + \langle n_i\gamma^*_j\rangle_\theta + \langle\epsilon^s_{\delta,i}n^*_j\rangle_\theta + \langle n_i\epsilon^{s*}_{\delta,j}\rangle_\theta. \end{aligned}\tag{5.13}$$

i and j correspond to galaxies in different tomographic redshift bins ($i \leq j$). Using the identity in Eqn. 2.51 we can equivalently write:

$$\begin{aligned}\hat{\xi}^{\rm TOT} &= \hat{\xi}^{GG} + \hat{\xi}^{II} + \hat{\xi}^{GI} + \hat{\xi}^{IG} + \hat{\xi}^{NN} + \\ &\quad \hat{\xi}^{GN} + \hat{\xi}^{NG} + \hat{\xi}^{IN} + \hat{\xi}^{NI}, \end{aligned}\tag{5.14}$$

where G, I and N refer to the shear, correlated intrinsic and uncorrelated noise components of the total ellipticity respectively. We define our notation such that e.g. $\hat{\xi}^{GI}$ refers to the correlation between a low-z G field and an I field from the same or higher redshift. Since the shot noise is uncorrelated, we expect all terms in $\hat{\xi}^{\rm TOT}$ containing a shot noise component (an 'N' term) to be consistent with

zero. $\hat{\xi}^{IG}$ may be non-zero due to the overlapping nature of the tomographic redshift bins; some galaxies in the high-z bin may be at a lower redshift than galaxies in the low-z bin, leading to an apparent reversal of the signal. We will be able to estimate the value of $\hat{\xi}^{\rm TOT}$ if we are able to measure each of the individual components in Eqn. 5.14.

5.3.2 Shot noise

The shot noise inherent in measuring galaxy shapes is an important component of the uncertainty in lensing measurements. To ensure the shot noise is modelled as realistically as possible, the average shot noise for the full 154 sq. degrees of CFHTLenS data was used to determine the components of $\hat{\xi}^{\rm TOT}$ (and its covariance) that contain a shot noise contribution. The shot noise was determined for both ellipticity components ϵ_1 and ϵ_2 from the $N_{\rm gal}$ survey galaxies as

$$\sigma_\epsilon^2 = \frac{\sum_{j=1}^{N_{\rm gal}} w_j^2 \epsilon_j^2}{\sum_{j=1}^{N_{gal}} w_j^2} \tag{5.15}$$

where w is the `Lensfit` weight assigned to each galaxy. The weights are related to how reliably `Lensfit` can assign a shape measurement to a given galaxy, and span an arbitrary range in the CFHTLenS catalogue from 0 to 15.6.

The average shot noise was found to be $\sigma_{\epsilon_1} = \sigma_{\epsilon_2} = 0.28$. Both ellipticity components have identical shot noise, which is expected as the shot noise is uncorrelated with direction. We also investigated dependence of σ_{ϵ_i} on photometric redshift and found no significant redshift dependence. Since the shot noise is uncorrelated, the value of σ_ϵ represents the Gaussian standard deviation in each intrinsic ellipticity component from CFHTLenS galaxies.

In lensing data, the total shot noise observed is due to the combined correlated and uncorrelated intrinsic ellipticities, so that $\sigma_\epsilon = \epsilon_\delta^s + n$. However, for simplicity we assume $\sigma_\epsilon = n$ and hence overestimate the size of the uncorrelated shot noise. n is much larger in magnitude than ϵ_δ^s (by a factor of $\sim 10^3$ as measured from the Gaussian I fields, see previous section) meaning that we can safely ignore its contribution to the total noise in this way.

5.3.3 Producing galaxy catalogues

As in Chapter 4, the lensing simulations and corresponding galaxy catalogues we generate use the CFHTLenS PDF redshift distribution and galaxy density. We generate $n = 736$ Gaussian shear and intrinsic fields as described in §4.2.2 and §5.2.1 to match the N-body simulations in number and input cosmology. We use only a portion of the fields with equivalent area to the clone lines of sight (after dividing into subfields) of 2.70 square degrees, so that they match the clones in area and number.

We produce galaxy catalogues in a similar manner to that described in §4.2.3. We construct three separate galaxy catalogues: a shear catalogue which we denote G, a correlated intrinsic ellipticity catalogue I and an uncorrelated shot noise catalogue N. These three catalogues contain clone galaxy positions and ellipticities corresponding to each of the three components of the total observed ellipticity ϵ_i described in Eqn. 5.12 (γ_i for the G catalogue, $\epsilon^s_{\delta,i}$ for the I catalogue and n_i for the N catalogue). We have already constructed G catalogues from the N-body and Gaussian simulations in the previous chapter. The N catalogues contain galaxy positions that map exactly to the G catalogues, with randomly assigned shot noise n_i taken from a 1D Gaussian with width $\sigma_{\epsilon_i} = 0.28$ per ellipticity component. Gaussian intrinsic ellipticity fields are used to create the I catalogues by measuring the strength of the intrinsic field at galaxy positions corresponding to the G catalogues (from the Gaussian simulations). We use the Gaussian G catalogues only to calculate the amplitude of the GI correlation function; the clone G catalogues are used to calculate the GG correlation function (see next section).

5.3.4 Measuring $\hat{\xi}^{\mathrm{TOT}}$

We estimate all the correlation function terms in Eqn. 5.14 separately in order to find $\hat{\xi}^{\mathrm{TOT}}$. The estimator for $\hat{\xi}^{GG}$ was measured from the clone simulations as described in Chapter 4 . Every other term in Eqn. 5.14 was found by using `ATHENA` (see §4.3) to measure the correlations between galaxies on different angular scales within and between the G, I and N catalogues. Ensuring the galaxy positions in the G, I and N catalogues are the same allows us to easily and reliably estimate the cross terms in Eqn. 5.14. We note that to calculate $\hat{\xi}^{GI}$ and $\hat{\xi}^{IG}$, it is necessary

to measure the correlations from G and I fields that were created simultaneously from the Gaussian simulation code, since the fields would then be correlated with each other as described in §5.2.1.

We plot $\hat{\xi}^{\rm TOT}$ as well as some of the terms in Eqn. 5.14 for one and five tomographic bins in Figs. 5.1-5.2. In general we find good agreement with theory for $\hat{\xi}^{GG}$, $\hat{\xi}^{II}$, $\hat{\xi}^{GI}$ and the total correlation function with the caveat that at large scales the clones lose power as described in §4.3. $\hat{\xi}^{\rm TOT}$ is lower than $\hat{\xi}^{GG}$ because at all scales, the GI (and IG) correlation is negative. We do not plot $\hat{\xi}^{IG}$ as this is exactly equal to $\hat{\xi}^{GI}$ in auto-correlation bins (also $\hat{\xi}^{GN} = \hat{\xi}^{NG}$, $\hat{\xi}^{IN} = \hat{\xi}^{NI}$). $\hat{\xi}^{NN}$ and all noise cross terms are consistent with zero at all scales, as expected since the noise is uncorrelated in both angular scale and redshift. We only plot $\hat{\xi}^{NN}$ since the variance on this term is larger than for any of the noise cross terms.

For $n_{\rm z} = 5$, the IA signal is behaving as expected. $\hat{\xi}^{II}$ is highest in auto-correlation and in the lowest redshift bins (since at low z, a given angular scale represents a smaller physical separation than at high z, hence the strength of the alignment will be greater). The noise terms in the correlation function are still consistent with zero, however the variance on them is large and hence dominates our measure of $\hat{\xi}^{\rm TOT}$ for low-z correlations. This is because at low redshift, the shear signal is low and difficult to detect in the presence of shot noise, which is unaffected by redshift[3]. It is worth noting that $\hat{\xi}^{II}$ ($\hat{\xi}^{GI}$) in cross (auto) correlations is non-zero. This is because of the highly overlapping nature of the redshift bins highlighted in Fig. 4.2. If two different tomographic bins overlap, it is possible to get a significant II signal in the cross-correlation between them since some of the galaxies in bin i will have similar redshifts to some of the galaxies in bin j. Similarly, catastrophic photo-z errors mean that galaxies from within one redshift bin can have quite different redshifts, leading to a significant GI signal even in auto-correlation.

[3]Strictly, this is only true if the tomographic redshift bins contain the same number of galaxies, as in this analysis. For tomographic bins with equal width, there will be more galaxies in the high-z bins as they cover a larger volume so the shot noise will decrease.

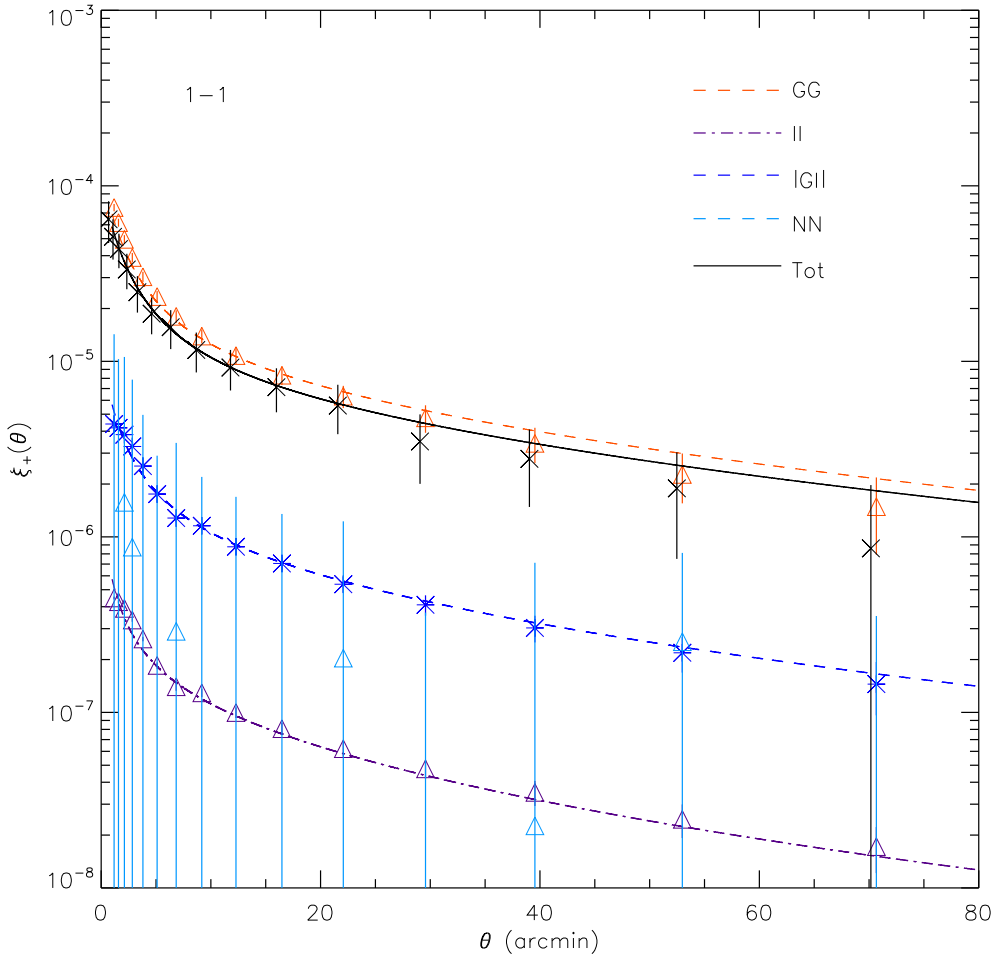

Figure 5.1: GG and NN correlation functions measured from the N-body simulations, and GI and II measured from the Gaussian fields (with the same redshift distribution, galaxy density and input cosmology) for $n_z = 1$ and $n_\theta = 15$. Points correspond to values measured from simulations, and error bars are found from the standard deviation between all 736 fields rescaled to match the standard deviation expected from a CFHTLenS-like survey (see §4.4.1). Theory curves plotted as lines are found from `NICAEA`. Also plotted is $\hat{\xi}^{\rm TOT}$ and its theoretical value (found from summing the individual components from `NICAEA`); the points have been offset slightly to visually differentiate them from $\hat{\xi}^{GG}$.

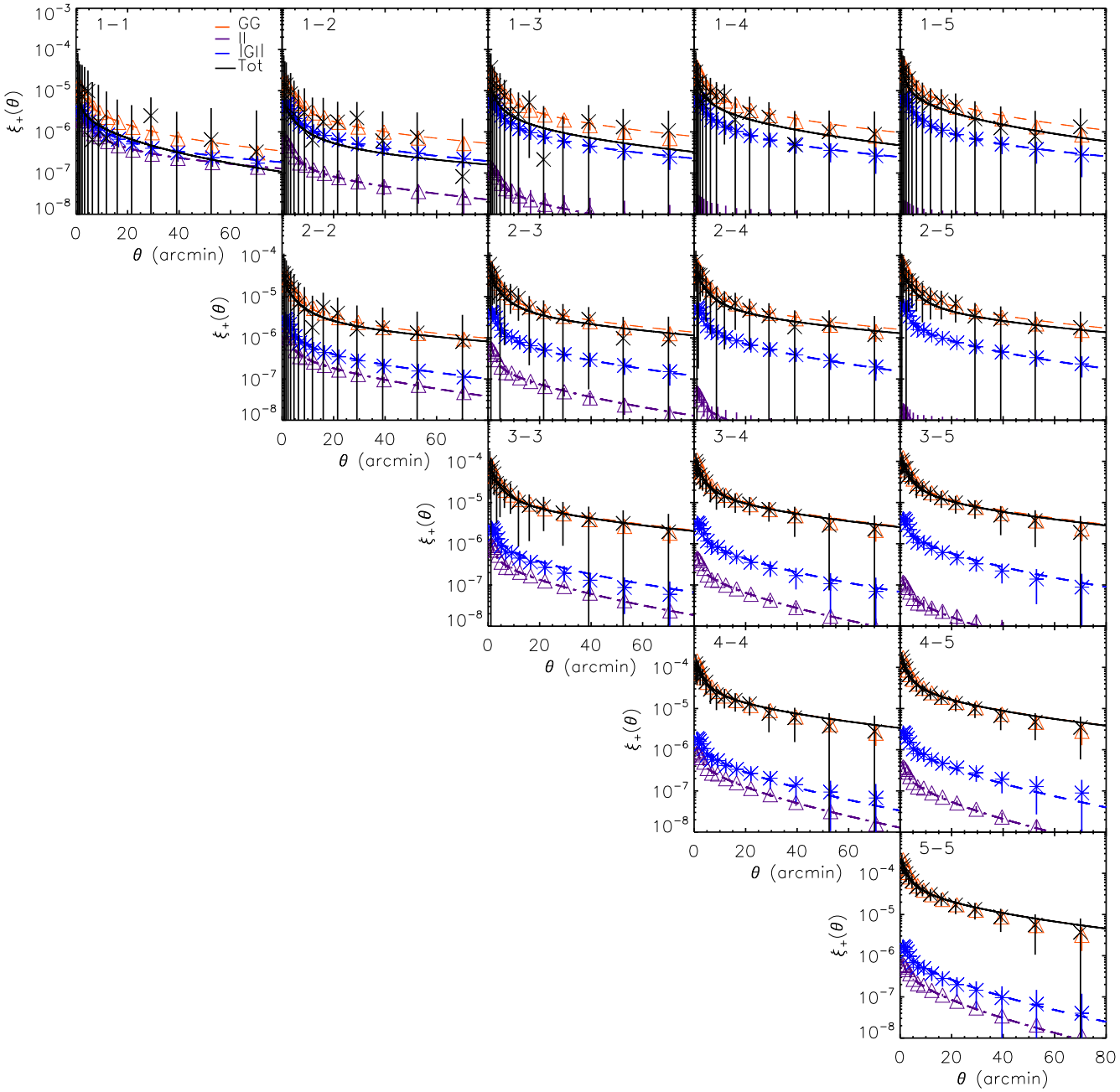

Figure 5.2: As in Fig. 5.1 for $n_z = 5$, $n_\theta = 15$. $\hat{\xi}^{NN}$ has been omitted for clarity. The redshift correlation being plotted is shown in the top left of each subplot. $\hat{\xi}^{II}$ is highest in auto-correlations and $\hat{\xi}^{GI}$ is strongest for far cross-correlations. Except in the lowest bins, the cosmic shear correlation dominates over the IA terms.

5.4 The total covariance

As we have an estimate of the total correlation function, $\hat{\xi}^{\rm TOT}$, we can construct a covariance matrix which incorporates the uncertainty in every term in Eqn. 5.14 by simply modifying Eqn. 4.1 to include $\hat{\xi}^{\rm TOT}$:

$$\hat{\mathbf{C}}_{ij}^{\rm TOT} = \frac{1}{N-1}\sum_{k=1}^{N}\left(\hat{\xi}_i^{{\rm TOT},(k)} - \bar{\xi}_i^{\rm TOT}\right)\left(\hat{\xi}_j^{{\rm TOT},(k)} - \bar{\xi}_j^{\rm TOT}\right). \tag{5.16}$$

Here, i, j represent the dimension of the data vector. It is important to note that $\hat{\mathbf{C}}^{\rm TOT}$ is not equal to the sum of the covariance of the individual correlation functions, that is

$$\begin{aligned}\hat{\mathbf{C}}^{\rm TOT} \neq\ & \hat{\mathbf{C}}^{GG,GG} + \hat{\mathbf{C}}^{II,II} + \hat{\mathbf{C}}^{GI,GI} + \hat{\mathbf{C}}^{IG,IG} + \hat{\mathbf{C}}^{NN,NN} + \\ & \hat{\mathbf{C}}^{GN,GN} + \hat{\mathbf{C}}^{NG,NG} + \hat{\mathbf{C}}^{IN,IN} + \hat{\mathbf{C}}^{NI,NI},\end{aligned} \tag{5.17}$$

where $\hat{\mathbf{C}}^{GG,GG}$ is the covariance of $\hat{\xi}^{GG}$ and so on. This is because the covariance matrix will also contain contributions from cross terms between the correlation functions in Eqn. 5.14. For example, there will be a $\hat{\mathbf{C}}^{GG,II}$ term that arises from the covariance between $\hat{\xi}^{GG}$ and $\hat{\xi}^{II}$. There are 36 of these covariance cross-terms, and it is simpler to calculate $\hat{\mathbf{C}}^{\rm TOT}$ directly from $\hat{\xi}^{\rm TOT}$, which naturally accounts for the total covariance.

We plot the diagonals of the individual covariances of the correlation functions in Eqn. 5.14 and $\hat{\mathbf{C}}^{\rm TOT}$ for $n_{\rm z} = 1$, 5 and 10 in Figs. 5.3-5.5. All terms in the covariance matrix are scaled to CFHTLenS area as described in §4.4.1. For $n_{\rm z} = 1$, we find that the dominant contribution to the total covariance on scales of $\lesssim 30'$ is the shot noise through the variance on $\hat{\xi}^{NN}$, as expected. Above this scale, sample variance is the dominant contributor to the covariance through the GG term. For higher bin tomography, however, the shot noise dominates in every redshift correlation, and its amplitude is independent of redshift, as expected since the number of galaxies per redshift bin is constant. This result is significant; the covariance of the noise (and of the cross terms between the N and G catalogues) is the largest contributor to the total covariance at all angular scales and in all redshift bins, by up to two orders of magnitude. Thus, inclusion of a realistic shot noise component in the covariance is vital if we are to construct a covariance matrix that can be used to constrain cosmology.

Terms containing a contribution from the I field have very low variance for all n_z. This means that explicit calculation of the covariance matrix contribution from IAs may not be necessary as the shear and noise covariance terms are larger by a factor of 10^1-10^4 in almost every redshift correlation, and the covariance calculated from the sampling variance and the shot noise alone is a good approximation to the total covariance. The caveat to this is that if the amplitude of the normalisation constant A in the NLA model is significantly higher than 1, the I covariance terms may become significant. For example, if $A = 10$ (e.g. for low-z LRGs as found in [199]), the amplitude of $\hat{\xi}^{II}$ and its covariance will increase by a factor of $\sim A^2 = 100$, making it a significant contributor to the total covariance, particularly in low-z bins. As LRGs make up such a small fraction of the galaxy population however, it is unlikely that A will be this high on average.

It is worth noting that since the amplitude of the covariance matrix scales inversely with survey area, the relative amplitudes of each of the covariance terms in Figs. 5.3-5.5 will be the same for surveys of different areas for a given redshift distribution and galaxy density. This means that the results presented in this thesis are applicable to future large-scale lensing surveys, if the galaxy density and redshift distributions are similar.

In Figs. 5.6-5.7 we plot the ratio of the total covariance to that found from the sum of the individual covariances in Eqn. 5.17 for $n_z = 1$, 5 and $n_\theta = 15$. Error bars are the 1σ standard deviation on $\hat{\mathbf{C}}^{\rm TOT}$ and were estimated from bootstrapping to generate multiple covariance matrices as described in §4.4.2. This plot shows the significance of the cross terms in the covariance matrix ($\hat{\mathbf{C}}^{GG,II}$ etc.). At angular scales above $\sim 7'$, the ratio deviates from one, implying that the cross terms are contributing significantly to the covariance matrix. At high angular scales, the covariance matrix is boosted by up to $\sim 25\%$ by these terms. We find this trend for all values of n_z. Figs. 5.3-5.4 show that most of the covariance comes from the G and N correlations, so it is likely that the cross terms from these correlations (such as $\hat{\mathbf{C}}^{GG,NN}$, $\hat{\mathbf{C}}^{GN,NN}$) are responsible for most of the additional variance. [203] perform a similar 1-bin analysis of the clone simulations (without IAs) and also find the cross terms between G and N to be subdominant

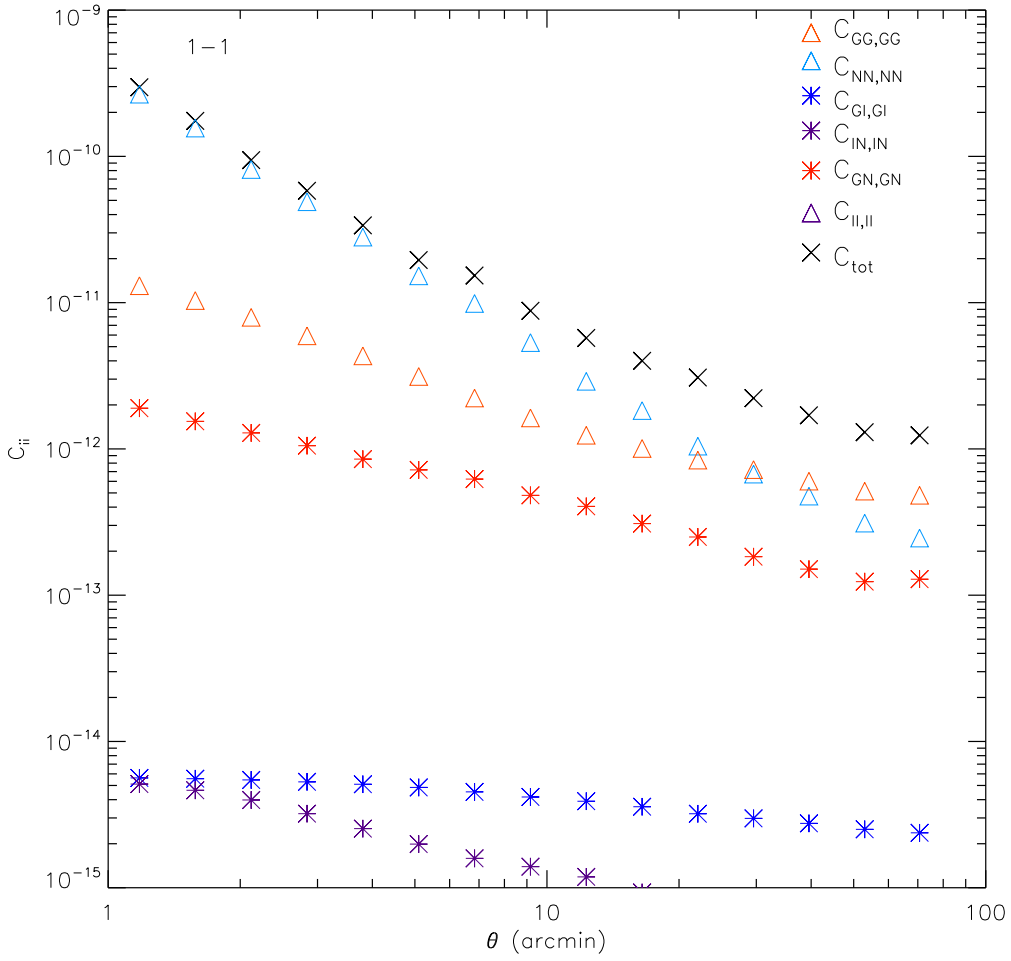

Figure 5.3: Diagonal of the total covariance $\hat{\mathbf{C}}^{\mathrm{TOT}}$ as a function of angular scale for $n_z = 1$, $n_\theta = 15$. Also plotted are some of the individual components of the covariance as defined in Eqn. 5.17. Where the covariance terms are symmetric in the auto-correlation, for example $\hat{\mathbf{C}}^{IG,IG} = \hat{\mathbf{C}}^{GI,GI}$, we plot only one of the two terms. $\hat{\mathbf{C}}^{II,II}$ is of the order 10^{-18} and is not visible on the plot. The redshift correlation being plotted is shown at the top of the plot.

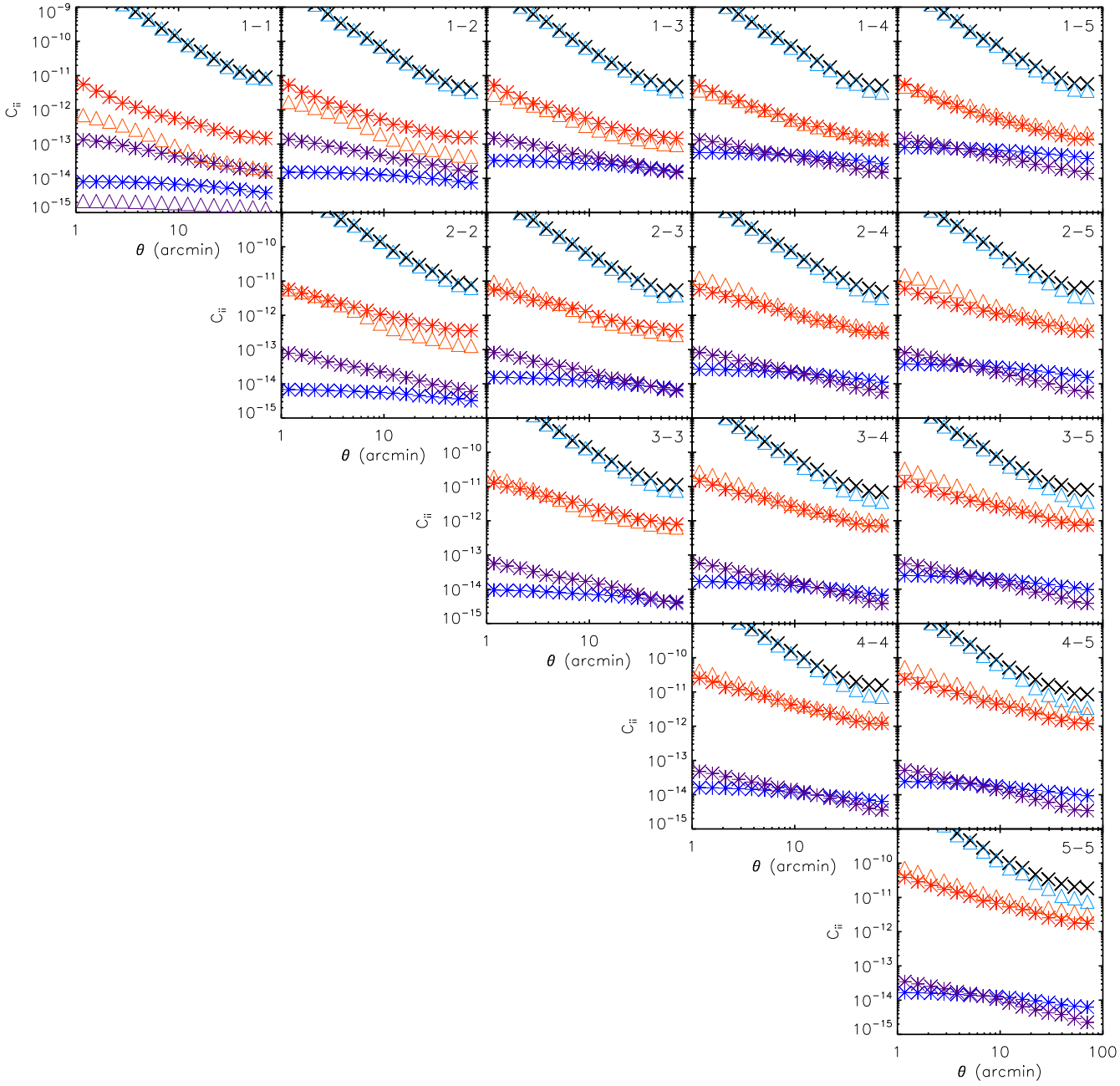

Figure 5.4: As in Fig. 5.3 for $n_z = 5$, $n_\theta = 15$. Legend is the same as in Fig. 5.3 and has been omitted to avoid crowding.

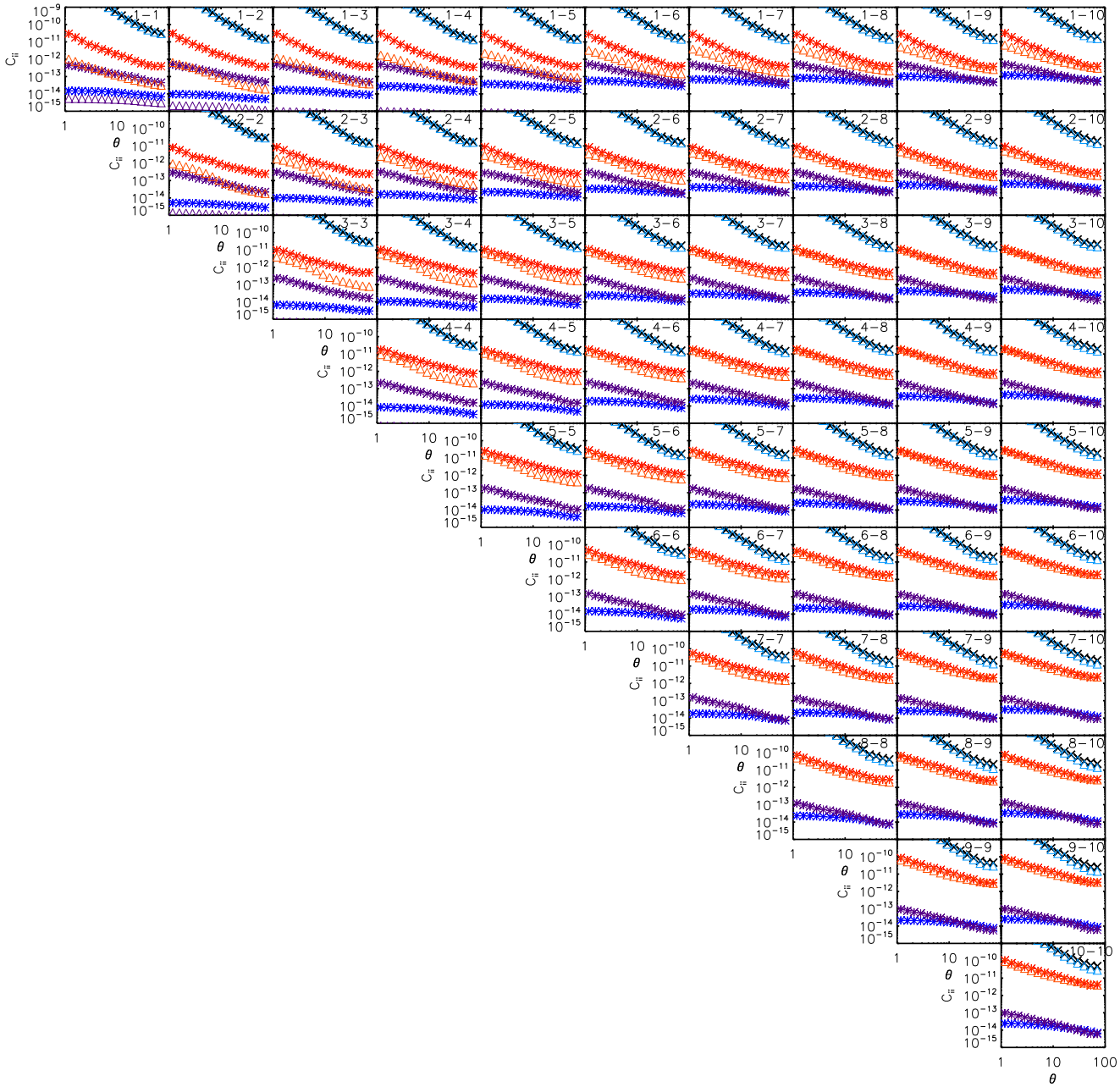

Figure 5.5: As in Fig. 5.3 for $n_z = 10$, $n_\theta = 15$. Legend is the same as in Fig. 5.3 and has been omitted to avoid crowding.

to $\hat{\mathbf{C}}^{GG,GG}$ and $\hat{\mathbf{C}}^{NN,NN}$, but still a potentially important contributor to the total noise. The effect of these terms cannot therefore be neglected, and our implementation of $\hat{\mathbf{C}}^{\rm TOT}$ as an estimator for the total covariance is necessary. Interestingly, for $n_{\rm z} > 1$, the cross terms contribute most significantly in high-z auto-correlations. This is probably because both the shear signal and its variance are highest in these correlations (see Fig. 5.4), so any covariance terms containing a contribution from the G field are significant.

5.4.1 Noise-to-Signal Ratio

In Figs. 5.8-5.9 we plot the diagonal of the covariance components over the total shear correlation function to investigate the relative noise contribution of each term at different angular scales for $n_{\rm z} = 1$, 5 and $n_\theta = 15$. The noise-to-signal ratio for the total covariance is lowest between $\sim 10'$ and $50'$. This means that it is important to measure the shear covariance on these scales as the higher signal-to-noise ratio may enable us to better constrain parameters through lensing. For $n_{\rm z} = 1$, $\hat{\mathbf{C}}^{NN,NN}$ contributes proportionally the most to the noise on scales below $\sim 30'$, and $\hat{\mathbf{C}}^{GG,GG}$ the most above these scales, which is consistent with the behaviour of these covariance terms in Fig. 5.3. The other signals also behave as expected based on Fig. 5.3.

For $n_{\rm z} = 5$, the low-z correlations behave somewhat less smoothly. In general, $\hat{\mathbf{C}}^{NN,NN}$ contributes the most to the amplitude of $\hat{\mathbf{C}}^{\rm TOT}$, however the behaviour with angular scale is erratic for the correlations involving the lowest redshift bin. This is due to the high variance on the noise terms affecting the estimate of $\hat{\xi}^{\rm TOT}$ as shown in Fig. 5.2. Although the components of the variance at low-z are well behaved, as shown in Fig. 5.4, the noise-to-signal ratio is not since $\hat{\xi}^{\rm TOT}$ is dominated by shot noise. The high-z correlations are well behaved, although we do not see the crossover at $\sim 30'$ between the NN,NN and GG,GG covariance terms, as the noise dominates on all scales.

The relative contributions to the covariance matrix shown in these plots will be the same for a lensing survey of any area, due to the fact that the covariance scales inversely with area. This means that the results shown here are also applicable to

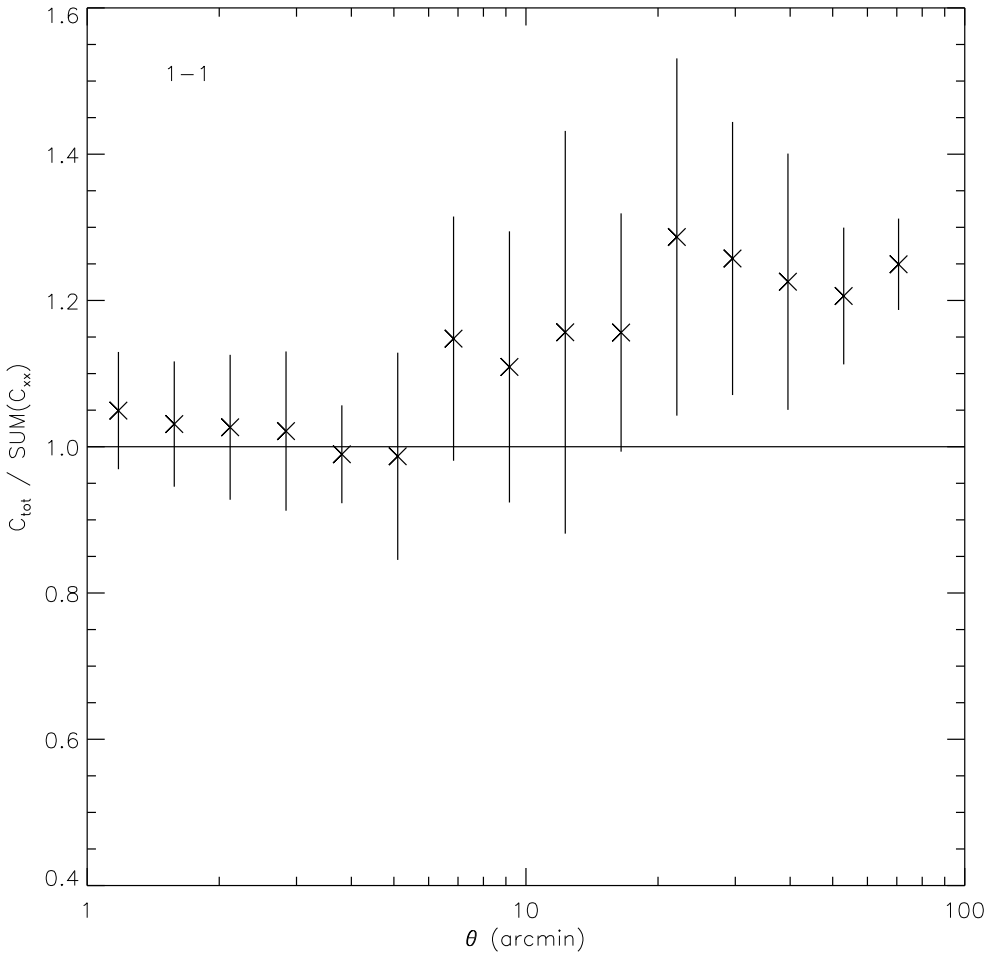

Figure 5.6: The total covariance matrix $\hat{\mathbf{C}}^{\mathrm{TOT}}$ measured from Eqn. 5.16 as a ratio of the sum of all covariance terms in Eqn. 5.17, for $n_{\mathrm{z}} = 1$ and $n_\theta = 15$. The redshift correlation being plotted is shown at the top of the plot.

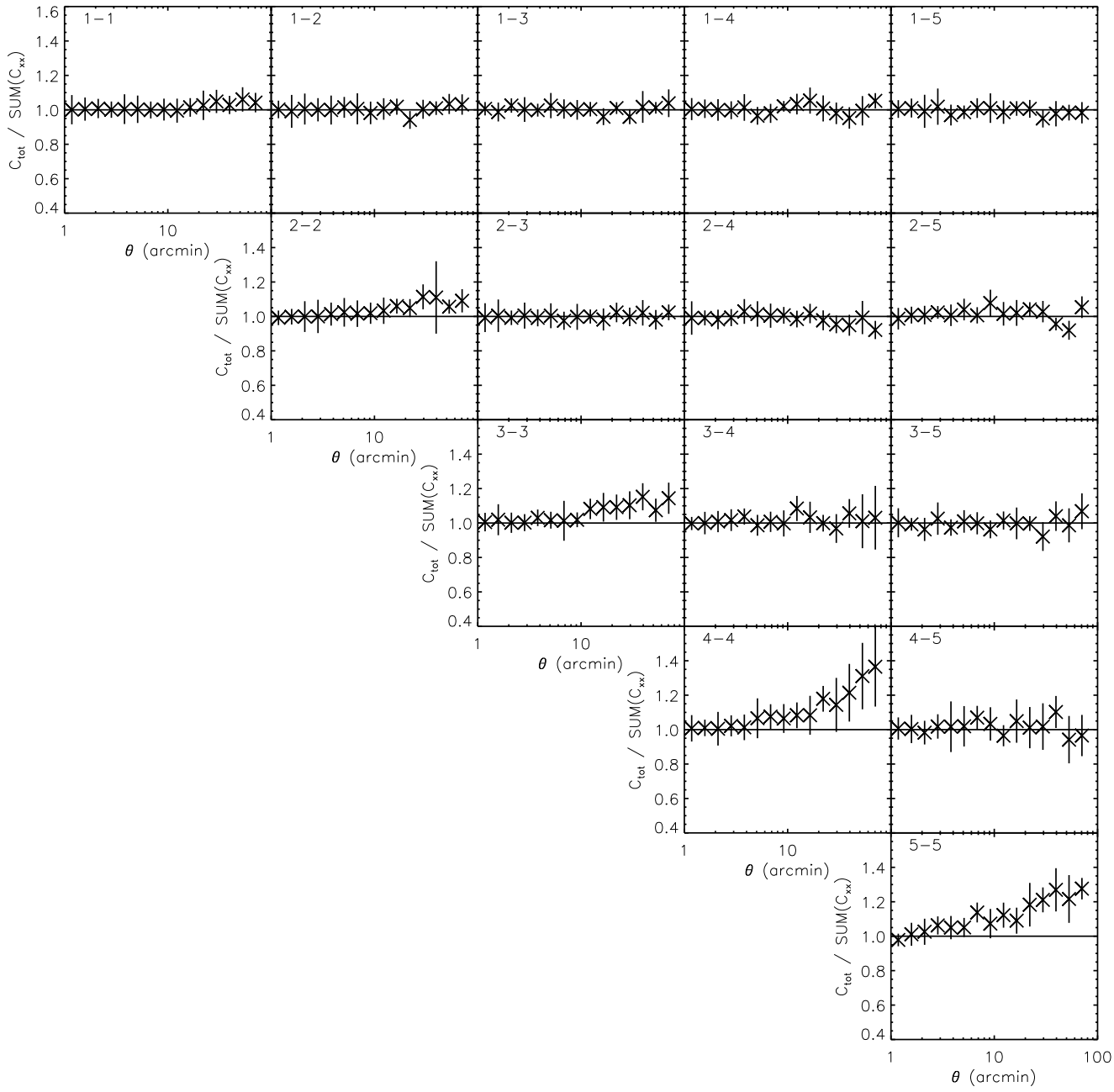

Figure 5.7: As in Fig. 5.6 for $n_z = 5$, $n_\theta = 15$.

future lensing surveys with larger survey areas, if the galaxy density and redshift distributions are similar.

5.4.2 Invertibility of $\hat{\mathbf{C}}_{\mathbf{TOT}}$

To enable us to use the covariance matrices constructed in this chapter in a likelihood analysis, we must ensure that they are non-singular and pseudo-invertible. We use a technique known as Singular Value Decomposition (SVD) to aid us in this section. SVD is a useful method for diagnosing and pseudo-inverting singular or close-to-singular matrices. It is especially well suited to linear least-squares fitting, a fact we exploit in Chapter 6. We may decompose a square matrix $\mathbf{A}$ of dimensions N_u x N_u into

$$\mathbf{A} = \mathbf{U}\mathbf{V}^T\mathbf{W}. \tag{5.18}$$

Here, $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{W}$ are also necessarily N_u x N_u matrices. $\mathbf{W}$ is a diagonal matrix with positive elements (the 'singular values') that are the eigenvalues of $\mathbf{A}$. The eigenvalues of a matrix are those that satisfy the equations $\mathbf{A}\boldsymbol{E} = \lambda\boldsymbol{E}$ ($\boldsymbol{E}$ are the eigenvectors of $\mathbf{A}$). $\mathbf{U}$ and $\mathbf{V}$ are both unitary matrices (that is, the matrix multiplied by its conjugate transpose equals the identity matrix, $\mathbf{U}\mathbf{U}^T = \mathbf{I}$) whose columns form the left- and right-eigenvectors of the eigenvalues $\boldsymbol{\lambda}$ respectively. $\mathbf{W}$ can be thought of as a scaling matrix and $\mathbf{U}$ and $\mathbf{V}^T$ as rotation matrices which together work analogously to the mapping vector $\boldsymbol{\pi}^{\rm map}$ and rotation matrix $\mathbf{R}$ used in principal component analysis in Eqn. 3.12. From Eqn. 5.18, the inverse of $\mathbf{A}$ will be

$$\mathbf{A}^{-1} = \mathbf{V}\left[\frac{1}{\boldsymbol{\lambda}}\right]\mathbf{U}^T \tag{5.19}$$

Eqn. 5.19 will fail if any of the eigenvalues are negative which occurs if $\mathbf{A}$ is close to singular; SVD provides a diagnostic of the invertibility of a matrix [141].

The condition number of the matrix is defined as the ratio of the largest to the smallest eigenvalues, and a matrix with a condition number close to one (the identity matrix has a condition number of 1) is said to be well-conditioned and its inverse can be computed to a high degree of accuracy [204]. Conversely, a

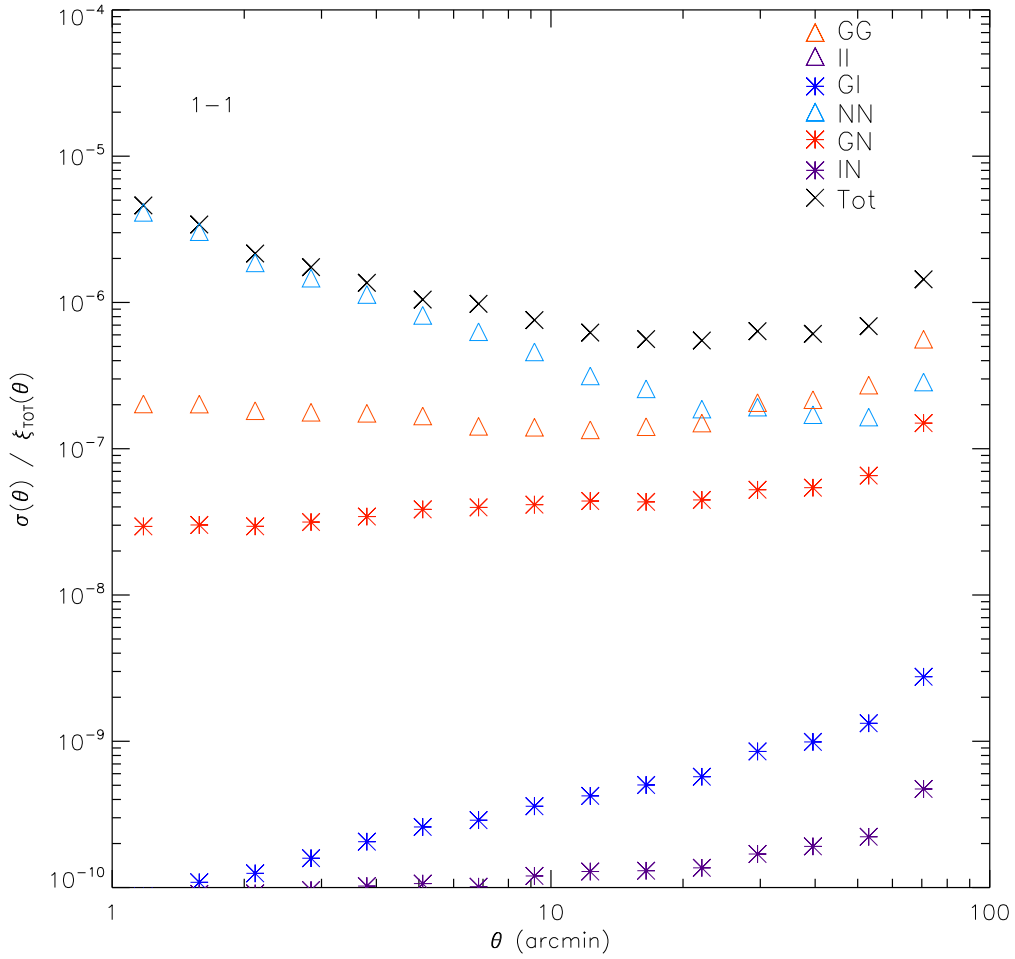

Figure 5.8: The variance-to-signal ratio for different terms contributing to the total covariance $n_z = 1$ and $n_\theta = 15$. $\sigma^{II}/\hat{\xi}^{\rm TOT}$ is of the order $\sim 10^{-13}$ and is not visible. The redshift correlation being plotted is shown at the top of the plot.

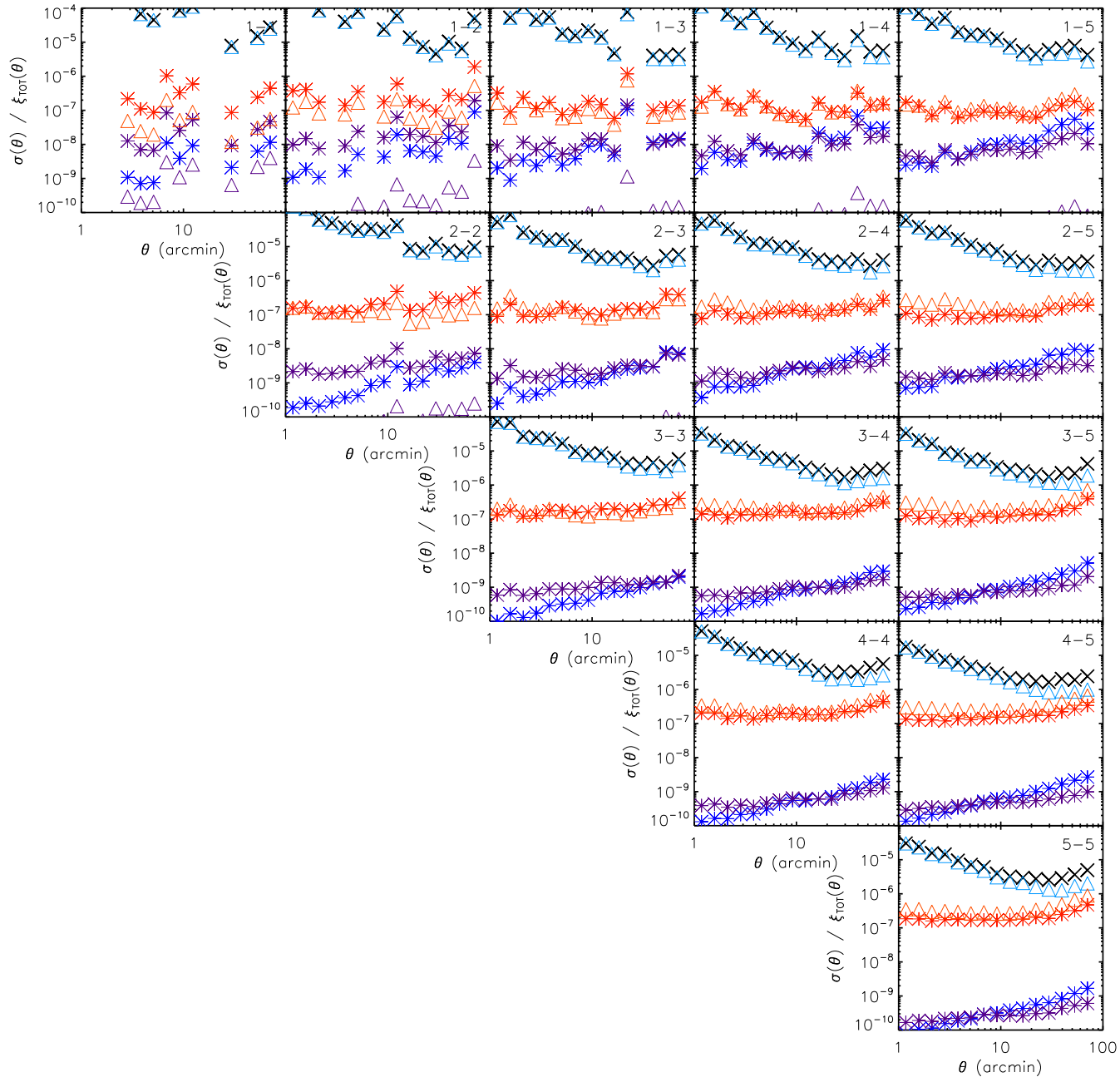

Figure 5.9: As in Fig. 5.8 for $n_z = 5$, $n_\theta = 15$.

singular matrix will have an infinite condition number. If the reciprocal of the condition number is of the order of the level of computing precision available (typically $\sim 10^{-6}$ for single precision calculations, $\sim 10^{-12}$ for double precision), the accuracy required to produce an unbiased estimator of the inverse will not be present and a likelihood analysis may result in biased or imprecise constraints.

We apply the Anderson correction to $\hat{\mathbf{C}}_{\rm TOT}$ as described in §4.1.2. The Anderson correction is applicable to any covariance matrix that has Gaussian noise, so our inclusion of shot noise (which is Gaussian) in the total covariance means that the correction is still valid. We then measure the condition number of $\hat{\mathbf{C}}_{\rm TOT}$. for all $n_{\rm z}$ and n_θ to investigate whether the covariance will be (pseudo) invertible and hence useable in a tomographic analysis. The results are shown in Fig. 5.10. Note that for the higher bin tomography, the condition number cannot be computed for high values of n_θ since this violates the $p/N \leq 1.0$ condition (i.e. the number of data points must not exceed the number of realisations of the data). The highest value of the condition number is $\sim 10^5$ for $n_{\rm z} = 15$, $n_\theta = 6$, which is unsurprising as this bin combination results in the largest possible covariance matrix in our parameter set. In general, since we are able to compute using double precision, the condition number is low enough to ensure that inversion of the matrix is possible to a high enough degree of accuracy to avoid introducing a bias. The condition number increases rapidly with both $n_{\rm z}$ and n_θ, since the size of the covariance matrix grows. The increase in condition number with $n_{\rm z}$ for $n_\theta = 15$ becomes faster than exponential as the limiting case of $p/N = 1.0$ is reached for $n_{\rm z} = 9$.

To obtain a more accurate estimate of the covariance, [203] generate an additional $\sim$ 60 noise-only realisations of the clone catalogues. They then calculate the noise covariance term for each of the 60 extra realisations, and average to get a 'smoothed' covariance estimator. Such a method ensures that any estimate of $\hat{\mathbf{C}}_{\rm TOT}$ will indeed be closer to the true covariance, but it is computationally expensive to generate many additional catalogues. A more accurate estimate of $\hat{\mathbf{C}}_{\rm TOT}$ does not necessarily translate into a less singular and better-behaved matrix, since the true covariance itself may be close to singular. It is, however, important to have enough lines of sight to ensure that the covariance matrix estimate is close to the true value, as the use of just a few lines of sight in the determination of the covariance could result in a noisy matrix that is very

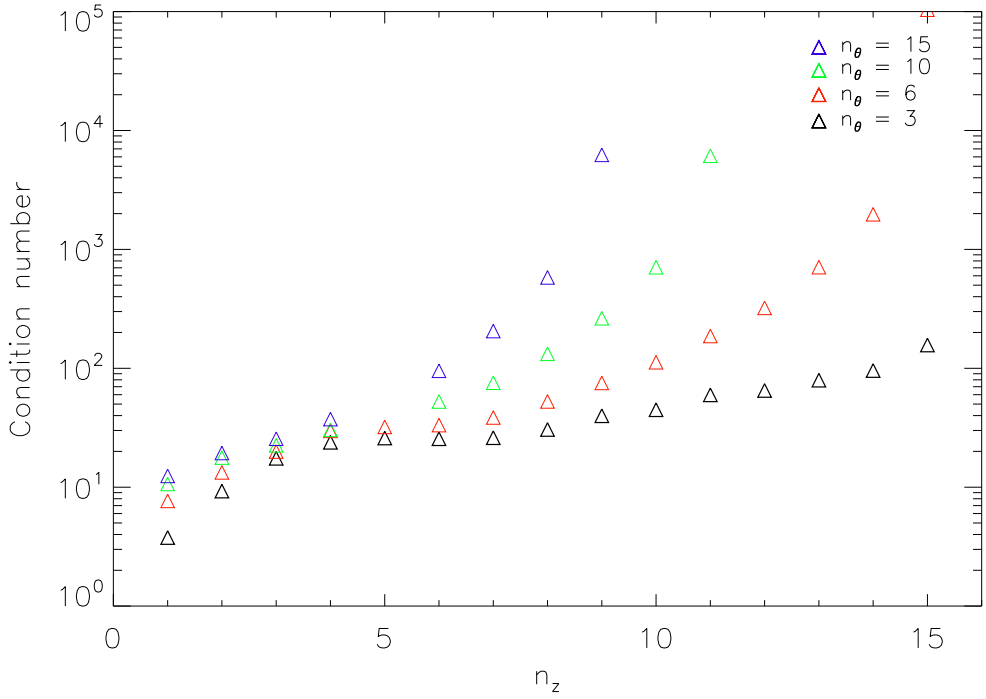

Figure 5.10: The condition number of $\hat{\mathbf{C}}_{\rm TOT}$ for $n_{\rm z} = 1$ to 15 and $n_\theta = 3$, 6, 10 & 15.

different from the true covariance. For this reason, we calculate the condition number of $\hat{\mathbf{C}}_{\rm TOT}$ for different numbers of lines of sight from subsets of the total 736 for $1 \leq n_{\rm z} \leq 15$. The results are plotted in Fig. 5.11. We find that the condition number converges well below the number of lines of sight available to us for all low-bin tomographic covariance matrices. For $n_{\rm z} = 1$, the covariance stabilises by $N \sim 200$, consistent with the results presented in Fig. 4.7 where we investigated the stability of the shear covariance matrix with N. As $n_{\rm z}$ increases, it takes ever more lines of sight to achieve a stable condition number. This is simply a consequence of the fact that the covariance matrix for high $n_{\rm z}$ contains more elements than for low $n_{\rm z}$ (if n_θ = constant), meaning more lines of sight are required to ensure we are in the regime where $p/N < 1.0$ and the covariance matrix inverse is non-singular. For this reason, we also plot the condition number for $n_{\rm z} = 2$, $n_\theta = 5$. The covariance matrix from this bin combination will have the same dimensions as that of $n_{\rm z} = 1$, $n_\theta = 15$. The condition number for this case is even more stable than for the $n_{\rm z} = 1$, $n_\theta = 15$ case, due to the stronger diagonal component from the shot noise for higher $n_{\rm z}$. It is clear from this plot that increasing the number of tomographic bins does not result in a less well-conditioned matrix unless the size of the matrix itself also increases, indicating that the additional (non GG, GG) terms in the total covariance matrix that we have introduced in this chapter do not make it less stable or less invertible. Since the covariance condition number in every tomographic and angular bin combination we wish to run is well below machine precision, we can be confident that $\hat{\mathbf{C}}_{\rm TOT}$ is non-singular and pseudo-invertible, in preparation for the likelihood analysis presented in Chapter 6.

5.5 Summary and Conclusion

In Chapter 4 we estimated the correlation function of cosmic shear only, while incorporating a realistic galaxy redshift distribution and number density as determined from CFHTLenS data. In this chapter, we have extended this work to produce a covariance matrix that incorporates the effects of intrinsic galaxy alignments and shot noise, which are present in any real shear data set. We have then investigated the relative importance of each component of the covariance,

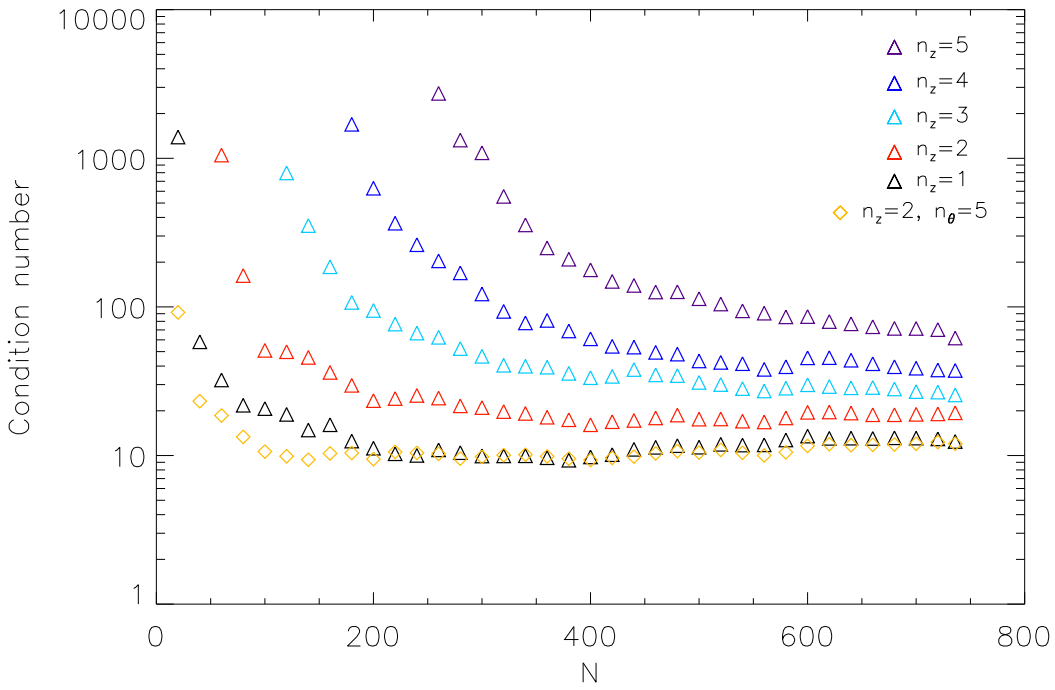

Figure 5.11: The condition number of $\hat{\mathbf{C}}_{\rm TOT}$ for $1 \leq n_{\rm z} \leq 5$ and $n_\theta = 15$ (triangles). Also plotted is the condition number with N for $n_{\rm z} = 2$, $n_\theta = 5$ (diamonds); this covariance matrix has the same dimensions as for $n_{\rm z} = 1$, $n_\theta = 15$.

and the effect of the new covariance matrix on the number of tomographic bins we can use in a likelihood analysis. To achieve this we:

- Generate correlated Gaussian cosmic shear and intrinsic fields incorporating the NLA model for intrinsic alignments and the CFHTLenS redshift distribution and cosmology as described in Chapter 4.
- Measure the shot noise r.m.s. source ellipticity per component from the CFHTLenS as $\sigma_e = 0.28$.
- Construct cosmic shear (G), correlated intrinsic ellipticity (I) and uncorrelated shot noise (N) galaxy catalogues from both the N-body and Gaussian simulations, using $\sigma_e = 0.28$ for the amplitude of the shot noise and CFHTLenS galaxy density.
- Measure $\hat{\xi}$ for every combination of the catalogues by correlating them with themselves and each other, then summing to calculate $\hat{\xi}^{\rm TOT}$.
- Estimate the total covariance, $\hat{\mathbf{C}}_{\rm TOT}$, from $\hat{\xi}^{\rm TOT}$ and the covariance of individual components of the total correlation function ($\hat{\mathbf{C}}_{GG}$ measured from $\hat{\xi}^{GG}$ etc.), finding the shot noise to be the dominant contributor to the covariance matrix and the IA covariance to be negligible in most cases.
- Determine the contribution of cross terms to $\hat{\mathbf{C}}_{\rm TOT}$, finding that they are significant for some redshift correlations on large scales.
- Measure the noise-to-signal ratio and find a minimum between $10' \lesssim \theta \lesssim 50'$ where the noise affects the signal the least.
- Calculate the condition number of the covariance matrix for different tomographic and angular bin numbers and determine that it is stable under pseudo-inversion in all considered cases.

One of the limitations of this analysis is the Gaussian nature of the IA covariance. This could be improved upon if an IA model such as the NLA model is included in N-body simulations, however the small-scale resolution required to measure the IA signal accurately becomes computationally expensive. If IAs were included in N-body simulations, it would then be possible to assign a realistic observed ellipticity ϵ_i to each galaxy which incorporates the shear and IA signal

for that galaxy and a shot noise component. If one is only interested in the total covariance, then simply measuring $\hat{\xi}_+$ from such a catalogue will result in an estimator for $\hat{\mathbf{C}}_{\rm TOT}$. This is computationally less expensive than correlating separate G, I and N fields as we have done in this analysis. Since the IA signal in this analysis comes from a different set of simulations than the shear signal, with different galaxy positions (due to the fact that there are masks present in the clone fields which are absent in the Gaussian fields), we were not able to estimate $\hat{\mathbf{C}}_{\rm TOT}$ in this way. Additionally, we have investigated the relative contributions of different terms in the covariance matrix in this chapter, which is information that would have been lost to us had we taken this approach.

Finally, we have investigated the invertibility of the covariance matrix in preparation for the likelihood analysis in Chapter 6, and found that for all $n_{\rm z}$ and n_θ that are permissible, the condition number inverse is lower than machine precision and the matrix is non-singular (we note however that in the region of $p/N \to 1$, the condition number is much larger than for low values of p/N, which may affect the relative accuracy of the covariance matrix inversion). This means that the conclusion we drew in Chapter 4 - that 10-bin to 15-bin tomography should be achievable with the clone simulations (for at least some values of n_θ) - is still valid.

Chapter 6

Optimal Tomography with the CFHTLenS Covariance

6.1 Introduction

In Chapter 2, we discussed how weak lensing can be used to constrain cosmological parameters. Incorporating a least-squares fit into an importance sampling method can provide estimates and corresponding credibility intervals of the cosmological parameters that are sensitive to the dark matter distribution. We described the construction of such an importance sampling method - the Markov Chain Monte Carlo - in Chapter 3. In this chapter, we employ this algorithm to obtain credibility intervals expected from a tomographic CFHTLenS-like survey for some key cosmological parameters. Our aim is to determine which combination of redshift and angular bins produces the tightest constraints on the parameters of interest, and to explore the limitations of such a tomographic approach.

In the previous two chapters, we have described a method for estimating covariance matrices for cosmic shear analysis from simulations. We now have a realistic estimate of the covariance matrix for a tomographic CFHTLenS-like survey that includes all redshift auto- and cross-correlations. The covariance matrix represents the expected error on our measurements, and any likelihood analysis involving correlated data (such as the shear correlation function) must take it into account. The combined effect of IAs and realistic photo-z errors on the optimal number of tomographic and angular bins for shear analysis has not been investigated before, so in this chapter we incorporate the covariance matrix

estimated in the previous chapters into a likelihood analysis to investigate this. Since we are using the redshift distribution and galaxy density from CFHTLenS, the optimisation we perform in this chapter can be applied to the CFHTLenS data, which we do in Chapter 7. We perform likelihood fits to estimate the posterior distribution of some key cosmological parameters for every viable combination of tomographic and angular bins. We then determine the optimal bin combination that produces the tightest constraints on these parameters. We also investigate the effect of neglecting IAs and photo-z errors on the size and position of cosmological likelihood contours to see if this biases parameter constraints.

6.2 Least-squares fitting

The key calculation in the MCMC algorithm is the determination of the likelihood through its dependence on χ^2, as described in Eqn. 3.7 and Eqn. 4.4. However, as we have discussed in §4.1.2, inverting a covariance matrix, especially one that is close-to-singular, can result in a biased and unreliable estimator for the inverse. This can lead to one incorrectly estimating the size of the credibility intervals one would expect from a given covariance. A solution to this is to pseudo-invert the covariance matrix through singular value decomposition. We introduced SVD in §5.4.2 as a means to diagnose the invertibility of the covariance matrix through its condition number. SVD can also be used to calculate an estimate of χ^2; we detail in this section how this is achieved.

We follow the method of [205] which exploits the properties of SVD described in [141]. First, the covariance matrix is normalised such that the diagonal is rescaled to unity and the off-diagonals fall between ± 1. We define the variance $\sigma_i = \mathbf{C}_{i,j}\delta_{i,j}$ such that the normalised covariance is

$$\mathbf{C}_N = \frac{\mathbf{C}}{\sigma^T \sigma}. \tag{6.1}$$

$\mathbf{C}_N$ is also known as the correlation matrix. SVD is a special case of eigenvalue decomposition that produces a system of linear equations that satisfy

$$\mathbf{C}_N \mathbf{E}_i = \lambda_i \mathbf{E}_i \quad i = 1...N_u\,, \tag{6.2}$$

for a covariance matrix of dimensions N_u x N_u, where as in §5.4.2, λ_i and $\mathbf{E}_i$ are

the eigenvalues and eigenvectors of the (normalised) covariance matrix, and some eigenvalues are set to zero. The eigenvectors form a complete orthonormal basis, and can be used to recover $\mathbf{C}$.

With the eigenvectors and eigenvalues, one can redefine χ^2 in Eqn. 4.4 to be

$$\begin{aligned} \chi^2 &= \sum_{i=1}^{N_u} \frac{(\boldsymbol{y}_i - \boldsymbol{y}(\pi))\,(\boldsymbol{y}_i - \boldsymbol{y}(\pi))^T}{\lambda_i} \\ \boldsymbol{y}_i &= \mathbf{E}^T \sigma^{-1} \bar{\boldsymbol{\xi}}_+. \end{aligned} \tag{6.3}$$

Summing every linear equation from 1 to N_u gives exactly the same result as the standard χ^2 expression, while circumventing the need to directly invert the covariance matrix. It is possible to sum over fewer than N_u modes to calculate χ^2. This may be desirable in the case where the condition number is very high (close to machine precision). This is because in such a scenario the last few modes (corresponding to the smallest eigenvalues) will be corrupted by round-off error, and the inclusion of these in the linear combinations of equations can bias the solution we derive for the whole set. Therefore it can be beneficial to construct χ^2 from only those modes for which the condition number lies well above machine precision. In §5.4.2 we found that the condition number is much smaller than the machine precision available to us for every tomographic and angular bin combination we can explore, so we do not impose a mode cutoff in this analysis.

6.3 Optimisation

Our optimisation parameter space consists of every statistically viable combination of tomographic and angular bins in the region $1 \leq n_z \leq 15$ and $3 \leq n_\theta \leq 15$ as determined from Chapters 4 & 5. We incorporate the Bayesian estimates of the redshift distribution (z(PDF) in Chapter 4) into the theoretical prediction of the total shear correlation function. We optimise on $\hat{\xi}_+^{\mathrm{TOT}}$, which incorporates the expected signal from the II and GI alignment correlation functions. The lensing signal is sensitive to multiple cosmological parameters, however we must choose the parameters we wish to include in our likelihood analysis carefully, as there is limited constraining power available to us with current lensing surveys such as CFHTLenS. We choose to vary three parameters of interest: Ω_m, σ_8 and the

NLA model amplitude parameter A. As described in §3.6, preliminary MCMC chains showed that there is very little constraining power present for h_0, w_0 or w_a, so we choose not to constrain these parameters as this would vastly increase the dimensions of the parameter space and hence the time to convergence, with very little information return. h_0 is well constrained from HST supernovae data [12] to be 0.738 ± 0.024 (to 1σ), so we constrain it with this Gaussian prior in our likelihood analysis. This is the approach used by [149] and [206] in similar cosmological likelihood analyses. Practically, this means that we allow the MCMC to step within the region that encompasses $h_0 = 0.738 \pm 3\sigma = 0.738 \pm 0.072$, and then multiply the likelihood at this point by its probability as determined from a Gaussian distribution. This means we are truncating the Gaussian at the 3σ level rather than letting the parameter space extend to $\pm\infty$ as it formally should, however by definition only $\sim 1\%$ of the MCMC points will attempt to step outside this region so it is highly unlikely to affect our parameter estimates.

We assume flat ΛCDM cosmology to match the clone simulation, such that $\Omega_m + \Omega_\Lambda = 1$ and the dark energy equation of state, parametrised as $w(z) = w_0 + w_a(1-a)$, is set to $w_0 = -1.0$ and $w_a = 0.0$. All cosmological parameters that we do not constrain are set to the fiducial values of the clone input as described in Table 4.1. We assign tophat priors to the parameters we wish to constrain. Ω_m and σ_8 are allowed to vary between $0.1 \leq \Omega_m \leq 1.0$ and $0.1 \leq \sigma_8 \leq 1.5$ (the ranges for which `NICAEA` has been tested and can reliably predict the correlation function). We set a hard prior such that $A > 0$ as we wish to optimise on the fiducial, positive value of A. We allow A to vary within the range $0 < A < 15$; this was determined from trial MCMC runs with $n_z = 1$ and 10, where we find there is very little posterior probability in the region $A > 15$.

6.3.1 Optimisation metrics

$\xi_+(\theta)$ is relatively featureless and produces a highly degenerate, curved contour for Ω_m and σ_8. This contour is, however, quite narrow, and almost orthogonal to constraints from analysis of cosmic microwave background data. This makes it possible to combine confidence regions from both data sources and lift this degeneracy [81, 207]. For this reason, we measure the area of the Ω_m – σ_8 68.3% contour with tomography as an optimisation metric (we denote this metric 1). [152] find that two-bin tomography partially breaks the degeneracy between the

two parameters, indicating that tomography has the potential to put tight limits on the size of this contour. We also measure the width of the Ω_m (metric 2) and σ_8 (metric 3) 68.3% likelihood contours around the fiducial values for different bin combinations, to see if the optimal binning for these parameters differs.

We also investigate the effect of tomography on constraints achievable on the NLA model normalisation constant A. We measure the 68.3% likelihood width of A (marginalised over every cosmological parameter) for every tomographic and angular bin combination (metric 4). It is unknown whether the optimal combination that minimises the error on A will be the same as that which minimises the optimisation metrics for Ω_m and σ_8; our analysis will determine if this is the case.

6.3.2 The data vector

Since the input correlation functions are poorly recovered from the simulations above $\sim 30'$ for the clone and below $\sim 5'$ for the Gaussian simulations (see Figs. 4.3-4.4), using them in the likelihood calculation may result in a bias in cosmological parameter estimates. Much of the constraining power of the shear correlation function exists on these scales, so simply excluding these angular scales will reduce the precision of our parameter estimates considerably. This would not reflect the precision we can expect to achieve with real data, which is not limited by pixel resolution or edge effects as simulations are. For this reason, we set the mean of the data vector equal to the input correlation function calculated from `NICAEA`, $\xi_{\rm th}^{\rm TOT}(\theta)$, on all angular scales (where the subscript 'th' denotes the theoretical correlation function prediction for a given cosmology). This means that $\hat{\xi}^{\rm TOT} = \xi_{\rm th}^{\rm TOT}$, and for the fiducial cosmology we will measure $\chi^2 = 0$. This represents an ideal measurement of the correlation function, but since we are optimising the number of tomographic and angular bins in this chapter, this enables us to avoid introducing biases in the likelihood contours and hence drawing potentially erroneous conclusions about the optimum bin combination(s). This approach also accounts for the fact that the WMAP prior we introduce on h_0 is slightly different from the value used in the N-body simulations, ensuring that this does not introduce a bias into the cosmological constraints.

6.3.3 Running MCMC chains

We decide the minimum chain length using the *Gelman-Rubin* statistic R mentioned in §3.4.1. R is formally defined as follows. For M chains of length $N_{\rm c}$, where chain elements are denoted y_i^j and $i = 1...N_{\rm c}$, $j = 1...M$, the mean of a chain is

$$\bar{y}^j = \frac{1}{N_{\rm c}} \sum_{i=1}^{N_{\rm c}} y_i^j. \tag{6.4}$$

The mean from all chains is

$$\bar{y} = \frac{1}{MN_{\rm c}} \sum_{i,j=1}^{MN_{\rm c}} y_i^j. \tag{6.5}$$

Therefore the variance between chains is defined as

$$V = \frac{1}{M} \sum_{j=1}^{M} (\bar{y}^j - \bar{y})^2, \tag{6.6}$$

and the variance within a chain as

$$W = \frac{1}{M(N_{\rm c} - 1)} \sum_{i,j=1}^{MN_{\rm c}} (\bar{y}_i^j - \bar{y}^j)^2. \tag{6.7}$$

The *Gelman-Rubin* statistic is then defined as [142, 143]

$$R = \frac{W\left(\frac{N_{\rm c}-1}{N_{\rm c}}\right) + V\left(1 + \frac{1}{M}\right)}{W}. \tag{6.8}$$

The right hand side of Eqn. 6.8 contains two estimates of the variance. The numerator in R is an estimator of the variance in the entire distribution that will be overestimated if the chains are not convergent. Conversely, the denominator will be an underestimate if convergence is not reached. From trial chains, we find that we need between ~ 2 x 10^5 and 5 x 10^5 points in total for each least-squares fit to achieve an acceptable value of $R \leq 1.03$ (see §3.4.1). The longest chains are generally needed for bin combinations that result in the largest credibility intervals; since the points are spread more thinly over the parameter space it takes more points to ensure we are sampling representatively from the posterior.

An important factor to consider when using an iterative sampling method such

as MCMC is the computing time needed. The largest computing overhead comes from the calculation of theoretical shear correlation functions in `NICAEA`, which must be calculated for each point in the MCMC chain. Increasing the number of tomographic bins in the analysis results in a more complicated calculation and slows down at a rate that is slightly worse than linear, i.e. for 10-bin tomography, the calculation is ~ 12 times slower than for 1-bin tomography (~ 6 seconds on a standard CPU vs. ~ 0.5 s)[1]. Since many iterations of the MCMC are needed to achieve convergence, we utilise multiple CPUs and run many chains in parallel. We then recombine the resulting chains to produce long, convergent chains. We can adopt this approach since we select a random starting point for each chain that is close to the input cosmology, removing the need for burn-in as described in §3.4.2.

We follow the procedure to determine an optimal proposal distribution as described in §3.6.1. Once Gaussian proposal distributions were generated for each value of n_z, we ran some short trial chains to observe their behaviour. In every case, the new proposal distributions resulted in a much higher acceptance ratio than our prelimiary tophat distribution ($\gtrsim 50\%$ compared to $\sim 25\%$). We scaled the variances on each fitted parameter to ensure the acceptance rate for each n_z is approximately 50%, as this generally results in the shortest time to convergence [136]. This also ensures that all our chains will need to be about the same length to achieve convergence, as long as the shape of the posterior does not change dramatically over the optimisation parameter space.

6.3.4 Credibility intervals

The output from the MCMC algorithm is a (combined) chain of length N_c, consisting of χ^2 values and their coordinates in cosmological parameter space. To produce credibility intervals for all possible parameter combinations, we exploit the characteristic of importance sampling mentioned in §3.4: the density of points in a region of parameter space is directly proportional to the likelihood of that region (assuming convergence is reached). Therefore we estimate the density of points by gridding the multidimensional parameter space into fine bins. Next,

[1]The same slow-down does not apply for increased angular scales as `NICAEA` calculates ξ_{th} on a wide range of scales simultaneously and interpolates to produce the required number of angular scales.

we count the number of MCMC points in each grid space. This produces a 3-dimensional density array (corresponding to Ω_m, σ_8 and A), which allows us to estimate joint 2D likelihood contours for all combinations of parameters by marginalising over the third parameter and drawing contours around the first 68.3%, 95.4% and 99.0% of the grid spaces with the highest density of points. Marginalisation means we simply ignore the value of the third parameter, essentially collapsing this dimension of the gridded likelihood array down. This is roughly equivalent to, but better than, using standard $\Delta\chi^2$ values to determine 1, 2 and 3 σ credibility intervals as this is only accurate when applied to Gaussian probability distributions [143]. Due to the strong degeneracy between parameters, this assumption is not valid.

6.4 Optimal Binning

In Fig. 6.1 we show the predicted constrains on the three fitted parameters achievable from a 2D analysis with $n_z = 1$ and $n_\theta = 8$ as a result of the least-squares MCMC fit. We find that the Ω_m-σ_8 contour is curved and degenerate as expected, and the size of the 68.3% contour is smaller than the constraints from the early analysis of CFHTLS data by [124] and [154], which we expect since the total survey area of CFHTLenS is larger than the areas used in their analyses (see §7.2). We predict 1D credibility intervals (marginalising over every other parameter)[2] of $\Omega_m = 0.241^{+0.124}_{-0.080}$ and $\sigma_8 = 0.847^{+0.143}_{-0.188}$ (all values quoted are 68.3% CL). Our predicted constraints are similar to those achieved by [208] and [209] who use the full CFHTLenS catalogues and which we will review in §7.2. Additionally, we find that even a 2D analysis will have some constraining power on the NLA amplitude parameter, predicting $A = 0.746^{+1.426}_{-0.663}$ and a 68.3% upper bound of 2.17 when marginalising over Ω_m and σ_8.

We plot the four optimisation metrics as a function of the number of tomographic bins in Fig. 6.2 for a representative range of angular bin numbers ($n_\theta = 4$, 8 & 12). There is a strong dependence on n_z for all metrics. This shows

[2]We note that the means recovered are not exactly equal to the simulation inputs due to the random stepping nature of the MCMC and the gridding procedure used to determine contour intervals, but they are well within the tolerance of the error bars in every case.

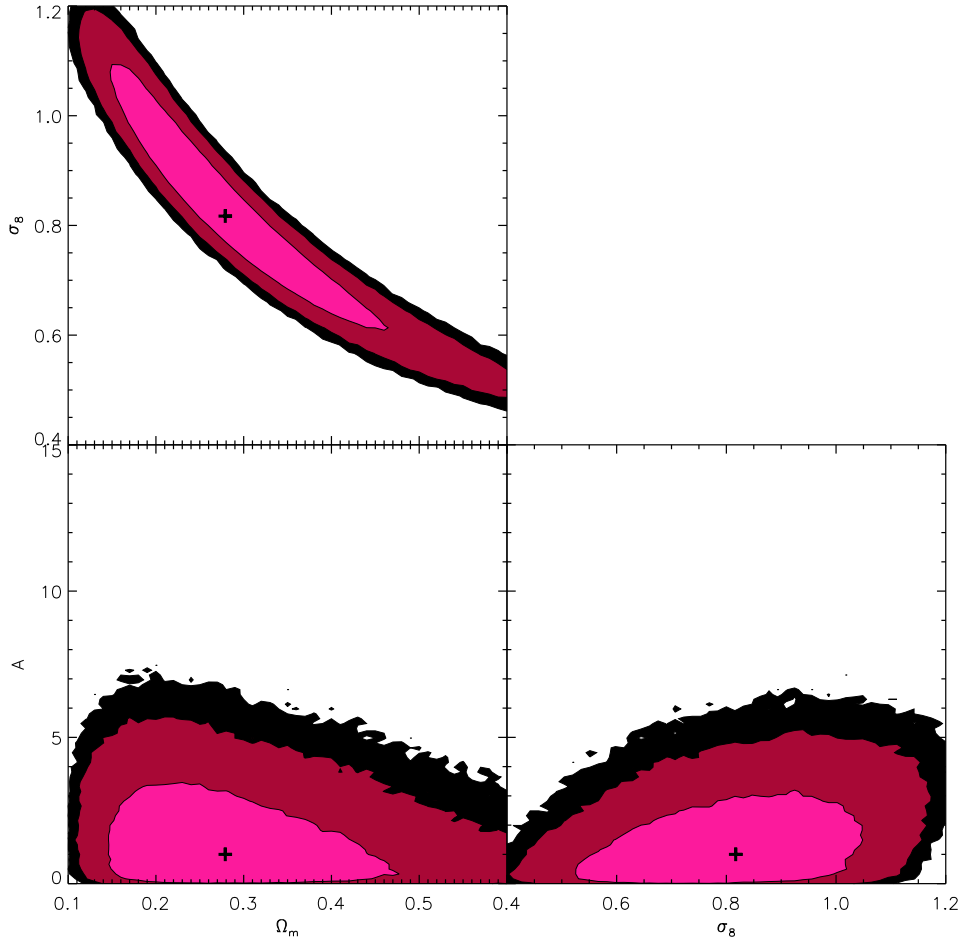

Figure 6.1: 68.3%, 95.4% and 99.0% contour intervals on Ω_m, σ_8 and A for $n_z = 1$, $n_\theta = 8$ from a chain of length ~ 5 x 10^5. The input cosmology is indicated with a cross.

that the Ω_m-σ_8 contour and the NLA model amplitude are both highly sensitive to tomography. For each n_θ there is an optimum number of tomographic bins that minimises the metrics. In every case, the size of the optimisation metrics begins to increase rapidly at high n_z as the data vector and covariance increase in size. There is also a shift along the x-axis in the optimisation curve with n_θ; when n_θ is high the data vector is large and hence the optimum tomographic bin number is lower than for smaller n_θ. This means that there is a partial degeneracy between the optimal n_z and n_θ - increasing the number of tomographic bins will decrease the number of angular bins needed to reach the optimal metric values, and vice versa.

In Fig. 6.3 we plot the optimisation metrics as a function of the number of angular bins for $n_z = 3$, 6 & 9. We find in general that the metrics are much less sensitive to the number of angular bins than to the number of tomographic bins. This is to be expected since the addition of tomographic bins rapidly increases the size of the data vector p and recovers a larger amount of information, whereas adding additional angular bins increases p only linearly. The slight exception to this trend is the behaviour of $n_z = 9$, which rises steeply at high n_θ because the covariance matrix becomes prohibitively large.

Because of the limited statistical precision of the mock survey, we also investigate the optimal size of p for each of the metrics, as shown in Fig. 6.4. We find that there is a general trend for the metrics to reach a minimum at $100 \lesssim p \lesssim 200$ for all bin combinations. The points from different values of n_θ follow the same trend, although the points from $n_\theta = 4$ sit slightly above those from higher angular bins numbers. This is due to the fact that for such a low number of angular bins, there is less overall constraining power present than for higher angular bin numbers. These results indicate that middling values for both n_z and n_θ minimise the overall amount of correlation in the data vector, so as long as n_θ is neither exceptionally small nor large (i.e. the combination of n_z and n_θ is close the the optimum), the ideal binning is more strongly dependent on the total number of data points used and hence the dimensions of the covariance matrix than on the exact combination of n_z and n_θ used.

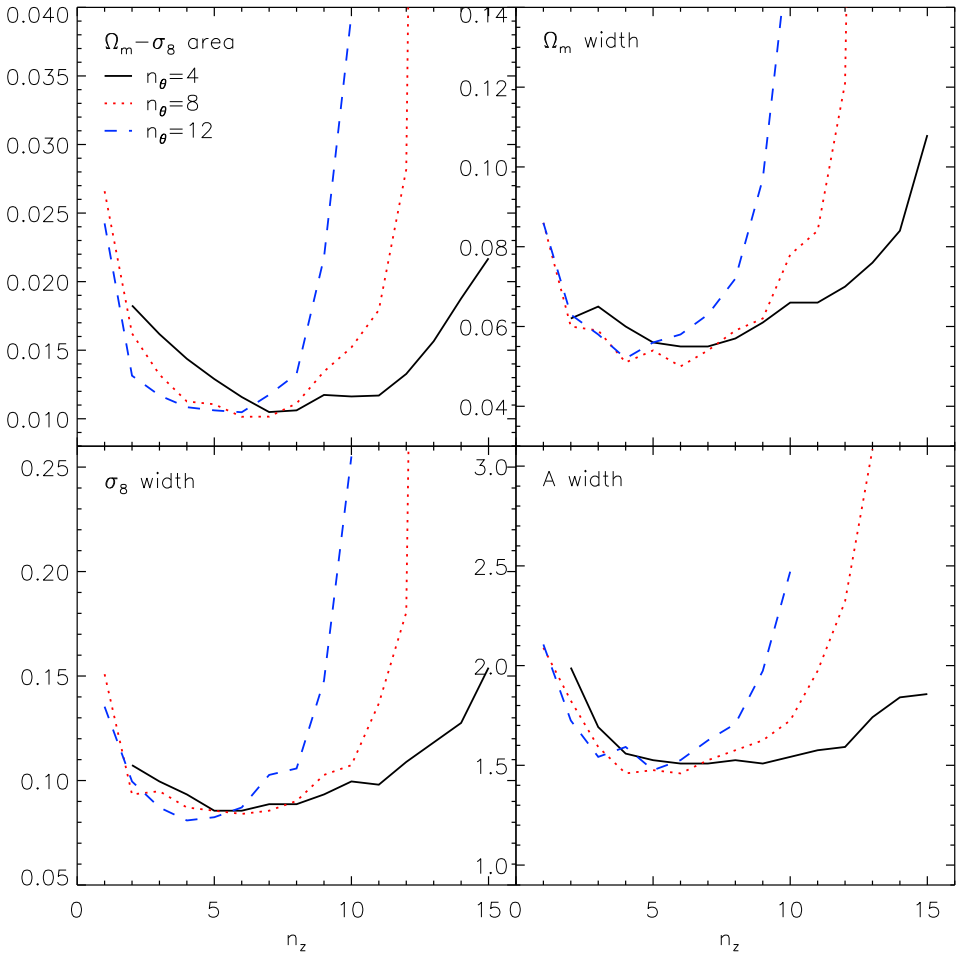

Figure 6.2: The value of the four optimisation metrics as a function of n_z, for $n_\theta = 4$, 8 & 12. *Top left*: The area of the Ω_m-σ_8 68.3% likelihood contour. *Top right*: The width of the Ω_m 68.3% contour at the fiducial value of σ_8. *Bottom left*: The width of the σ_8 68.3% contour at the fiducial value of Ω_m. *Bottom right*: The marginalised width of the A 68.3% contour. All units are dimensionless.

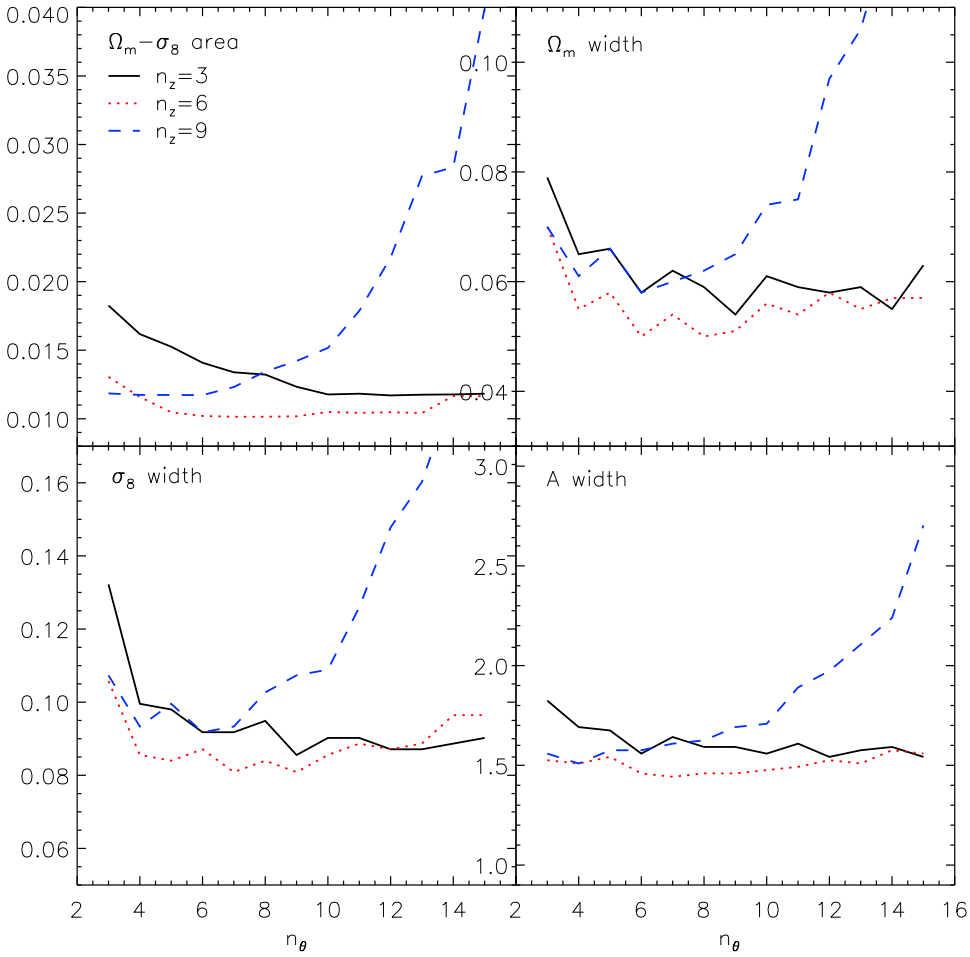

Figure 6.3: The value of the four optimisation metrics as a function of n_θ, for $n_z = 3$, 6 & 9. *Top left*: The area of the Ω_m-σ_8 68.3% likelihood contour. *Top right*: The width of the Ω_m 68.3% contour at the fiducial value of σ_8. *Bottom left*: The width of the σ_8 68.3% contour at the fiducial value of Ω_m. *Bottom right*: The marginalised width of the A 68.3% contour. All units are dimensionless.

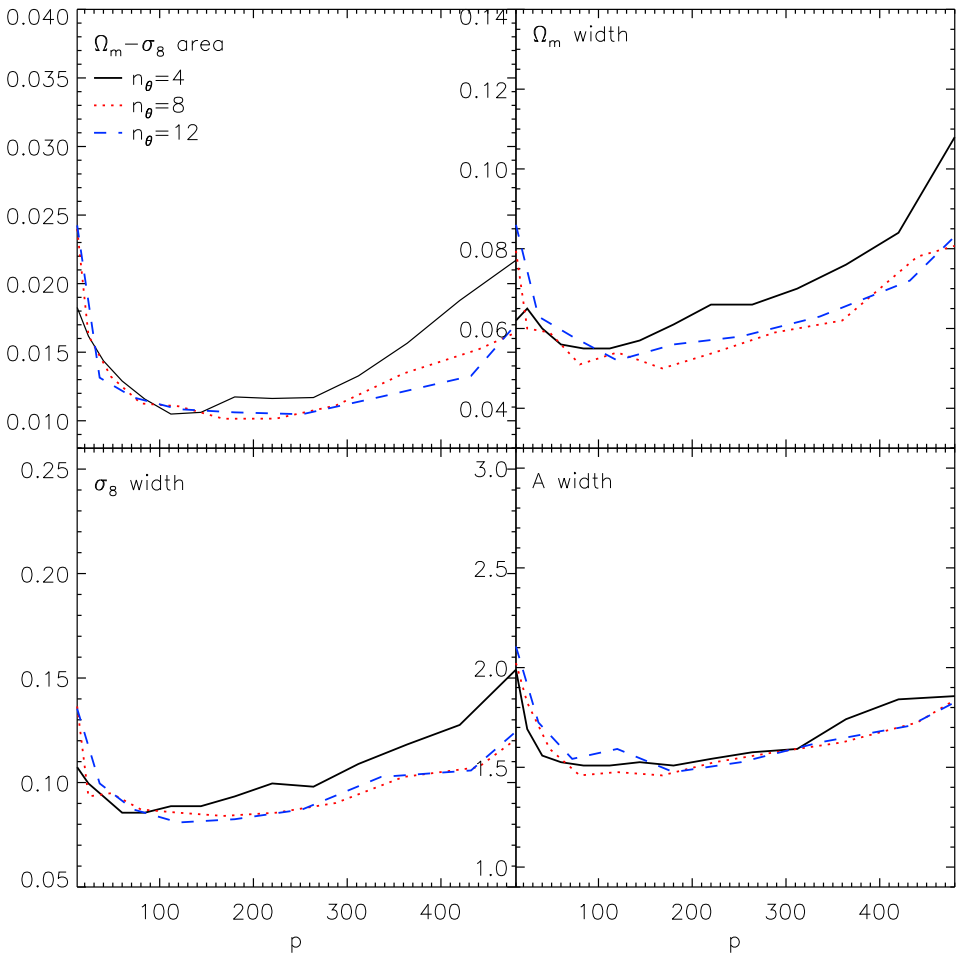

Figure 6.4: The value of the four optimisation metrics as a function of p, for n_z = 4, 8 & 12. *Top left*: The area of the Ω_m-σ_8 68.3% likelihood contour. *Top right*: The width of the Ω_m 68.3% contour at the fiducial value of σ_8. *Bottom left*: The width of the σ_8 68.3% contour at the fiducial value of Ω_m. *Bottom right*: The marginalised width of the A 68.3% contour. All units are dimensionless.

Overall, the optimum bin combination is similar for the different optimisation metrics. We summarise our findings in Table 6.1. Metric 1 reaches a minimum (i.e. the gradient of the curve is approximately zero) for $n_z = 7$, $n_\theta = 7$. Metric 2 is at a minimum at $n_z = 6$, $n_\theta = 8$, metric 3 at $n_z = 6$, $n_\theta = 9$ and metric 4 at $n_z = 6$, $n_\theta = 7$. There is a factor of approximately 3 decrease in the size of the Ω_m-σ_8 contour for the optimal case compared to the 2D constraint, and an approximate 40% reduction in the width of the contour about the fiducial value. This is a somewhat smaller improvement than that found by [126] and [152], who find up to a factor of 10 decrease in parameter uncertainties with 3- and 4-bin tomography respectively. However, their analyses do not incorporate realistic photometric redshift bins, IAs or the Anderson correction into their results, which most likely account for the difference in improvement. In agreement with [152], we find that most of the additional constraining power from tomography is achieved from splitting the data into two or three redshift bins, with only incremental gains after this point. The constraint on A is improved by only around 30% by tomography, indicating that it is slightly less sensitive to tomography than the fitted cosmological parameters.

Table 6.1: Optimisation metrics

Metric	Quantity	Optimal value	n_z	n_θ	p	Fraction of 2D value
1	Ω_m-σ_8 area	1.01 x 10^{-2}	7	7	196	0.338
2	Ω_m width	5.00 x 10^{-2}	6	8	168	0.581
3	σ_8 width	8.09 x 10^{-2}	6	9	189	0.565
4	A width	1.443	6	7	147	0.707

Since $n_z = 6$, $n_\theta = 8$ is the optimal bin combination for one of the metrics and sits in between two of the other optimal bin combinations, we show the credibility intervals from this bin combination for the fitted parameters in Fig. 6.5. We find marginalised predicted 1D credibility intervals for this binning of $\Omega_m = 0.265^{+0.082}_{-0.048}$, $\sigma_8 = 0.809^{+0.090}_{-0.096}$ and $A = 1.277^{+0.663}_{-0.796}$. Our results are considerably more precise than the values achieved from previous tomographic analyses such as the 6-bin analysis of COSMOS data in [156] who found $\Omega_m = 0.32^{+0.34}_{-0.11}$ assuming flat ΛCDM cosmology. The improvement is due to the substantial ($\sim 10^2$)

increase in survey area of the CFHTLenS compared to COSMOS.

We note that with tomography and CFHTLenS uncertainties, there is moderate constraining power on A such that very high values (e.g. $A \sim 10$) could be ruled out. This is promising as it shows that CFHTLenS may have enough constraining power on A to detect the presence of a significant IA signal and may yield information on the galaxy population that is thought to contribute to IAs (see §5.2).

6.5 Effect of covariance matrix errors

We now turn to the issue of the upward trend at high bin values in Figs. 6.2-6.4. This result is contrary to what one might expect from increasing the size of the data vector; given a sufficiently accurate covariance, the optimisation metrics would be expected to plateau at high n_z and n_θ since the addition of more data points always mean there is more information even in the presence of degeneracies. This is the result found by BK07 for power spectrum tomography. Many of our results begin to plateau before rising again at high bin numbers when the error in the covariance matrix becomes prohibitively large. Since the measurement of tomographic correlation functions requires the estimation of a noisy covariance matrix from a finite number of lines of sight, these plots show the limit at which the clone covariance matrix estimate breaks down. This manifests as a sharp increase in parameter uncertainties.

We verify that this is the cause of the upturn in Fig. 6.6, where we plot the values of the optimisation metrics from a likelihood analysis that utilises only half the available clone lines of sight to calculate the covariance for $n_\theta = 8$. When using only 368 lines of sight, the optimisation metrics reach minimum values at lower n_z than when all 736 lines of sight are used. This is to be expected, since the estimate of the covariance is less reliable and the limiting case of $p/N = 1$ is reached at a lower bin number. We see some plateauing behaviour for the case of $N = 368$, although the plateau extends to higher n_z when more lines of sight are included in the covariance. Since we see partial plateaus for all the optimisation metrics for $N = 736$, this indicates that the addition of more lines of sight will probably not minimise the metrics any further, and will have the

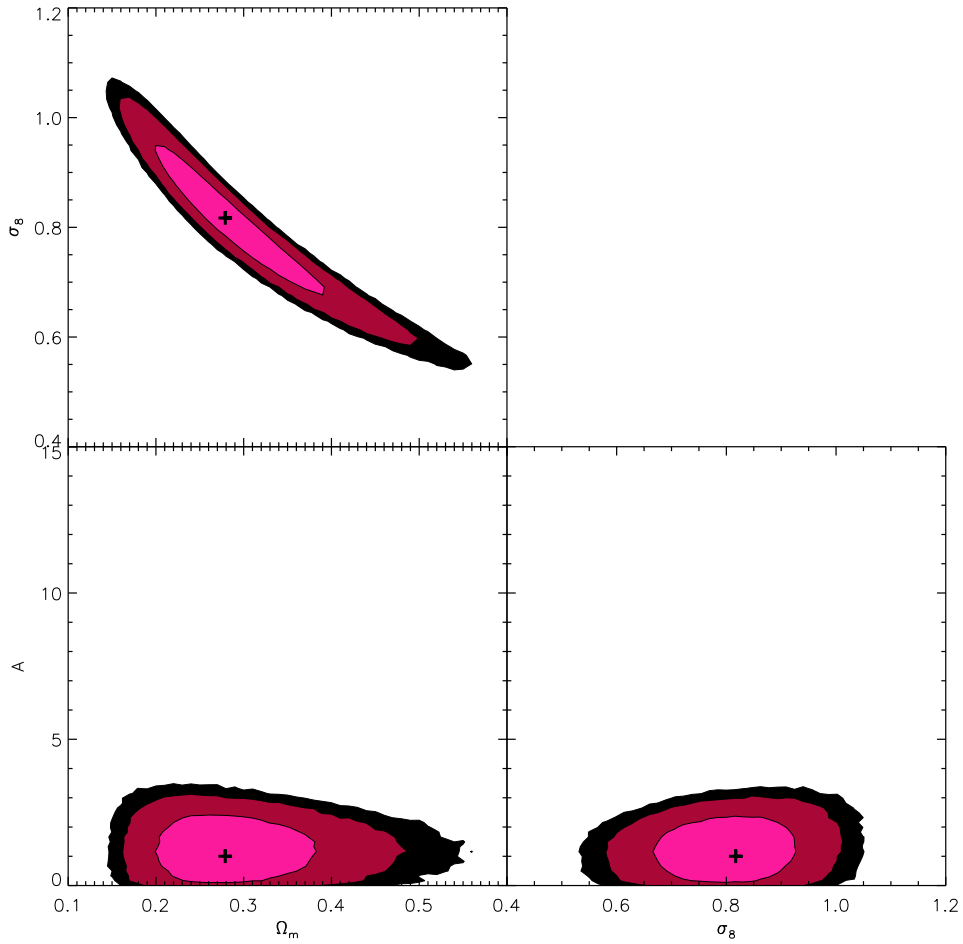

Figure 6.5: 68.3%, 95.4% and 99.0% contour intervals on Ω_m, σ_8 and A for the optimal case of $n_z = 6$, $n_\theta = 8$ from a chain of length $\sim$ 5 x 10^5. The input cosmology is indicated with a cross.

effect of extending the plateaus to higher bin numbers. Overall, the metrics for the two cases reach similar minimum values, however the optimal value for metric 1 is higher for $N = 368$ than for $N = 736$ (1.23 x 10^{-2} for $N = 368$, $n_z = 4$ as opposed to 1.01 x 10^{-2} for $N = 736$, $n_z = 6$), meaning that further analysis with a greater number of simulation realisations is needed to test whether the plateau has truly been reached or if there may be some further gain in the values of the metrics with higher N.

6.6 Effect of neglecting IAs and photo-z errors

Now we have determined the optimal binning for a realistic CFHTLenS-like data set we characterise the bias produced on cosmological parameter estimates by neglecting IAs and photo-z errors in a tomographic analysis. To investigate the bias from neglecting IAs, we perform a likelihood fit on the optimal bin combination utilising $\hat{\xi}^{\rm TOT}$ as the data vector but neglecting the IA contribution to the shear correlation function in the theoretical value we fit it to (i.e. we fit to just $\xi_{\rm th}^{GG}$). Similarly, to ignore photo-z errors, we perform another likelihood analysis, this time fitting $\hat{\xi}^{\rm TOT}$ calculated using z(PDF) to the `NICAEA` prediction of $\xi_{\rm th}^{\rm TOT}$ for the raw photo-z distribution, z_p (see §4.2.3).

We plot the resulting constraints on Ω_m and σ_8 in Fig. 6.7. On the left of the plot, one can see the bias towards lower σ_8 introduced in the position of the credibility intervals for the case where IAs are neglected. The bias is small and consistent with the fiducial cosmology to within 68.3%. The marginalised 68.3% confidence limits are now $\Omega_m = 0.272^{+0.066}_{-0.058}$, $\sigma_8 = 0.804^{+0.084}_{-0.096}$ (since our likelihood fit has no constraining power on A in this case, we ignore it). The credibility intervals have not changed in shape or size from those in Fig. 6.5, as is to be expected since the covariance matrix is identical. This plot shows that, if the NLA model is correct and $A \sim 1$, the effect of IAs in a CFHTLenS-like survey will be small. However, these are significant assumptions that may not be true. Additionally, when fitting to the NLA model as we have done in Fig. 6.5, we have the luxury of perfect knowledge of the IA model in the simulations; we will not have this advantage when fitting to CFHTLenS data and hence additional, unforeseen biases may occur in the real data due to our incomplete knowledge of

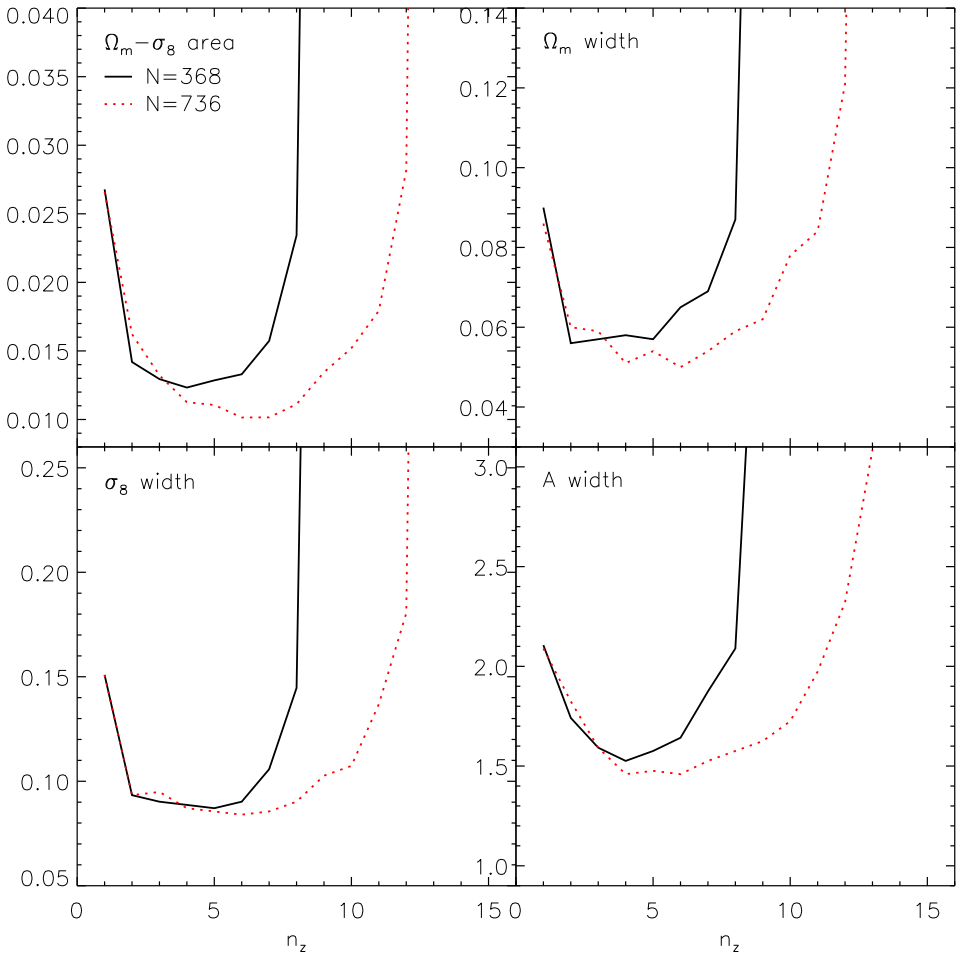

Figure 6.6: The value of the four optimisation metrics as a function of n_z, for $n_\theta = 8$ for covariances constructed from 368 (solid black line) and 736 (dotted red line) lines of sight. *Top left*: The area of the Ω_m-σ_8 68.3% likelihood contour. *Top right*: The width of the Ω_m 68.3% contour at the fiducial value of σ_8. *Bottom left*: The width of the σ_8 68.3% contour at the fiducial value of Ω_m. *Bottom right*: The marginalised width of the A 68.3% contour. All units are dimensionless.

IAs. Our result shows that for the CFHTLenS, an imperfect treatment of IAs is unlikely to bias cosmological constraints catastrophically.

The right hand side of Fig. 6.7 shows the constraints achieved when neglecting photo-z errors. The 1D constraints on the cosmological parameters are $\Omega_m = 0.280^{+0.070}_{-0.065}$, $\sigma_8 = 0.784^{+0.101}_{-0.090}$. The credibility intervals and bias are very similar to those from neglecting IAs. Again, the bias from neglecting photo-z errors is within the 68.3% credibility interval from the fiducial case. We add the caveat that we have used the covariance matrix estimated from the overlapping redshift distributions (z(PDF)), however this is a simplification and the covariance distribution measured from z_p will be different. We would expect the covariance from non-overlapping redshift bins to be smaller since the signal in different bins will be less strongly correlated, hence the size of the credibility intervals on this plot would decrease and the bias may become more significant. The plot shows that photo-z errors in CFHTLenS are not large enough to significantly bias constraints on the fitted cosmological parameters. Since the photo-z errors in CFHTLenS are smaller than for previous surveys (see §7.2), this is not a surprising result. It is unknown whether this conclusion holds for other cosmological parameters, however; in particular strongly redshift dependent parameters such as the dark energy equation of state parameters may be more strongly biased by this effect. Future surveys with smaller statistical errors will be more sensitive to a shift in best-fit parameters, since smaller credibility intervals mean that they may exclude the true cosmology to a higher degree of significance than CFHTLenS if such biases are not accounted for, meaning these effects may be significant for upcoming lensing surveys.

6.7 Summary and Conclusion

In this chapter we have used the covariance matrix estimated in Chapter 5 to determine the optimum tomographic binning for cosmic shear analysis of CFHTLenS data. To achieve this we have:

- Used singular value decomposition to construct a pseudo-inverse of the covariance matrix.

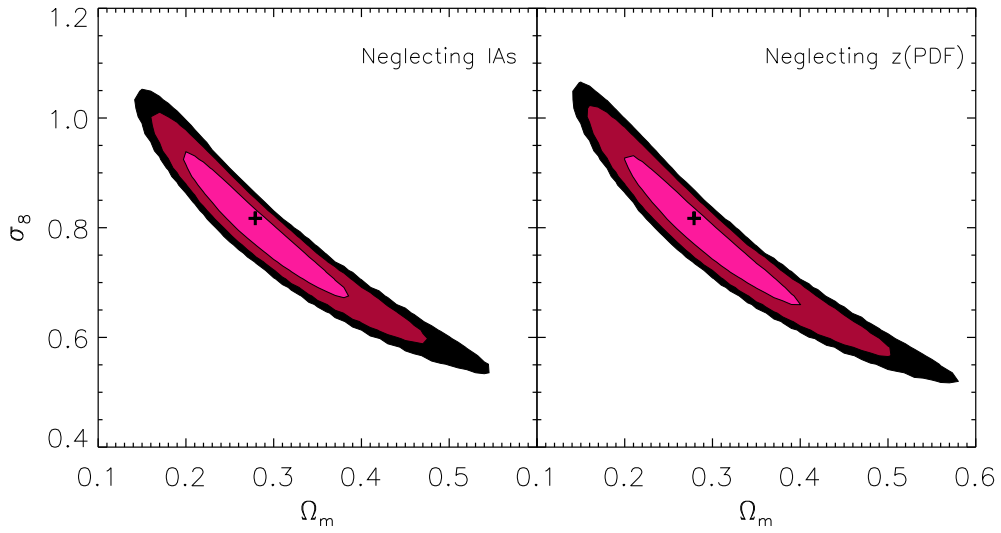

Figure 6.7: *Left*: Ω_m-σ_8 credibility intervals when the presence of IAs is ignored, for the optimal bin combination of $n_z = 6$, $n_\theta = 8$ from a chain of length ~ 5 x 10^5. *Right*: credibility intervals for the same bin combination assuming the raw z_p distribution represents the underlying redshift redshift distribution.

- Performed a least-squares fit using the pseudo-inverse to the theoretical prediction of the shear correlation function with the MCMC code described in Chapter 3, for $1 \leq n_z \leq 15$ tomographic bins and $3 \leq n_\theta \leq 15$ angular bins. We use a Gaussian WMAP7 prior on h_0 and allow Ω_m, σ_8 and A to vary, fixing all other parameters to their fiducial values.

- Finely gridded the cosmological parameter space and summed the points in each grid space to estimate the posterior directly from the relative density of points.

- Determined the optimal bin combination to be degenerate and equal to approximately $n_z = 6$, $n_\theta = 8$ for the following metrics: the area of the 68.3% Ω_m-σ_8 likelihood contour, the width and height of the 68.3% Ω_m-σ_8 contour at the fiducial cosmology, and the width of the marginalised A 68.3% likelihood contour.

- Additionally investigated the bias from neglecting IAs or photo-z errors in the likelihood analysis, finding a statistically insignificant bias in parameter

constraints for both cases (assuming a perfectly-known NLA model with $A \sim 1$).

A key result in this chapter is the dependence of the minimisation of the metrics on the number of simulated lines of sight in the covariance matrix. The 'optimal' binning we present here is only optimal for $N = 736$ lines of sight. In the future, it would be desirable to investigate the behaviour of the optimisation metrics with a greater number of realisations to see if they clearly begin to plateau. It is currently unknown whether the optimisation metrics can be significantly minimised further, or what value of N is required to cause the metrics to plateau, beyond which no further gain in precision occurs due to the finite area of a given survey. One could also investigate the relationship between the desired precision on a parameter (or combination of parameters) and the minimum number of lines of sight needed to attain it, which may have relevance to future analyses that wish to maximise information gain with minimal computing time.

The results we have presented in this chapter are calibrated to the CFHTLenS, however it is not known whether we would find the same optimum binning for a survey with different redshift distribution and/or galaxy distribution (we reiterate that the optimum binning is, however, independent of survey area since the covariance matrix scales inversely with area assuming a constant survey depth, see §4.4.1). The covariance matrix from such a survey may be quite different from the estimate we present in Chapter 5, which in turn will affect the behaviour of the credibility intervals we measure. This is pertinent since upcoming wide-field lensing surveys such as KIDS, DES and Euclid will have the power to place tighter constraints on cosmological parameters than previous lensing surveys, and tomography will be invaluable in making lensing a competitive cosmological probe. With this in mind, an extension to this work would be to investigate how the optimum binning changes with key survey parameters such as median redshift or galaxy density.

Finally, we note that in this analysis we have only allowed three parameters to vary, since these are the parameters of interest to us. We cannot say with certainty that marginalising over other cosmological parameters does not affect the optimisation. For this reason, an improvement to this work would be to perform the optimisation over more cosmological parameters (such as the dark energy equation of state parameters w_0 and w_a) although this approach will

result in some loss of constraining power on Ω_m and σ_8, and since the parameter space is larger, the MCMC chains would take longer to achieve convergence. Additionally, one could fit different intrinsic alignment models to the data and investigate whether the optimal binning is altered. Since Ω_m and σ_8 are only weakly dependent on the choice of alignment model, however, it is likely that only the constraints on the IA model itself would be substantially altered.

Overall we have found that tomography has the ability to substantially improve cosmological constraints when applied to a realistic covariance matrix from a contemporary lensing survey, although the improvement is not as large as previous studies with simpler treatment of the redshift distribution and IA contamination such as those by [126] and [152]. Even in the presence of IAs and photometric redshift errors, we can expect a factor ~ 3 improvement in cosmological constraints using optimal binning, although this comes with the caveat that the NLA model may be an incomplete one, and future tomographic lensing analyses will have to ensure that any IA model used is as accurate and well-tested as possible. In the next chapter, we test the results of this optimisation on CFHTLenS data to produce preliminary tomographic parameter constraints.

Chapter 7

Tomography with CFHTLenS

7.1 Introduction

In the previous chapter we used N-body weak lensing simulations to establish the optimum combination of tomographic and angular bin numbers to best constrain Ω_m, σ_8 and A. Using this optimum binning improved parameter constraints by a factor of three compared to a 2D analysis. In this chapter, we apply the result of this optimal tomographic binning to data from the CFHTLenS survey. With a preliminary analysis of CFHTLenS data, we aim to examine whether the optimisation derived in Chapter 6 performs as well on real data as it does on simulated data.

We perform a least-squares fit as described in Chapter 6 on the CFHTLenS $\hat{\xi}_+$ data vector, measured over $n_z = 6$ tomographic bins and $n_\theta = 8$ angular scales for which the optimisation metrics are the smallest. Since we have developed a covariance matrix that closely matches the error expected on CFHTLenS in Chapters 4 & 5, we utilise this in the likelihood analysis. This covariance matrix represents the closest estimate we have to the true tomographic covariance of CFHTLenS. As in the previous chapter, we fit only the cosmological parameters Ω_m and σ_8, as well as the normalisation constant for the NLA model, A. Since the optimisation was performed on only these three parameters, a higher-dimensional analysis may have a different bin combination as its optimum. We do not know how well the NLA model represents IAs in a survey like CFHTLenS. Even if the model is insufficient to accurately describe galaxy alignments, the constraints we achieve on A will give us valuable information on the magnitude of the effect of

IAs in CFHTLenS. A value of A that is significantly higher than 1, for example, tells us that there is strong evolution between measures of A at low redshift and at the redshift of the CFHTLenS galaxies (see §5.2) and may also mean that the IA covariance terms investigated in Chapter 5 may not be negligible in all cases. The analysis presented in this Chapter represents an important step towards a full and accurate simultaneous treatment of IAs and photo-z uncertainties in cosmic shear analysis.

In this chapter, we briefly review the status of CFHTLenS. We summarise the effect of systematic errors in a lensing survey of such unprecedented size and the steps the CFHTLenS collaboration have taken to minimise them. We discuss the method used for measuring $\hat{\xi}_+$ from the data catalogues, and the usage of the data vector in our MCMC least-squares algorithm. We then present cosmological and IA constraints, and draw our conclusions.

7.2 CFHTLens Systematics

CFHTLenS is the largest lensing survey conducted to date, and the resulting reduction in statistical errors that come from such a large survey area has uncovered systematic sources of error in standard lensing analysis techniques. These systematics are large enough to potentially cause catastrophic biases in parameter estimates in both CFHTLenS and future wide-field surveys. For this reason, the CFHTLenS collaboration was formed to develop improved analysis techniques to reduce systematics down to an acceptable level, i.e. below the level of the statistical errors. The CFHTLenS catalogues will be made publicly available in November 2012, after the collaboration releases the results of its analysis of the full 154 square degrees of data.

The upcoming release of CFHTLenS results represents the third in a series. The first results were of the 4 square degrees of CFHTLS-Deep data [154] and of $\sim$ 22 square degrees of CFHTLS-Wide data [124], in which joint 2D constraints were placed on σ_8. Both analyses found B-modes consistent with zero, implying an absence of significant systematic errors. However, the analysis of a later data release of 57 square degrees of CFHTLS-Wide data by [50] had small enough statistical errors to uncover a significant B-mode on large angular scales, which was the first hint that systematic errors would be a problem for the

full CFHTLenS data set. Despite this, they did not find a bias in cosmological constraints when compared to WMAP3 results. A subsequent analysis of the same data by [86] found that there was significant variation in the shear signal between different MegaCAM pointings, implying that there were problems with the PSF calibration. Additionally, they found that the shear correlation function did not properly scale with redshift when the data were combined with photo-z information.

To correct for such systematics, CFHTLenS have re-developed every stage of the data analysis pipeline. The `THELI` data reduction pipeline [210] was first applied to CFHTLS data in [211] and several improvements were made for CFHTLenS. `THELI` performs photometric and astrometric corrections to data, as well as automating the masking process to identify saturated stars, satellite trails, image ghosts and other artefacts. All CFHTLS fields were visually inspected to ensure the masking was of sufficient quality. This was one of the main data-related tasks undertaken by the author for the CFHTLenS collaboration. A key development in `THELI` for CFHTLenS is the improvement of its cosmic ray rejection algorithm to prevent a significant fraction of stars needed for PSF measurement being wrongly masked as cosmic rays.

One of the key improvements in the lensing analysis is the more accurate measurement of both the PSF and galaxy shapes. Previous CFHTLS analyses have been conducted using the KSB+ shape measurement method (an improved version of KSB discussed in §2.3.2) [50] and `SHAPELETS` [212, 155, 92]. CFHTLenS use `LENSFIT` (see §2.3.2) to measure galaxy shapes, which is the only shape measurement method that can be applied optimally to individual exposures rather than to image stacks. Since the PSF varies from exposure to exposure, this allows one to fit complex spatially- and temporally-varying PSFs to each galaxy individually. `LENSFIT` was found to be the best shape measurement method for reducing the exposure-to-exposure variation in the shear detected by [86]. The `LENSFIT` algorithm has been improved for use with CFHTLenS; key improvements include better size and ellipticity priors, the use of a two-component (bulge + disk) galaxy model and the simultaneous analysis of individual exposures rather than averaging on a stack [213, 102].

Another major improvement that the CFHTLenS collaboration have made to the analysis pipeline is the determination of accurate photo-zs for every survey

galaxy using multiband data. BPZ [178] is used to estimate photo-zs; [214] found this to be one of the most accurate photo-z determination codes. Accounting for the colour variation across each source through a new technique called PSF homogenisation is key to improving the photo-z accuracy [177]. Calibration of photo-zs with available spectroscopic redshifts show CFHTLenS has an outlier rate of less than 4% and a scatter of $\sigma \sim 0.04(1+z)$ [177, 102]. The scatter is on a par with that measured from HST COSMOS and COMBO-17 and the outlier rate is much lower (see Chapter 5).

The net effect of the CFHTLenS analysis is a slight residual bias to the shear $\epsilon^{\rm true}$ on the sub-percent level [213]. This takes the form of both a multiplicative m and an additive c bias to the observed complex ellipticity,

$$\epsilon^{\rm obs} = (1+m)\epsilon^{\rm true} + c + n, \tag{7.1}$$

where n is the shot noise term. The residual 'm' systematic was estimated by least-squares comparison with the signal measured from simulations used in the STEP and GREAT challenges (see Chapter 2). The multiplicative bias is thought to be caused by a high fraction of small, noisy galaxies for which accurate shape measurement is difficult. The additive bias, c, is likely caused by imperfect PSF correction, and is very small since the PSF correction from `LENSFIT` is accurate to a high degree. The CFHTLenS collaboration correct for the bias empirically since its causes are currently not fully understood.

A 2D cosmic shear analysis of the full CFHTLenS catalogues is presented in [208]. Several ΛCDM cosmological parameters are fitted with the Population Monte Carlo code of [148] (see §3.5.2). Joint lensing and WMAP7 constraints result in $\Omega_m = 0.263^{+0.013}_{-0.012}$ and $\sigma_8 = 0.805^{+0.014}_{-0.016}$. Flatness is not assumed, but the results are consistent with flat ΛCDM and WMAP7 results. The authors assume that the IA signal will be subdominant to the shear in a 2D analysis and hence neglect it. Another key difference between the analysis of [208] and this thesis is the construction of the covariance matrix. [208] estimate $\hat{\mathbf{C}}^{NN,NN}$ from the clones in a similar method to that presented in this thesis. They then 'graft' this on to an analytically-determined Gaussian covariance $\mathbf{C}^{GG,GG}$. The cross term $\mathbf{C}^{GG,NN}$ is determined using the fitting formula of [182]. The Gaussian covariance terms can be calculated exactly in this way, however as before no analytical expressions for the tomographic covariance that include non-Gaussianity exist so this approach

is not applicable in a tomographic analysis.

[209] present a 2-bin tomographic shear analysis of CFHTLenS using the full photometric PDFs but neglecting IAs. They construct the covariance matrix identically to [208]. The authors use PMC to fit parameters in a non-flat ΛCDM paradigm. They find lensing-only constraints of $\Omega_m = 0.23 \pm 0.08$, $\sigma_8 = 0.80 \pm 0.15$ and $\Omega_\Lambda = 0.27 \pm 0.26$. Under a general ΛCDM model, the authors find the dark energy density parameter is unusually low and not consistent with flat ΛCDM to 99% confidence. Overall, the 2-bin constraints are consistent with the 2D results in [208], with tighter constraints on the joint Ω_m-σ_8 contour as one would expect from a tomographic analysis.

The CFHTLenS systematic tests have been calibrated for both 2D and 2-bin tomographic analysis. It is expected that high-bin tomography, although not thoroughly tested, will be capable of producing accurate cosmological constraints. However, it is possible that high-bin tomographic analysis of the data will uncover redshift-dependent systematics that have previously been masked by coarse redshift binning. If this is the case, we may see significant tension in best-fit cosmological parameter estimates and their credibility intervals when compared to the 2D and 2-bin constraints of [208] and [209] or with the independent WMAP constraints of [26].

7.3 Measuring $\hat{\xi}_+$

All CFHTLenS galaxies with a `LENSFIT` shape with magnitude $i_{AB} < 24.7$ (the limiting magnitude for which spectroscopic redshifts are available for calibration, [102]) are used to measure $\hat{\xi}_+$. The additive bias only affects ϵ_2 and is removed by subtraction. To account for the multiplicative bias, one does not simply divide $\epsilon^{\rm obs}$ by $(1+m)$. This could result in instabilities when $(1+m) \to 0$ and more importantly such an approach relies on the estimate of m for each galaxy being accurate and unbiased, an assumption which has not been verified. Instead, we must calculate the uncalibrated shear correlation function as described in Eqn. 2.52, where the number of pairs $N_p(\theta)$ is replaced by the `LENSFIT`-weighted pair sum,

$$\hat{\xi}_\pm(\theta) = \frac{\sum_{ij} w_i w_j (\epsilon_{i\rm t}\epsilon_{j\rm t} \pm \epsilon_{i\rm r}\epsilon_{j\rm r})}{\sum_{ij} w_i w_j}. \tag{7.2}$$

The correction due to the m is then estimated as [113]

$$1 + K(\theta) = \frac{\sum_{ij} w_i w_j [1 + m\,(\nu_{SN,i}, r_i)][1 + m\,(\nu_{SN,j}, r_j)]}{\sum_{ij} w_i w_j}. \tag{7.3}$$

We omit the delta function from Eqn. 2.52 for clarity. m has the functional form

$$m(\nu_{SN}, r) = \frac{\beta}{\log\,(\nu_{SN})} \exp^{-r\,\nu_{SN}\alpha}, \tag{7.4}$$

where ν_{SN} is the signal-to-noise, r is position and α and β are constants with best-fit values of $\alpha = 0.057$ and $\beta = -0.37$ [113]. The bias-corrected shear correlation function estimator is then

$$\xi_\pm^{\rm cal}(\theta) = \frac{\xi_\pm(\theta)}{1 + K(\theta)}. \tag{7.5}$$

This approach is less likely to introduce new bias as it assumes that we can correct the survey as an ensemble. Since simulations were used to calibrate the ensemble through the shear correlation function, this is a good assumption. As we use the bias-calibrated shear in the analysis, we will refer to $\xi_+^{\rm cal}$ as ξ_+ from now on. The correlation function from each of the four Wide fields was calculated with `ATHENA`, correlating every galaxy pair in the field on the same angular scales used in the optimisation ($1' \leq \theta \leq 85'$). Since each field is a different size and has a unique number density, the mean correlation function was found from the pair-weighted average of all four fields,

$$\bar{\xi}_+(\theta) = \frac{\sum_i N_p(\theta)\,\xi_{+,i}(\theta)}{\sum_i N_p(\theta)}, \tag{7.6}$$

for $i = 1...4$.

We produce tomographic catalogues for $n_{\rm z} = 6$ by cutting in z_p as described in §4.2.3. It is important to construct a data vector for which the covariance matrix developed in Chapters 4 & 5 is a good estimate of the error. We therefore measure $\hat{\xi}_+$ on the same angular scales as in previous chapter, for the optimal case of $n_\theta = 8$.

Of the 171 MegaCAM pointings that make up the effective 154 sq. degrees of CFHTLenS data, recent work by the CFHTLenS collaboration has determined that only 139 of these (equivalent to ~ 124 sq. degrees) are able to pass a range of systematic tests and are hence of high enough quality for lensing analysis [102].

The results presented in the previous three chapter include a covariance matrix scaled to a survey area of 154 sq. degrees, meaning that the covariance matrices and corresponding credibility intervals we have measured are underestimated. To produce credibility intervals for the CFHTLenS, we must therefore rescale the covariance matrix by a factor of $171/139 = 1.23$. This will increase the size of the credibility intervals measured in Chapter 6, but it will not change the results of the optimisation.

To incorporate the CFHTLenS data into the χ^2 calculation, we simply replace the data vector from the previous chapter (the `NICAEA` prediction for the fiducial cosmology) with the CFHTLenS data vector measured in this chapter. Since the covariance is that used in the tomographic optimisation but rescaled for the reduced survey area, we expect the likelihood contours from CFHTLenS to be similar in shape and size for this analysis. The best-fit values of the measured parameters may be different, and this will be the key result in this analysis.

7.4 Cosmological Constraints

In Fig. 7.1 we plot the 2D measurement of $\hat{\xi}_+$ from the CFHTLenS catalogues measured on 8 angular scales. Error bars are 1σ variances as determined from the covariance matrix estimated in Chapter 5 and rescaled by a factor of 1.23 to take into account the fraction of fields that pass systematics tests (see previous section). The results are consistent with the CFHTLenS estimate of $\hat{\xi}_+$ in Fig. 2 of [208], and can be seen to be slightly above the correlation function predicted from WMAP 7-year mean parameters [26] and the nonlinear model for the power spectrum of [28] (incorporating the NLA model intrinsic alignment signal with $A = 1$). In Fig. 7.2 we show the tomographic correlation function from the CFHTLenS when the data are split into 6 tomographic bins and measured over 8 angular bins. Again, the WMAP 7 + NLA model prediction is plotted, and the points tend to lie slightly above the model prediction on average.

Our procedure for the likelihood analysis is identical to that in Chapter 6; we again constrain only Ω_m, σ_8 and A, using the same priors. We use a WMAP

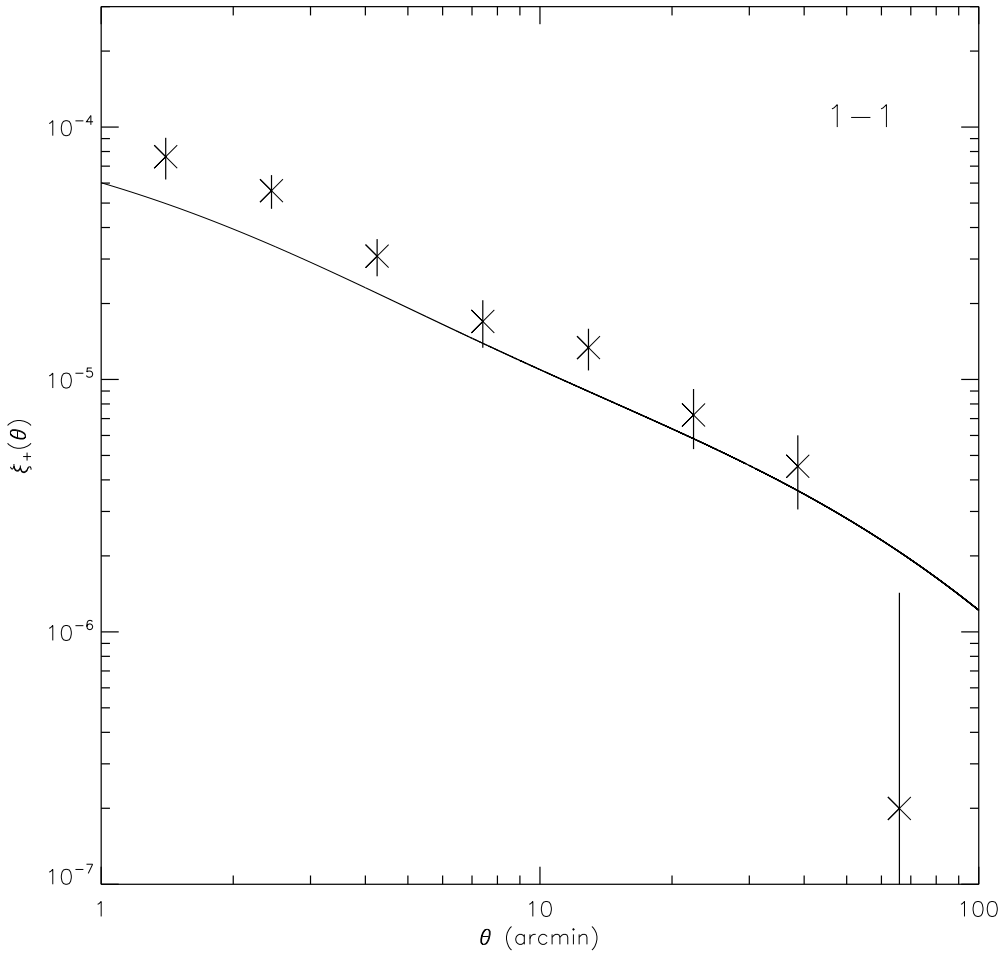

Figure 7.1: The shear correlation function measured from CFHTLenS for the 2D case of $n_z = 1$, $n_\theta = 8$. Error bars are the (scaled) covariance determined from the clone simulations in Chapter 5. Solid line represents WMAP7 cosmology [26] incorporating the NLA model intrinsic alignment signal with $A = 1$.

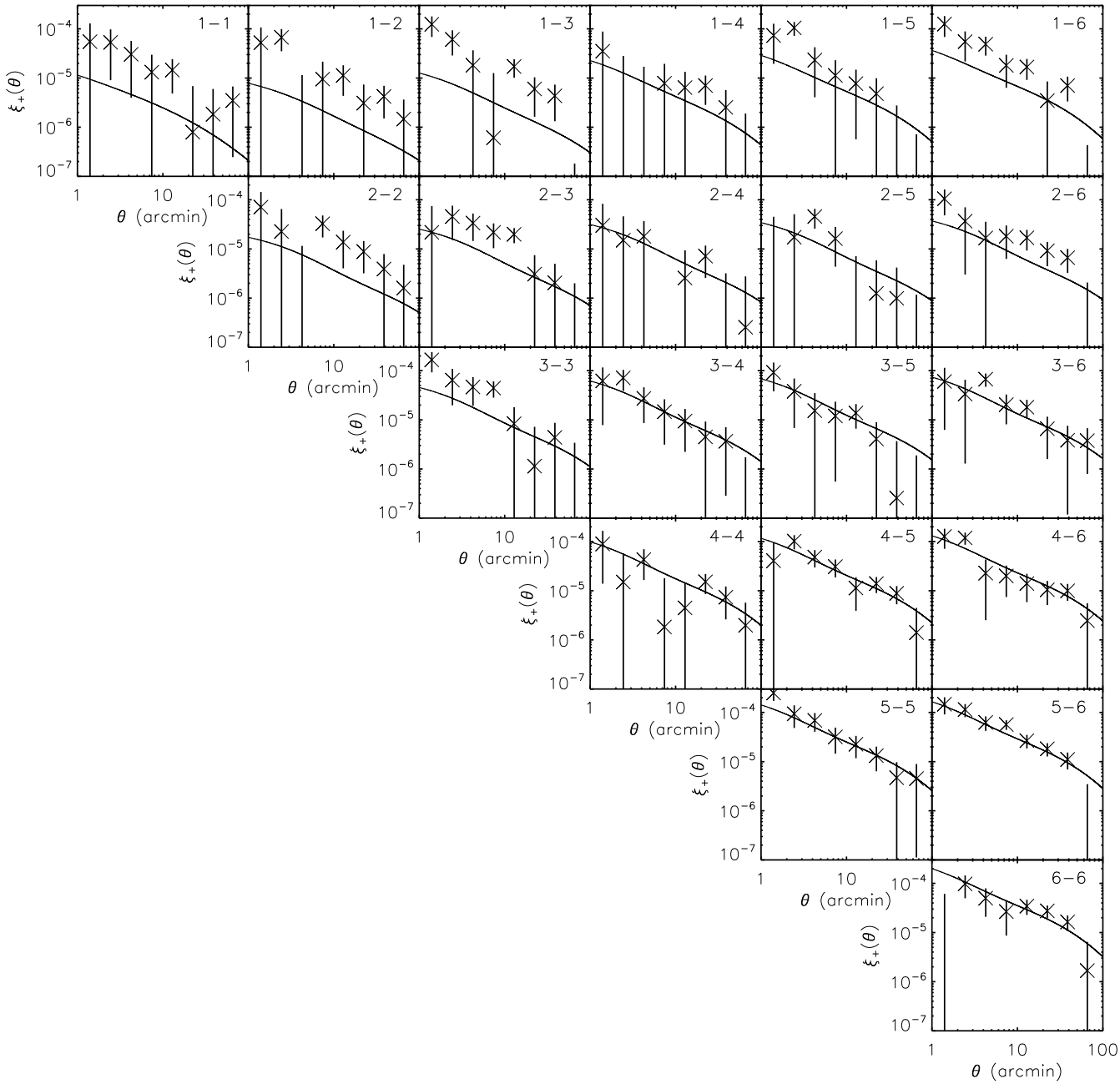

Figure 7.2: As in Fig. 7.1 for $n_z = 6$, $n_\theta = 8$. The redshift correlation being plotted is shown in the top right corner of each subplot.

prior on h_0 and fixing all other cosmological parameters to their fiducial values as inputted into the clone simulations. In Fig. 7.3 we plot the constraints from the 2D CFHTLenS shear correlation function with $n_\theta = 8$. The contours are of similar shape and size to those presented in the tomography optimisation in Fig. 6.1, as expected. We find marginalised constraints on the three parameters of $\Omega_m = 0.388^{+0.137}_{-0.129}$, $\sigma_8 = 0.728^{+0.157}_{-0.143}$ and $A = 0.448^{+1.277}_{-0.415}$ (all 68.3% CL). Ω_m and σ_8 are consistent to the WMAP 7-year means of $\Omega_m = 0.266 \pm 0.029$ and $\sigma_8 = 0.811^{+0.030}_{-0.031}$ to within 68.3% likelihood. Fig. 7.3 shows that the joint constraint on Ω_m and σ_8 favours a higher value than WMAP. It is currently unknown whether this contour is accurate, or if there is some unforeseen systematic in the data or simply tension in the fit due to assuming an incorrect prior for some parameter(s). Since [209] find under a general ΛCDM model that the CFHTLenS data prefer a low value of Ω_Λ, it is possible that by assuming flatness, we have introduced tension in the fit.

Our constraints on the size of the Ω_m-σ_8 contour broadly match the 2D CFHTLenS results of [208], although their results prefer a slightly lower value of Ω_m. [208] measure both $\hat{\xi}_+$ and $\hat{\xi}_-$ over a wider range of angular scales than us, which may account for the difference in best-fit value. We note that our 2D 68.3% Ω_m and σ_8 constraints overlap with the 2-bin tomography constraints for CFHTLenS from [209] for a flat ΛCDM universe. Our marginalised constraints on Ω_m and σ_8 are similar in precision to those in Chapter 6, and the area of the Ω_m-σ_8 contour is similar (2.95 x 10^{-2} compared to 2.66 x 10^{-2} in Fig. 6.1). The small difference is likely due to the smaller final CFHTLenS survey area.

The constraining power present on A is similar to that in the optimisation in Chapter 6. Even with a 2D analysis, very high values of $A \sim 10$ are ruled out to an extremely high degree of confidence. The asymmetrical confidence limits on A are most likely a result of the prior value of $A > 0$, meaning that not only are our constraints on A consistent with zero to 95.4% confidence, but it is possible that there is substantial probability in the posterior below $A = 0$. If the data favour a negative value of A, this means that galaxies are anti-correlated on average on the scales probed.

In Fig. 7.4 we plot CFHTLenS constraints from the optimal bin combination found in Chapter 6 of $n_z = 6$, $n_\theta = 8$. The marginalised constraints are now

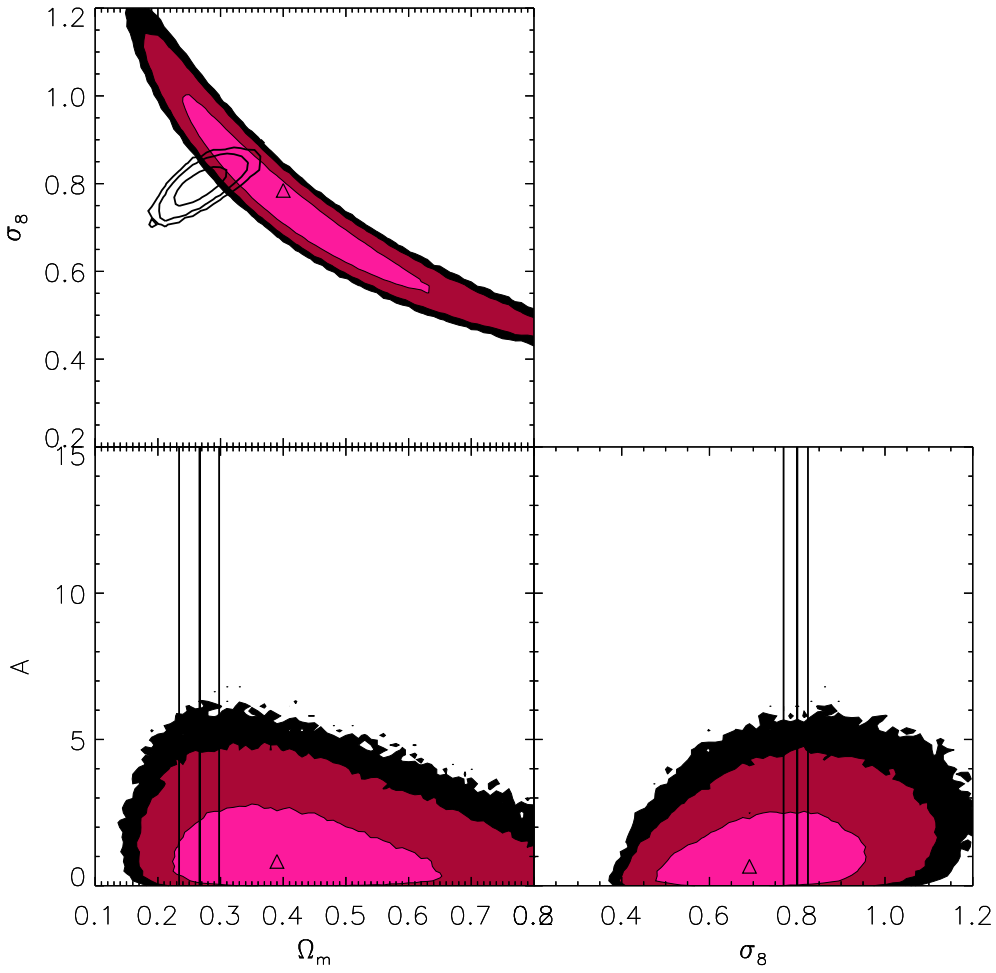

Figure 7.3: 68.3%, 95.4% and 99.0% contour intervals on Ω_m, σ_8 and A for the 2D case $n_z = 1$, $n_\theta = 8$ from a chain of length ~ 5 x 10^5. WMAP 7 confidence intervals are shown by black solid lines; the 68.3%, 95.4% and 99.0% contours are shown in the Ω_m-σ_8 plane and the best fit and 68.3% interval are shown for the Ω_m-A and σ_8-A planes as vertical lines. The peak of the likelihood surface from CFHTLenS is indicated with a triangle.

$\Omega_m = 0.412^{+0.122}_{-0.094}$, $\sigma_8 = 0.643^{+0.106}_{-0.092}$ and $A = 0.003^{+0.498}_{-0.003}$. The improvement in constraining power is similar to that found in Chapter 6 between the 2D and the optimal binning cases. 1D constraints on Ω_m and σ_8 have improved slightly with the optimal binning, and the area of the Ω_m-σ_8 contour is 1.47 x 10^{-2}, a factor of ~ 2 smaller than for the 2D case (compared to a factor of ~ 3 for the optimisation). This more modest improvement in constraining power may be due to the fact that in the optimisation presented in Chapter 6, we fit to the theoretical prediction of the shear correlation function rather than the correlation function measured from data or simulations. This may alter the constraints in a tomography-dependent way. A lower value of A is favoured compared to the 2D case, however they are both consistent with zero IA signal under the NLA model. The credibility intervals on A for $n_z = 6$ are smaller than for the optimal contours in Chapter 6. This may be caused by the truncation at $A = 0$ masking some of the posterior and the true difference in constraining power present between the two cases. The best-fit values of all fitted parameters are consistent to within 68.3% confidence of the 2D constraints, indicating that high-bin tomography has not introduced or uncovered significant bias in the data.

7.4.1 Relaxing the prior on A

Since we see some tension in the 2D joint constraint on Ω_m and σ_8 when compared to WMAP values, and since the data favour a low value of A that is consistent with zero, we relax the hard prior on A to investigate whether this alters and improves the constraints. We allow A to vary within the range $-15 \leq A \leq 15$ and re-run the analysis for $n_z = 1$ and $n_z = 6$. The results are plotted in Figs. 7.5-7.6. The marginalised constraints for the 2D analysis in Fig. 7.5 are now $\Omega_m = 0.475^{+0.205}_{-0.183}$, $\sigma_8 = 0.528^{+0.219}_{-0.118}$ and $A = 0.199^{+1.658}_{-2.488}$. The constraints on the cosmological parameters are consistent with the constraints for the case of $A > 0$, although the parameter estimates and joint credibility interval have been degraded somewhat. This is to be expected since we have increased the size of the parameter space. A is still consistent with zero to 68.3% confidence, although there is a large amount of posterior in the region $A < 0$.

Fig. 7.6 shows the constraints for the optimal tomographic binning. From this, $\Omega_m = 0.419^{+0.123}_{-0.090}$, $\sigma_8 = 0.623^{+0.101}_{-0.084}$ and $A = -1.161^{+1.163}_{-0.597}$. The constraints

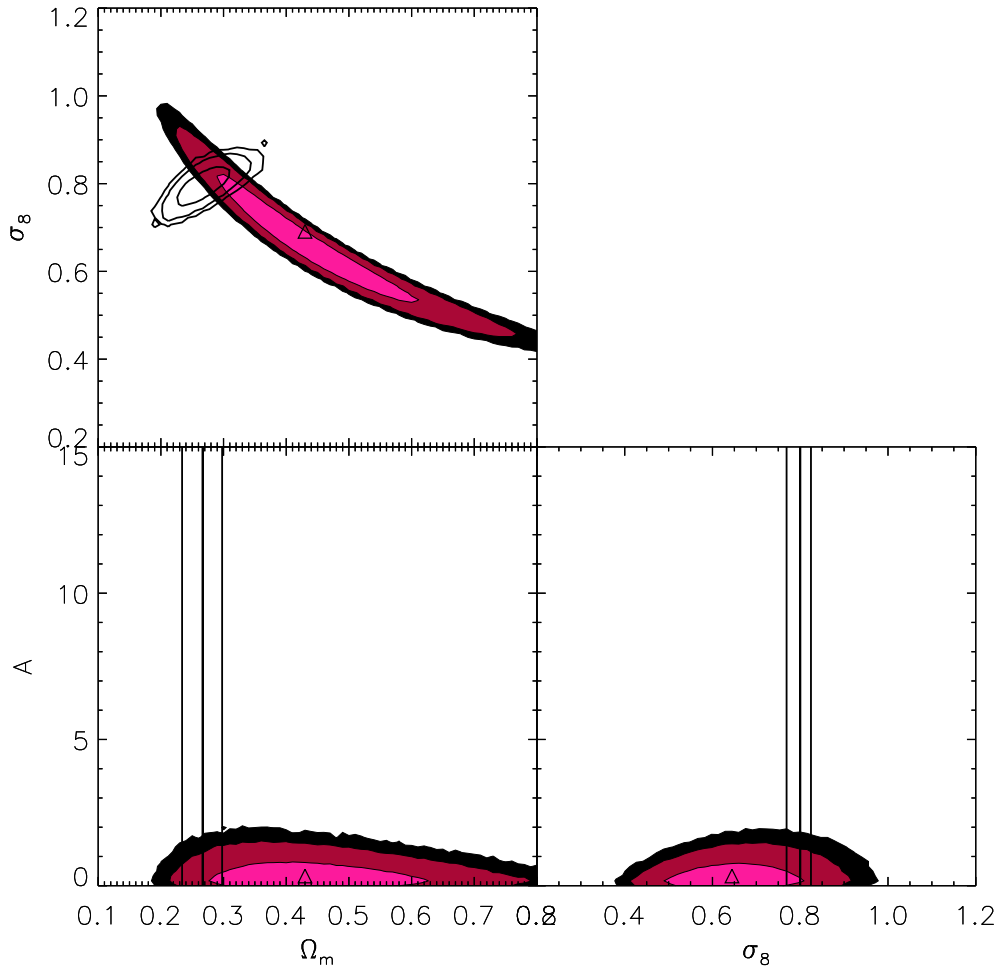

Figure 7.4: 68.3%, 95.4% and 99.0% contour intervals on Ω_m, σ_8 and A for the optimal case $n_z = 6$, $n_\theta = 8$ from a chain of length $\sim 5 \times 10^5$. WMAP 7 confidence intervals are shown by black solid lines; the 68.3%, 95.4% and 99.0% contours are shown in the Ω_m-σ_8 plane and the best fit and 68.3% interval are shown for the Ω_m-A and σ_8-A planes as vertical lines. The peak of the likelihood surface from CFHTLenS is indicated with a triangle.

on Ω_m and σ_8 are consistent with the results that use the prior $A > 0$ for the optimal binning (Fig. 7.4), within similar joint and marginalised credibility limits. The degenerate 95.4% likelihood contour encompasses the WMAP 7-year mean. Interestingly, the tomographic results now show a trend towards a slightly negative value of A. This may explain the unexpectedly large improvement in the constraint on A between Fig. 7.3 and Fig. 7.4; some of the posterior was unavailable to us as it lay below $A = 0$. A more in-depth analysis will be needed to determine whether the trend towards negative A is significant. Overall, we note that the result shown in Fig. 7.6 may be interpreted to mean that we find no significant detection of IAs under the fiducial NLA model. We also note that the 2D and tomographic 68.3% credibility intervals for the case of $-15 \leq A \leq 15$ are still consistent with each other, indicating that the previous prior of $A > 0$ was not masking a tomography-specific systematic bias.

7.5 Conclusion

In this chapter, we have performed a preliminary analysis of the two-point correlation function from CFHTLenS data to determine whether the optimisation derived in Chapter 6 performs as well on data as it does on simulations. We have discussed some of the early CFHTLS lensing analyses that hinted at the presence of serious systematic errors in the data. We have briefly reviewed some of the key improvements the CFHTLenS collaboration has made in an attempt to mitigate these systematics, in anticipation of both the CFHTLenS full-survey results and for future wide-field lensing surveys. We then described our method for extracting the correlation function from the CFHTLenS catalogues and presented the shear correlation functions for the 2D and optimal tomographic case. Finally, we perform a linear least-square fit to the data using the MCMC algorithm developed in Chapter 3 to produce constraints on Ω_m, σ_8 and the amplitude parameter of the NLA model A for both the 2D and the optimal tomographic case, with and without the hard prior $A > 0$, to determine whether the optimisation of Chapter 6 is capable of producing similar improvements in constraints on real data as it is on simulated data.

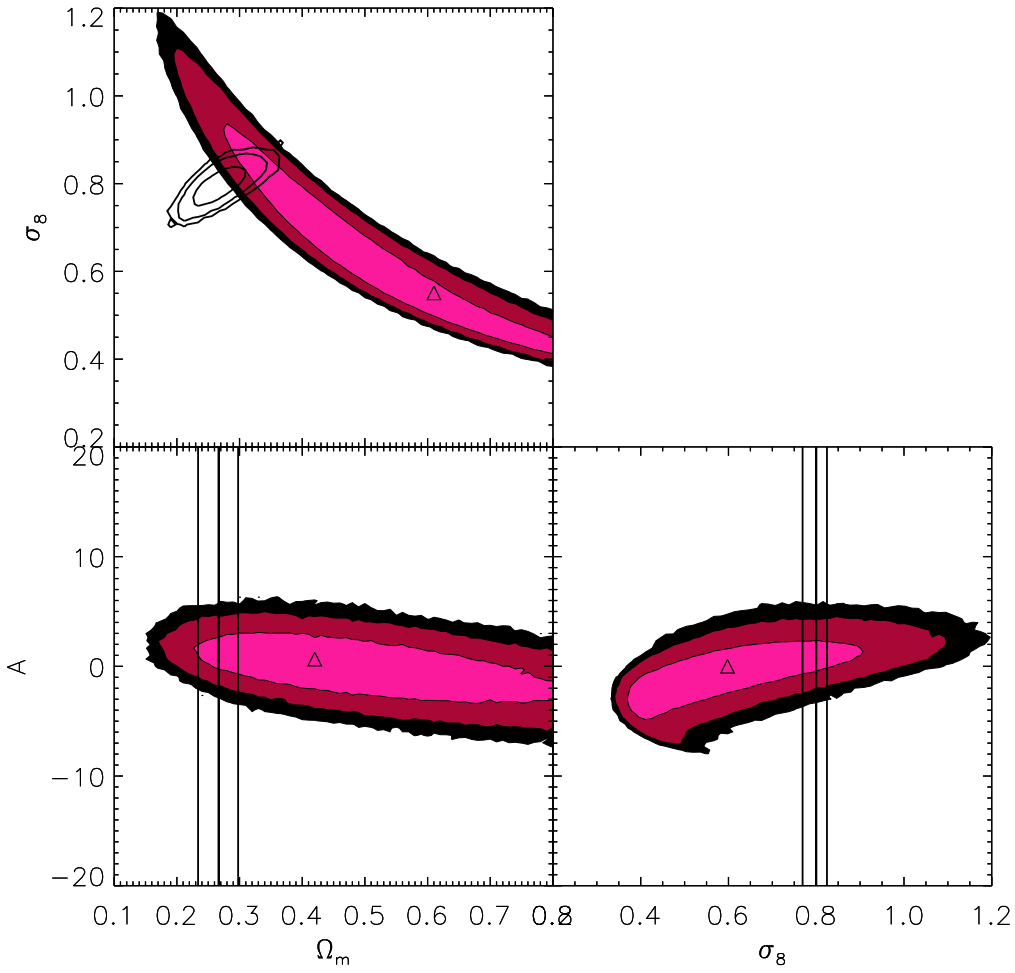

Figure 7.5: 68.3%, 95.4% and 99.0% contour intervals on Ω_m, σ_8 and A for the 2D case $n_z = 1$, $n_\theta = 8$ from a chain of length ~ 5 x 10^5, with $-15 \leq A \leq 15$. WMAP 7 confidence intervals are shown by black solid lines; the 68.3%, 95.4% and 99.0% contours are shown in the Ω_m-σ_8 plane and the best fit and 68.3% interval are shown for the Ω_m-A and σ_8-A planes as vertical lines. The peak of the likelihood surface from CFHTLenS is indicated with a triangle.

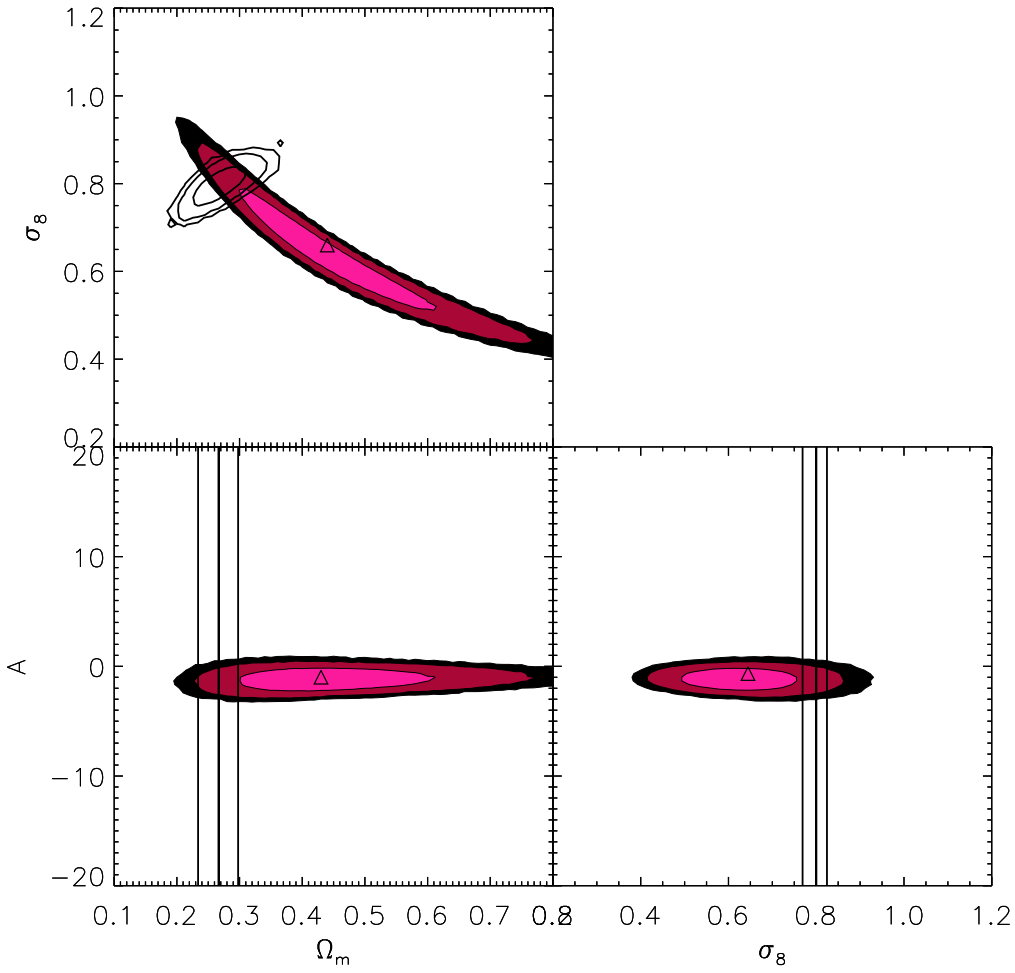

Figure 7.6: 68.3%, 95.4% and 99.0% contour intervals on Ω_m, σ_8 and A for the optimal case $n_z = 6$, $n_\theta = 8$ from a chain of length ~ 5 x 10^5, with $-15 \leq A \leq 15$. WMAP 7 confidence intervals are shown by black solid lines; the 68.3%, 95.4% and 99.0% contours are shown in the Ω_m-σ_8 plane and the best fit and 68.3% interval are shown for the Ω_m-A and σ_8-A planes as vertical lines. The peak of the likelihood surface from CFHTLenS is indicated with a triangle.

When using the same priors as in Chapter 6, our results are generally consistent with similar analyses of the CFHTLenS data by [208] and [209]. We find that Ω_m and σ_8 are consistent with WMAP 7-year constraints to within 68.3% likelihood, although the joint credibility interval favours somewhat higher values of these parameters than WMAP. Although it is important to let the data 'speak for themselves' and not make assumptions about what we believe the results ought to be, it is also important to check thoroughly for systematic errors when one finds an unexpected result. It is possible that inaccuracies in the hard priors we have assumed on other cosmological parameters are the cause of this result. We attempt to investigate this by relaxing the prior on A to allow negative values. This results in similar constraints on Ω_m and σ_8 as the values for the prior of $A > 0$. In every case, both the 2D and tomographic case favour a value of A that is consistent with zero to 95.4% CL, ruling out a significant IA signal under the NLA model. We find that the improvement in parameter constraints when going from 1 to 6 tomographic bins is somewhat smaller than the optimisation in Chapter 6 predicts, but the factor of improvement in Ω_m-σ_8 contour area is still substantial at ~ 2. Importantly, we find that all the constraints we achieve for both $n_\mathrm{z} = 1$ and $n_\mathrm{z} = 6$ are consistent with each other to within 68.3% likelihood, meaning that there is no evidence of tomography- or redshift-specific systematics in the CFHTLenS data.

The analysis presented in this chapter is preliminary, and further work will investigate the limitations of high-bin tomography with this data set. This will involve fitting more parameters to the data - in particular, highly redshift-dependent ones such as the dark energy equation of state parameters - to see how much constraining power CFHTLenS has over them, and to investigate how much constraints on Ω_m and σ_8 are degraded by their inclusion. To investigate the effect of different galaxy populations, the shear signal can be measured for red and blue galaxies separately. Additionally, other IA models can be fitted to the data, or the signal from IAs downweighted by one of the methods described in §5.1 and the effect on cosmological parameters measured. It is clear that the CFHTLenS data set has much to tell us about the nature of the Universe, and only by carefully considering how we can jointly account for systematic sources of error and optimise the analysis of the data will we realise the full potential of both the CFHTLenS and future lensing surveys.

Chapter 8

Conclusions

In this chapter, we summarise the importance of the work presented in this thesis and outline the main results. We then discuss the future prospects for cosmic shear studies and the current challenges faced by this field.

8.1 Weak lensing tomography

The observation of weak lensing is a field that is rapidly developing, with cosmic shear measurements producing interesting and useful constraints on cosmological parameters. In particular, weak lensing has the ability to put constraints on the Ω_m-σ_8 contour that are orthogonal to those achieved from CMB data, which shows its value as a cosmological probe that is complimentary to other established techniques. As photometric redshift information is now available for every survey galaxy as standard, it is possible to exploit the 3D information in the lensing signal to tighten cosmological constraints. A tomographic analysis in which data is sampled into discrete redshift bins recovers much of the statistical information that is lost in a 2D analysis. This allows us to produce tighter constraints on cosmological parameters, with highly redshift-dependent parameters such as the dark energy equation of state parameters, w_0 and w_a and the Hubble parameter, h_0, gaining the most.

Tomography is a conceptually simple technique that is starting to be fully exploited with the advent of CFHTLenS. It has been shown to be capable of producing parameter constraints that are improved by up to a factor of ten for a 4-bin analysis [152]. The previous tomographic analyses of [154, 155, 156] were

conducted on small fields of only a few square degrees. The statistical errors were therefore large enough to mask any systematic bias caused by intrinsic alignments or errors in the photo-z distribution. Surveys such as the CFHTLenS, with survey areas around two orders of magnitude (or more) larger than previous tomographic analyses, have diminishing statistical errors and will hence be affected by both these systematic sources of error. However, if systematics can be kept under control, tomography will be invaluable in producing tighter parameter constraints from weak lensing than ever before.

8.2 Intrinsic alignment contamination

When conducting a tomographic analysis, intrinsic galaxy alignments become one of the main sources of systematic error in weak lensing measurements. Our current understanding of IAs is hindered by incomplete knowledge of gas dynamics and the nonlinear evolution of the density field, which affect galaxy formation. Failure to take IAs into account properly may bias the lensing signal by up to $\sim 10\%$ [186], hence it is vital that we deal with IAs as carefully and accurately as possible. Currently, several methods exist to mitigate the effect of IAs, falling into two broad categories: optimisation methods to downweight or null the signal [191, 194, 195], which result in information loss and may not reliably account for bias; and modelling techniques, which if incorrect or incomplete may introduce further bias. Previous cosmic shear tomographic analyses have either ignored IAs or only partially accounted for their effects. For example, [156] attempt to account for IAs by downweighting the II signal by removing auto-correlations between narrow redshift bins, however this treatment is inadequate due to the heavy information loss incurred and the failure to remove the GI signal.

The non-linear alignment model of [119] and BK07 models IAs using an extension of the linear alignment model of [196] into the nonlinear regime. This model has been tested at low redshift by [199], who find constraints on the amplitude of the NLA model that are consistent with BK07. BK07 find that if IAs are accounted for, two or more times as many redshift bins are required to recover 80% of the lensing information, which highlights the need to incorporate their effects in a tomographic analysis. Current and upcoming lensing surveys like CFHTLenS will have the power to constrain the amplitude of the NLA model,

however we do not yet know how accurately the NLA model describes galaxy alignments and a more sophisticated model for IAs may be needed for future surveys to avoid systematic bias.

8.3 Covariance matrix estimation

In Chapter 4, we discussed the importance of measuring the covariance matrix of the shear correlation function, since it is highly correlated. Any likelihood analysis that aims to produce parameter constraints must use an estimator for the covariance matrix inverse, $\hat{\mathbf{C}}^{-1}$, and it is important that it is accurate and unbiased or we risk wrongly estimating the precision of our parameter constraints. In Chapters 4 & 5 we aimed to achieve two things: to investigate the behaviour of the different terms in the correlation function covariance, and to construct a stable estimate of the inverse for use in a likelihood analysis with as high a number of tomographic bins as possible. Although exact analytical expressions for the covariance of several weak lensing statistics exist [116], these do not account for non-Gaussianity and are hence unfit for a realistic analysis. For this reason, we used both Gaussian [176] and non-Gaussian [163] shear simulations to measure the covariance of the two-point shear correlation function. We attempted to use the empirical correction of [180] to add the contribution from non-Gaussianity into the correlation function from the Gaussian simulations, but found that the correction was not suitable for a tomographic analysis. This meant that the use of non-Gaussian simulations is necessary for a realistic tomographic analysis.

We tested the validity of the Anderson correction [169], which places a hard limit on the number of data points p (and hence tomographic bins) one can use in a likelihood analysis, for a given number of realisations of the data N. We found that the Anderson correction works well in the region where $p/N \leq 0.8$, and that it would be feasible to perform up to 15-bin tomography with the CFHTLenS data if the covariance is estimated from all N-body simulations available at this time.

In Chapter 5, we extended this analysis by incorporating both shot noise and correlated intrinsic ellipticity components into the measurement of $\hat{\xi}_+$. The intrinsic ellipticity was predicted from the NLA model for IAs of [119] and BK07. We measured the covariance of each component of $\hat{\mathbf{C}}^{\mathrm{TOT}}$ and found that in general

the dominant contribution is the shot noise covariance $\hat{\mathbf{C}}^{NN,NN}$ (although $\hat{\mathbf{C}}^{GG,GG}$ dominates on scales above $\sim 30'$ for the 2D case). The relative amplitude of the shot noise covariance increases with tomography since the number of galaxies in each bin falls. The IA covariance terms contribute very little to the total covariance, indicating that they can be neglected if the amplitude of the NLA model is not significantly higher than the fiducial value. Finally, we investigated the stability of the covariance matrix to inversion, and found that for all tomographic combinations that we wish to investigate, the covariance matrix has a low enough condition number to ensure it is invertible. The addition of shot noise and an IA model does not degrade the stability of the matrix, and therefore 15-bin tomography is still possible with our estimate of the full covariance.

The N-body simulations are as close to the CFHTLenS in input as possible - with galaxy positions, source clustering, galaxy density, shot noise and photometric redshift estimates that match the CFHTLenS. This means that, once scaled to the area of the CFHTLenS, the tomographic covariance matrix we measure is as good an estimate of the covariance of the CFHTLenS as it is currently possible to get. This ensures that the optimisation presented in Chapter 6, and the cosmological constraints from CFHTLenS we achieved in Chapter 7, are as accurate and unbiased as possible.

8.4 Optimal tomography

The optimal tomographic binning and relative improvement in constraints over a 2D analysis for a survey such as CFHTLenS has not previously been investigated. For this reason, we utilised the tomographic covariance matrix developed in Chapter 5 to determine the optimum tomographic binning for the CFHTLenS. To do this, we developed and tested an MCMC algorithm to efficiently explore the parameter space, which we discussed in Chapter 3. We investigated the behaviour of the Ω_m-σ_8 contour which is well constrained by cosmic shear and is sensitive to tomography, as well as the amplitude of the NLA model through the normalisation parameter A. We chose 4 values as our optimisation metrics - the area of the 68.3% Ω_m-σ_8 likelihood contour, the height and width of this contour about the fiducial value, and the width of the 1D 68.3% credibility interval on A.

We presented the results of this optimisation in Chapter 6. We found that

there is an optimum number of data points p that minimises the metrics and hence there is a partial degeneracy between the optimal n_z and n_θ. The optimisation metrics minimise at slightly different values of n_z and n_θ, with the combination $n_z = 6$, $n_\theta = 8$ in the middle of this range. Hence we took this combination as the overall optimum. We found that the limiting factor in determining the optimum bin combination is the number of realisations of the simulations used to calculate the covariance. This means that the optimum binning may shift to higher n_z and n_θ with more lines of sight, and that there will be a minimum number of realisations needed to achieve a desired degree of precision in parameter estimates. We saw a factor of three improvement in the tomographic area of the Ω_m-σ_8 contour compared to the 2D value, which is smaller than the improvement found by [152] who neglect IAs and photo-z errors. We also found that the data are precise enough to produce interesting constraints on A, indicating that CFHTLenS (and future wide-field surveys) will have significant constraining power on the NLA model. This in turn can tell us about the galaxy population that contributes to IAs, and potentially compare IA models with each other. Additionally, we investigated the biases in cosmological parameter estimates caused by neglecting photometric redshift errors or IAs, and found no statistically significant bias for a CFHTLenS-like survey. This is good news for the CFHTLenS survey, although we add the caveat that future ultra wide-field surveys with smaller statistical errors may find that this bias is more significant.

8.5 Cosmology with CFHTLenS and in the future

CFHTLenS is currently the widest deep lensing survey for cosmology. Spanning 154 sq. degrees with a median redshift of $z_m = 0.75$, CFHTLenS is ~ 100 times larger than all previous surveys used for tomographic analysis. Due to the large survey area and consequently small statistical uncertainties, preliminary analyses of the data indicated the presence of systematic errors [50, 86]. To remedy this, the CFHTLenS collaboration have redeveloped every stage of the analysis pipeline, with significant improvements in particular in photo-z error estimation, and galaxy shape and PSF measurement. CFHTLenS represents a stepping stone between previous small-field lensing surveys such as HST

COSMOS and CFHTLS-Deep and the next generation of ultra wide-field lensing surveys covering thousands to tens-of-thousands of square degrees, such as KIDS, DES and Euclid discussed in §2.3.6. The work undertaken by the CFHTLenS collaboration has changed the standard methods for lensing analysis. The lessons learned from CFHTLenS will be vital if we are to produce optimal constraints from future lensing surveys.

In Chapter 7, we performed a preliminary least-squares analysis of the two-point correlation function from CFHTLenS data. Using the MCMC algorithm developed in Chapter 3 and knowledge of the optimal tomographic and angular binning from Chapter 6, we produced constraints on Ω_m, σ_8 and A for both the 2D and the optimal tomographic case. The results were generally consistent with similar analyses of the CFHTLenS data by [208] and [209]. We found that although the marginalised constraints on Ω_m and σ_8 are consistent with WMAP 7-year constraints to within 68.3% likelihood, a slightly higher value for the joint constraint is preferred by the data. A is consistent with zero, ruling out a significant IA signal under the NLA model for both the 2D and tomographic cases. We found that the improvement in parameter constraints when going from 1 to 6 tomographic bins is slightly smaller than the optimisation in Chapter 6 predicts, but the factor of improvement in Ω_m-σ_8 contour area is quite substantial at ~ 2. Importantly, we find that all the constraints we achieve for both $n_z = 1$ and $n_z = 6$ are consistent with each other to within 68.3% likelihood, meaning that there is no evidence of tomography- or redshift-specific systematics in the CFHTLenS data. The results of Chapter 7 are interesting, highlighting some of the gaps in our knowledge that are still present. We leave a deeper, more thorough tomographic analysis of the data to the CFHTLenS collaboration, who will most likely improve upon the constraints and analysis techniques in this thesis to move lensing forward as a field.

Future lensing analyses will use a wide variety of techniques to maximise the information recovered from data. In addition to the two-point shear correlation function, other shear statistics may be used to constrain cosmological parameters, each with different benefits. These include the aperture mass dispersion [166] and the shear dispersion in circular apertures [165], as well as three-point statistics of these in which three galaxies at a time are correlated instead of two [215]. Additionally, the magnification κ was recently detected for the first time and,

since it is independent of the shear, may be used to break degeneracies between cosmological parameters [216].

Although the CFHTLenS collaboration has made great strides forward in reducing the effect of systematics in lensing data, some outstanding issues remain. Shape measurement is an ongoing issue, and despite the improvements made to `LENSFIT` by the CFHTLenS, it is important to develop better shape measurement methods for future surveys. To this end, the GREAT challenge is ongoing (see §2.3.2). More accurate photo-z estimation may also be required for ultra wide surveys, and IA contamination is likely to become a substantial contributor to the systematic errors until we better understand its cause. Despite these issues, weak lensing has already proven to be a useful and independent cosmological probe, and future ultra wide surveys will ensure it is a competitive one that can help cosmologists discover vital and fascinating information about our Universe.

Bibliography

[1] P. Coles, *Routledge Critical Dictionary of the New Cosmology.* Routledge, 2001.

[2] E. Hubble, "A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae," *Contributions from the Mount Wilson Observatory, vol. 3, pp.23-28*, vol. 3, pp. 23–28, 1927.

[3] E. Hubble and M. L. Humason, "The Velocity-Distance Relation for Isolated Extra-Galactic Nebulae," *Contributions from the Mount Wilson Observatory, vol. 3, pp.85-89*, vol. 3, pp. 85–89, 1927.

[4] A. A. Penzias and R. W. Wilson, "A Measurement of Excess Antenna Temperature at 4080 Mc/s.," *ApJ*, vol. 142, pp. 419–421, July 1965.

[5] P. J. E. Peebles, "The Black-Body Radiation Content of the Universe and the Formation of Galaxies.," *ApJ*, vol. 142, p. 1317, Nov. 1965.

[6] R. Maartens, "Is the Universe homogeneous?," *Royal Society of London Philosophical Transactions Series A*, vol. 369, pp. 5115–5137, Dec. 2011.

[7] J. Peacock, *Cosmological Physics.* Cambridge University Press, 1999.

[8] E. Thomas and D. Raine, *An introduction to the science of cosmology.* Institute of Physics, CRC Press, 2001.

[9] T. Kitching, *Constraining the Dark Energy Equation of State using Three Dimensional Weak Lensing.* PhD thesis, Institute for Astronomy, University of Edinburgh, 2006.

[10] A. Friedman, "Über die Krümmung des Raumes," *Zeitschrift fur Physik*, vol. 10, pp. 377–386, Dec. 1922.

[11] A. Friedmann, "Über die Möglichkeit einer Welt mit konstanter negativer Krümmung des Raumes," *Zeitschrift fur Physik*, vol. 21, pp. 326–332, Dec. 1924.

[12] A. G. Riess, L. Macri, S. Casertano, H. Lampeitl, H. C. Ferguson, A. V. Filippenko, S. W. Jha, W. Li, and R. Chornock, "A 3% Solution: Determination of the Hubble Constant with the Hubble Space Telescope and Wide Field Camera 3," *ApJ*, vol. 730, p. 119, Apr. 2011.

[13] L. A. Barnes, M. J. Francis, J. B. James, and G. F. Lewis, "Joining the Hubble flow: implications for expanding space," *MNRAS*, vol. 373, pp. 382–390, Nov. 2006.

[14] J. A. Peacock, "A diatribe on expanding space," *ArXiv e-prints*, Sept. 2008.

[15] M. D. Lehnert, N. P. H. Nesvadba, J.-G. Cuby, A. M. Swinbank, S. Morris, B. Clément, C. J. Evans, M. N. Bremer, and S. Basa, "Spectroscopic confirmation of a galaxy at redshift z = 8.6," *Nature*, vol. 467, pp. 940–942, Oct. 2010.

[16] P. J. E. Peebles and J. T. Yu, "Primeval Adiabatic Perturbation in an Expanding Universe," *ApJ*, vol. 162, p. 815, Dec. 1970.

[17] P. J. E. Peebles, *The large-scale structure of the universe*. 1980.

[18] J. M. Bardeen, "Gauge-invariant cosmological perturbations," *Phys. Rev. D*, vol. 22, pp. 1882–1905, Oct 1980.

[19] S. M. Carroll, W. H. Press, and E. L. Turner, "The cosmological constant," *Annual review of astronomy and astrophysics*, vol. 30, pp. 499–542, 1992.

[20] C. Heymans, *Weak Gravitational Lensing and Intrinsic Galaxy Alignments*. PhD thesis, St. Catherine's college, Oxford, 2003.

[21] K. Brand, S. Rawlings, G. J. Hill, and M. Lacy, "Using radio galaxies to find super-structures," *New Astronomy Reviews*, vol. 47, pp. 325–328, Sept. 2003.

[22] S. D. M. White, M. Davis, and C. S. Frenk, "The size of clusters in a neutrino-dominated universe," *MNRAS*, vol. 209, pp. 27P–31P, July 1984.

[23] T. Padmanabhan, *Structure Formation in the Universe*. June 1993.

[24] E. R. Harrison, "Fluctuations at the Threshold of Classical Cosmology," *Phys. Rev. D*, vol. 1, pp. 2726–2730, May 1970.

[25] Y. B. Zeldovich, "A hypothesis, unifying the structure and the entropy of the Universe," *MNRAS*, vol. 160, p. 1P, 1972.

[26] E. Komatsu, K. M. Smith, J. Dunkley, C. L. Bennett, B. Gold, G. Hinshaw, N. Jarosik, D. Larson, M. R. Nolta, L. Page, D. N. Spergel, M. Halpern, R. S. Hill, A. Kogut, M. Limon, S. S. Meyer, N. Odegard, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright, "Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation," *ArXiv e-prints*, Jan. 2010.

[27] J. A. Peacock and S. J. Dodds, "Non-linear evolution of cosmological power spectra," *MNRAS*, vol. 280, pp. L19–L26, June 1996.

[28] R. E. Smith, J. A. Peacock, A. Jenkins, S. D. M. White, C. S. Frenk, F. R. Pearce, P. A. Thomas, G. Efstathiou, and H. M. P. Couchman, "Stable clustering, the halo model and non-linear cosmological power spectra," *MNRAS*, vol. 341, pp. 1311–1332, June 2003.

[29] M. Bartelmann and P. Schneider, "Weak gravitational lensing," *Phys. Rep.*, vol. 340, pp. 291–472, Jan. 2001.

[30] L. Van Waerbeke and Y. Mellier, "Gravitational Lensing by Large Scale Structures: A Review," *ArXiv Astrophysics e-prints*, May 2003.

[31] L. Van Waerbeke, Y. Mellier, M. Radovich, E. Bertin, M. Dantel-Fort, H. J. McCracken, O. Le Fèvre, S. Foucaud, J. Cuillandre, T. Erben, B. Jain, P. Schneider, F. Bernardeau, and B. Fort, "Cosmic shear statistics and cosmology," *A&A*, vol. 374, pp. 757–769, Aug. 2001.

[32] J. Einasto, M. Einasto, E. Tago, A. A. Starobinsky, F. Atrio-Barandela, V. Müller, A. Knebe, P. Frisch, R. Cen, H. Andernach, and D. Tucker, "Steps toward the Power Spectrum of Matter. I. The Mean Spectrum of Galaxies," *ApJ*, vol. 519, pp. 441–455, July 1999.

[33] P. de Bernardis, P. A. R. Ade, J. J. Bock, J. R. Bond, J. Borrill, A. Boscaleri, K. Coble, B. P. Crill, G. De Gasperis, P. C. Farese, P. G. Ferreira, K. Ganga, M. Giacometti, E. Hivon, V. V. Hristov, A. Iacoangeli, A. H. Jaffe, A. E. Lange, L. Martinis, S. Masi, P. V. Mason, P. D. Mauskopf, A. Melchiorri, L. Miglio, T. Montroy, C. B. Netterfield, E. Pascale, F. Piacentini, D. Pogosyan, S. Prunet, S. Rao, G. Romeo, J. E. Ruhl, F. Scaramuzzi, D. Sforna, and N. Vittorio, "A flat Universe from high-resolution maps of the cosmic microwave background radiation," *Nature*, vol. 404, pp. 955–959, Apr. 2000.

[34] A. G. Riess, A. V. Filippenko, P. Challis, A. Clocchiatti, A. Diercks, P. M. Garnavich, R. L. Gilliland, C. J. Hogan, S. Jha, R. P. Kirshner, B. Leibundgut, M. M. Phillips, D. Reiss, B. P. Schmidt, R. A. Schommer, R. C. Smith, J. Spyromilio, C. Stubbs, N. B. Suntzeff, and J. Tonry, "Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant," *AJ*, vol. 116, pp. 1009–1038, Sept. 1998.

[35] G. Aldering, S. Perlmutter, R. A. Knop, P. Nugent, G. Goldhaber, D. E. Groom, M. Y. Kim, C. R. Pennypacker, S. Deustua, R. Quimby, A. Goobar, R. Pain, S. Fabbro, I. M. Hook, C. Lidman, A. Kim, B. E. Schaefer, R. Ellis, M. Irwin, N. Walton, P. Ruiz-Lapuente, A. S. Fruchter, N. Panagia, and Supernova Cosmology Project Collaboration, "Measurements of Omega and Lambda from High-Redshift Supernovae," in *Bulletin of the American Astronomical Society*, vol. 30 of *Bulletin of the American Astronomical Society*, pp. 1305–+, Dec. 1998.

[36] S. Carrol, "The cosmological constant," *Living Rev. Relativity*, vol. 4, 2001.

[37] P. J. Peebles and B. Ratra, "The cosmological constant and dark energy," *Reviews of Modern Physics*, vol. 75, pp. 559–606, Apr. 2003.

[38] NASA/WMAP, "Fate of the universe," Aug 2010. `http://www.map.gsfc.nasa.gov/universe`.

[39] S. Weinberg, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*. July 1972.

[40] R. A. Knop, G. Aldering, R. Amanullah, P. Astier, G. Blanc, M. S. Burns, A. Conley, S. E. Deustua, M. Doi, R. Ellis, S. Fabbro, G. Folatelli, A. S. Fruchter, G. Garavini, S. Garmond, K. Garton, R. Gibbons, G. Goldhaber, A. Goobar, D. E. Groom, D. Hardin, I. Hook, D. A. Howell, A. G. Kim, B. C. Lee, C. Lidman, J. Mendez, S. Nobili, P. E. Nugent, R. Pain, N. Panagia, C. R. Pennypacker, S. Perlmutter, R. Quimby, J. Raux, N. Regnault, P. Ruiz-Lapuente, G. Sainton, B. Schaefer, K. Schahmaneche, E. Smith, A. L. Spadafora, V. Stanishev, M. Sullivan, N. A. Walton, L. Wang, W. M. Wood-Vasey, and N. Yasuda, "New Constraints on Ω_M, Ω, and w from an Independent Set of 11 High-Redshift Supernovae Observed with the Hubble Space Telescope," *ApJ*, vol. 598, pp. 102–137, Nov. 2003.

[41] R. Scranton, A. J. Connolly, R. C. Nichol, A. Stebbins, I. Szapudi, D. J. Eisenstein, N. Afshordi, T. Budavari, I. Csabai, J. A. Frieman, J. E. Gunn, D. Johnson, Y. Loh, R. H. Lupton, C. J. Miller, E. S. Sheldon, R. S. Sheth, A. S. Szalay, M. Tegmark, and Y. Xu, "Physical Evidence for Dark Energy," *ArXiv Astrophysics e-prints*, July 2003.

[42] K. Glazebrook, C. Blake, W. Couch, D. Forbes, M. Drinkwater, R. Jurek, K. Pimbblet, B. Madore, C. Martin, T. Small, K. Forster, M. Colless, R. Sharp, S. Croom, D. Woods, M. Pracy, D. Gilbank, H. Yee, and M. Gladders, "The WiggleZ project: AAOmega and Dark Energy," *ArXiv Astrophysics e-prints*, Jan. 2007.

[43] M. Tegmark, "Cosmological lessons from MAP and SDSS," *APS April Meeting Abstracts*, pp. 1003–+, Apr. 2003.

[44] L. Anderson, E. Aubourg, S. Bailey, D. Bizyaev, M. Blanton, A. S. Bolton, J. Brinkmann, J. R. Brownstein, A. Burden, A. J. Cuesta, L. N. A. da Costa, K. S. Dawson, R. de Putter, D. J. Eisenstein, J. E. Gunn, H. Guo, J.-C. Hamilton, P. Harding, S. Ho, K. Honscheid, E. Kazin, D. Kirkby, J.-P. Kneib, A. Labatie, C. Loomis, R. H. Lupton, E. Malanushenko, V. Malanushenko, R. Mandelbaum, M. Manera, C. Maraston, C. K. McBride, K. T. Mehta, O. Mena, F. Montesano, D. Muna, R. C. Nichol, S. E. Nuza, M. D. Olmstead, D. Oravetz, N. Padmanabhan, N. Palanque-Delabrouille, K. Pan, J. Parejko, I. Paris, W. J. Percival, P. Petitjean, F. Prada, B. Reid, N. A. Roe, A. J. Ross, N. P. Ross, L. Samushia, A. G. Sanchez, D. J. S. D. P. Schneider, C. G. Scoccola, H.-J. Seo, E. S. Sheldon, A. Simmons, R. A. Skibba, M. A. Strauss, M. E. C. Swanson, D. Thomas, J. L. Tinker, R. Tojeiro, M. Vargas Magana, L. Verde, C. Wagner, D. A. Wake, B. A. Weaver, D. H. Weinberg, M. White, X. Xu, C. Yeche, I. Zehavi, and G.-B. Zhao, "The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: Baryon Acoustic Oscillations in the Data Release 9 Spectroscopic Galaxy Sample," *ArXiv e-prints*, Mar. 2012.

[45] B. Bassett and R. Hlozek, *Baryon acoustic oscillations*, pp. 246–+. 2010.

[46] D. J. Eisenstein, W. Hu, J. Silk, and A. S. Szalay, "Can Baryonic Features Produce the Observed 100 H -1 MPC Clustering?," *ApJ Lett.*, vol. 494, p. L1, Feb. 1998.

[47] D. J. Eisenstein, D. H. Weinberg, E. Agol, H. Aihara, C. Allende Prieto, S. F. Anderson, J. A. Arns, É. Aubourg, S. Bailey, E. Balbinot, and et al., "SDSS-III: Massive Spectroscopic Surveys of the Distant Universe, the Milky Way, and Extra-Solar Planetary Systems," *AJ*, vol. 142, pp. 72–+, Sept. 2011.

[48] C. Blake, T. Davis, G. B. Poole, D. Parkinson, S. Brough, M. Colless, C. Contreras, W. Couch, S. Croom, M. J. Drinkwater, K. Forster, D. Gilbank, M. Gladders, K. Glazebrook, B. Jelliffe, R. J. Jurek, I.-H. Li, B. Madore, D. C. Martin, K. Pimbblet, M. Pracy, R. Sharp, E. Wisnioski, D. Woods, T. K. Wyder, and H. K. C. Yee, "The WiggleZ Dark Energy Survey: testing the cosmological model with baryon acoustic oscillations at z= 0.6," *MNRAS*, vol. 415, pp. 2892–2909, Aug. 2011.

[49] SNAP Collaboration, "Probing Dark Energy via Weak Gravitational Lensing with the SuperNova Acceleration Probe (SNAP)," *ArXiv Astrophysics e-prints*, July 2005.

[50] L. Fu, E. Semboloni, H. Hoekstra, M. Kilbinger, L. van Waerbeke, I. Tereno, Y. Mellier, C. Heymans, J. Coupon, K. Benabed, J. Benjamin, E. Bertin, O. Doré, M. J. Hudson, O. Ilbert, R. Maoli, C. Marmo, H. J. McCracken, and B. Ménard, "Very weak lensing in the CFHTLS wide: cosmology from cosmic shear in the linear regime," *A&A*, vol. 479, pp. 9–25, Feb. 2008.

[51] A. Heavens, "Cosmology with Gravitational Lensing," *ArXiv e-prints*, Sept. 2011.

[52] E. V. Linder, "Dark Energy, Expansion History of the Universe, and SNAP," in *Particle Physics and Cosmology* (J. F. Nieves & C. N. Leung, ed.), vol. 655 of *American Institute of Physics Conference Series*, pp. 193–207, Feb. 2003.

[53] P. J. E. Peebles and B. Ratra, "Cosmology with a time-variable cosmological 'constant'," *ApJ Lett.*, vol. 325, pp. L17–L20, Feb. 1988.

[54] I. Zlatev, L. Wang, and P. J. Steinhardt, "Quintessence, Cosmic Coincidence, and the Cosmological Constant," *Phys. Rev. Lett.*, vol. 82, pp. 896–899, Feb. 1999.

[55] M. Li, X.-D. Li, S. Wang, and Y. Wang, "Dark Energy," *ArXiv e-prints*, Mar. 2011.

[56] S. Weinberg, "Anthropic bound on the cosmological constant," *Phys. Rev. Lett.*, vol. 59, pp. 2607–2610, Nov. 1987.

[57] S. Kachru, R. Kallosh, A. Linde, and S. P. Trivedi, "de Sitter vacua in string theory," *Phys. Rev. D*, vol. 68, p. 046005, Aug. 2003.

[58] M. B. Green, H. J. Schwarz, and E. Witten, *Superstring Theory.* Cambridge University Press, 1987.

[59] A. G. Riess, P. E. Nugent, R. L. Gilliland, B. P. Schmidt, J. Tonry, M. Dickinson, R. I. Thompson, T. Budavári, S. Casertano, A. S. Evans, A. V. Filippenko, M. Livio, D. B. Sanders, A. E. Shapley, H. Spinrad, C. C. Steidel, D. Stern, J. Surace, and S. Veilleux, "The Farthest Known Supernova: Support for an Accelerating Universe and a Glimpse of the Epoch of Deceleration," *ApJ*, vol. 560, pp. 49–71, Oct. 2001.

[60] H. Martel, P. R. Shapiro, and S. Weinberg, "Likely Values of the Cosmological Constant," *ApJ*, vol. 492, pp. 29–+, Jan. 1998.

[61] A. Barreira and P. P. Avelino, "Anthropic versus cosmological solutions to the coincidence problem," *Phys. Rev. D*, vol. 83, pp. 103001–+, May 2011.

[62] K. Nakamura and Particle Data Group, "Review of Particle Physics," *Journal of Physics G Nuclear Physics*, vol. 37, p. 075021, July 2010.

[63] F. Zwicky, "On the Masses of Nebulae and of Clusters of Nebulae," *ApJ*, vol. 86, pp. 217–+, Oct. 1937.

[64] D. Clowe, "Beyond the Bullet: Direct Detection of Dark Matter in Merging Galaxy Clusters," in *HST Proposal*, pp. 11194–+, July 2007.

[65] A. N. Taylor, D. J. Bacon, M. E. Gray, C. Wolf, K. Meisenheimer, S. Dye, A. Borch, M. Kleinheinrich, Z. Kovacs, and L. Wisotzki, "Mapping the 3D dark matter with weak lensing in COMBO-17," *MNRAS*, vol. 353, pp. 1176–1196, Oct. 2004.

[66] T. Hamana, Y. Ohyama, M. Chiba, and N. Kashikawa, "Dark and stellar matter in strong lensing galaxies from a joint lensing and stellar dynamics," *ArXiv Astrophysics e-prints*, July 2005.

[67] A. Coc, E. Vangioni-Flam, P. Descouvemont, A. Adahchour, and C. Angulo, "Updated Big Bang Nucleosynthesis Compared with Wilkinson Microwave Anisotropy Probe Observations and the Abundance of Light Elements," *ApJ*, vol. 600, pp. 544–552, Jan. 2004.

[68] W. Percival, "Baryon acoustic oscillations," *Cosmology, Galaxy Formation and Astroparticle Physics on the pathway to the SKA*, 2006.

[69] G. Bertone and D. Merritt, "Dark Matter Dynamics and Indirect Detection," *Modern Physics Letters A*, vol. 20, pp. 1021–1036, 2005.

[70] R. Bernabei, P. Belli, F. Cappella, R. Cerulli, C. J. Dai, A. d'Angelo, H. L. He, A. Incicchitti, X. H. Ma, F. Montecchia, F. Nozzoli, D. Prosperi, X. D. Sheng, Z. P. Ye, and R. G. Wang, "Particle Dark Matter in DAMA/LIBRA," *ArXiv e-prints*, July 2010.

[71] M. Felizardo, T. Morlat, A. C. Fernandes, T. A. Girard, J. G. Marques, A. R. Ramos, M. Auguste, D. Boyer, A. Cavaillou, C. Sudre, J. Poupeney, R. F. Payne, H. S. Miley, and J. Puibasset, "First Results of the Phase II SIMPLE Dark Matter Search," *Physical Review Letters*, vol. 105, pp. 211301–+, Nov. 2010.

[72] J. Wambsganss, *Gravitational Microlensing*, pp. 453–+. 2006.

[73] J. Beaulieu, D. P. Bennett, P. Fouqué, A. Williams, M. Dominik, U. G. Jørgensen, D. Kubas, A. Cassan, C. Coutures, J. Greenhill, K. Hill, J. Menzies, P. D. Sackett, M. Albrow, S. Brillant, J. A. R. Caldwell, J. J. Calitz, K. H. Cook, E. Corrales, M. Desort, S. Dieters, D. Dominis, J. Donatowicz, M. Hoffman, S. Kane, J. Marquette, R. Martin, P. Meintjes, K. Pollard, K. Sahu, C. Vinter, J. Wambsganss, K. Woller, K. Horne, I. Steele, D. M. Bramich, M. Burgdorf, C. Snodgrass, M. Bode, A. Udalski, M. K. Szymański, M. Kubiak, T. Więckowski, G. Pietrzyński, I. Soszyński, O. Szewczyk, Ł. Wyrzykowski, B. Paczyński, F. Abe, I. A. Bond, T. R. Britton, A. C. Gilmore, J. B. Hearnshaw, Y. Itow, K. Kamiya, P. M. Kilmartin, A. V. Korpela, K. Masuda, Y. Matsubara, M. Motomura, Y. Muraki, S. Nakamura, C. Okada, K. Ohnishi, N. J. Rattenbury, T. Sako, S. Sato, M. Sasaki, T. Sekiguchi, D. J. Sullivan, P. J. Tristram, P. C. M. Yock, and T. Yoshioka, "Discovery of a cool planet of 5.5 Earth masses through gravitational microlensing," *Nature*, vol. 439, pp. 437–440, Jan. 2006.

[74] S. S. Vogt, R. P. Butler, E. J. Rivera, N. Haghighipour, G. W. Henry, and M. H. Williamson, "The Lick-Carnegie Exoplanet Survey: A 3.1 M Planet in the Habitable Zone of the Nearby M3V Star Gliese 581," *ApJ*, vol. 723, pp. 954–965, Nov. 2010.

[75] P. Tisserand, L. Le Guillou, C. Afonso, J. N. Albert, J. Andersen, R. Ansari, É. Aubourg, P. Bareyre, J. P. Beaulieu, X. Charlot, C. Coutures, R. Ferlet, P. Fouqué, J. F. Glicenstein, B. Goldman, A. Gould, D. Graff, M. Gros, J. Haissinski, C. Hamadache, J. de Kat, T. Lasserre, É. Lesquoy, C. Loup, C. Magneville, J. B. Marquette, É. Maurice, A. Maury, A. Milsztajn, M. Moniez, N. Palanque-Delabrouille, O. Perdereau, Y. R. Rahal, J. Rich, M. Spiro, A. Vidal-Madjar, L. Vigroux, S. Zylberajch, and EROS-2 Collaboration, "Limits on the Macho content of the Galactic Halo from the EROS-2 Survey of the Magellanic Clouds," *A&A*, vol. 469, pp. 387–404, July 2007.

[76] L. L. R. Williams and P. L. Schechter, "Measurement of the Hubble Constant Via Gravitational Lensing— A Review of the Jodrell Bank "Golden Lenses" Workshop," *ArXiv Astrophysics e-prints*, Sept. 1997.

[77] B. M. Dobke, L. J. King, C. D. Fassnacht, and M. W. Auger, "Estimating cosmological parameters from future gravitational lens surveys," *MNRAS*, vol. 397, pp. 311–319, July 2009.

[78] J. A. Tyson, G. P. Kochanski, and I. P. dell'Antonio, "Detailed Mass Map of CL 0024+1654 from Strong Lensing," *ApJ Lett.*, vol. 498, pp. L107+, May 1998.

[79] H. Hoekstra, H. K. C. Yee, and M. D. Gladders, "Constraints on Ω_m and σ_8 from Weak Lensing in Red-Sequence Cluster Survey Fields," *ApJ*, vol. 577, pp. 595–603, Oct. 2002.

[80] B. Jain and U. Seljak, "Cosmological Model Predictions for Weak Lensing: Linear and Nonlinear Regimes," *ApJ*, vol. 484, p. 560, July 1997.

[81] C. R. Contaldi, H. Hoekstra, and A. Lewis, "Joint Cosmic Microwave Background and Weak Lensing Analysis: Constraints on Cosmological Parameters," *Phys. Rev. Lett.*, vol. 90, pp. 221303–+, June 2003.

[82] L. Hollenstein, D. Sapone, R. Crittenden, and B. M. Schäfer, "Constraints on early dark energy from CMB lensing and weak lensing tomography," *Journal of Cosmology and Astro-Particle Physics*, vol. 4, pp. 12–+, Apr. 2009.

[83] H. Hoekstra, H. K. C. Yee, and M. D. Gladders, "The Red-Sequence Cluster Survey: first lensing results," *ArXiv Astrophysics e-prints*, Sept. 2001.

[84] U. Pen, T. Lu, L. van Waerbeke, and Y. Mellier, "The three-dimensional power spectrum of dark and luminous matter from the VIRMOS-DESCART cosmic shear survey," *MNRAS*, vol. 346, pp. 994–1008, Dec. 2003.

[85] D. Munshi, P. Valageas, L. van Waerbeke, and A. Heavens, "Cosmology with weak lensing surveys," *Phys. Rep.*, vol. 462, pp. 67–121, June 2008.

[86] M. Kilbinger, K. Benabed, J. Guy, P. Astier, I. Tereno, L. Fu, D. Wraith, J. Coupon, Y. Mellier, C. Balland, F. R. Bouchet, T. Hamana, D. Hardin, H. J. McCracken, R. Pain, N. Regnault, M. Schultheis, and H. Yahagi, "Dark-energy constraints and correlations with systematics from CFHTLS weak lensing, SNLS supernovae Ia and WMAP5," *A&A*, vol. 497, pp. 677–688, Apr. 2009.

[87] S. Hilbert, J. Hartlap, S. D. M. White, and P. Schneider, "Ray-tracing through the Millennium Simulation: Born corrections and lens-lens coupling in cosmic shear and galaxy-galaxy lensing," *A&A*, vol. 499, pp. 31–43, May 2009.

[88] R. Narayan and M. Bartelmann, "Lectures on Gravitational Lensing," *ArXiv Astrophysics e-prints*, June 1996.

[89] M. Bartelmann, "Gravitational lensing lecture notes," 2006. `http://www.oact.inaf.it/gl/lectures/Bartelmann.pdf`.

[90] J. Schneider, P. Ehlers and F. E., *Gravitational lenses*. New York: Springer, 1999.

[91] M. Bartelmann, "TOPICAL REVIEW Gravitational lensing," *Classical and Quantum Gravity*, vol. 27, p. 233001, Dec. 2010.

[92] M. Velander, K. Kuijken, and T. Schrabback, "Probing galaxy dark matter haloes in COSMOS with weak lensing flexion," *MNRAS*, vol. 412, pp. 2665–2677, Apr. 2011.

[93] A. Leonard, L. J. King, and D. M. Goldberg, "New constraints on the complex mass substructure in Abell 1689 from gravitational flexion," *MNRAS*, vol. 413, pp. 789–804, May 2011.

[94] R. Scranton, B. Ménard, G. T. Richards, R. C. Nichol, A. D. Myers, B. Jain, A. Gray, M. Bartelmann, R. J. Brunner, A. J. Connolly, J. E. Gunn, R. K. Sheth, N. A. Bahcall, J. Brinkman, J. Loveday, D. P. Schneider, A. Thakar,

and D. G. York, "Detection of Cosmic Magnification with the Sloan Digital Sky Survey," *ApJ*, vol. 633, pp. 589–602, Nov. 2005.

[95] H. Hildebrandt, L. van Waerbeke, and T. Erben, "CARS: The CFHTLS-Archive-Research Survey. III. First detection of cosmic magnification in samples of normal high-z galaxies," *A&A*, vol. 507, pp. 683–691, Nov. 2009.

[96] L. van Waerbeke, "Shear and magnification: cosmic complementarity," *MNRAS*, vol. 401, pp. 2093–2100, Jan. 2010.

[97] A. Amara, A. Refregier, and S. Paulin-Henriksson, "Cosmic Shear Systematics: Software-Hardware Balance," *ArXiv e-prints*, May 2009.

[98] N. Kaiser, G. Squires, and T. Broadhurst, "A Method for Weak Lensing Observations," *ApJ*, vol. 449, pp. 460–+, Aug. 1995.

[99] G. A. Luppino and N. Kaiser, "Detection of Weak Lensing by a Cluster of Galaxies at Z = 0.83," *ApJ*, vol. 475, pp. 20–+, Jan. 1997.

[100] H. Hoekstra, M. Franx, K. Kuijken, and G. Squires, "Weak Lensing Analysis of CL 1358+62 Using Hubble Space Telescope Observations," *ApJ*, vol. 504, pp. 636–+, Sept. 1998.

[101] C. Heymans, L. Van Waerbeke, D. Bacon, J. Berge, G. Bernstein, E. Bertin, S. Bridle, M. L. Brown, D. Clowe, H. Dahle, T. Erben, M. Gray, M. Hetterscheidt, H. Hoekstra, P. Hudelot, M. Jarvis, K. Kuijken, V. Margoniner, R. Massey, Y. Mellier, R. Nakajima, A. Refregier, J. Rhodes, T. Schrabback, and D. Wittman, "The Shear Testing Programme - I. Weak lensing analysis of simulated ground-based observations," *MNRAS*, vol. 368, pp. 1323–1339, May 2006.

[102] C. Heymans et al., "CFHTLenS: The Canada-France-Hawaii Telescope Lensing Survey (MNRAS Submitted)," 2012. `http://www.roe.ac.uk/~heymans/foo/cfhtlens_sys.pdf`.

[103] K. Kuijken, "Weak weak lensing: correcting weak shear measurements accurately for PSF anisotropy," *A&A*, vol. 352, pp. 355–362, Dec. 1999.

[104] N. Kaiser, "A New Shear Estimator for Weak-Lensing Observations," *ApJ*, vol. 537, pp. 555–577, July 2000.

[105] G. M. Bernstein and M. Jarvis, "Shapes and Shears, Stars and Smears: Optimal Measurements for Weak Lensing," *AJ*, vol. 123, pp. 583–618, Feb. 2002.

[106] A. Refregier, T. Chang, and D. J. Bacon, "Shapelets: a new method to measure galaxy shapes," in *The Shapes of Galaxies and their Dark Haloes* (P. Natarajan, ed.), pp. 29–+, 2002.

[107] A. Refregier, "Shapelets - I. A method for image analysis," *MNRAS*, vol. 338, pp. 35–47, Jan. 2003.

[108] T. Kitching, S. Balan, G. Bernstein, M. Bethge, S. Bridle, F. Courbin, M. Gentile, A. Heavens, M. Hirsch, R. Hosseini, A. Kiessling, A. Amara, D. Kirk, K. Kuijken, R. Mandelbaum, B. Moghaddam, G. Nurbaeva, S. Paulin-Henriksson, A. Rassat, J. Rhodes, B. Schölkopf, J. Shawe-Taylor, M. Gill, M. Shmakova, A. Taylor, M. Velander, L. van Waerbeke, D. Witherick, D. Wittman, S. Harmeling, C. Heymans, R. Massey, B. Rowe, T. Schrabback, and L. Voigt, "Gravitational Lensing Accuracy Testing 2010 (GREAT10) Challenge Handbook," *ArXiv e-prints*, Sept. 2010.

[109] S. Bridle, J. Shawe-Taylor, A. Amara, D. Applegate, B. Balan, S. T. Joel, G. Bernstein, H. Dahle, T. Erben, M. Gill, A. Heavens, C. Heymans, F. W. High, H. Hoekstra, M. Jarvis, D. Kirk, T. Kitching, J. Kneib, K. Kuijken, D. Lagatutta, R. Mandelbaum, R. Massey, Y. Mellier, B. Moghaddam, Y. Moudden, R. Nakajima, S. Paulin-Henriksson, S. Pires, A. Rassat, A. Refregier, J. Rhodes, T. Schrabback, E. Semboloni, M. Shmakova, L. van Waerbeke, D. Witherick, L. Voigt, and D. Wittman, "Handbook for the GREAT08 Challenge: An image analysis competition for cosmological lensing," *Annals of Applied Statistics*, vol. 3, pp. 6–37, 2009.

[110] L. Miller, T. D. Kitching, C. Heymans, A. F. Heavens, and L. van Waerbeke, "Bayesian galaxy shape measurement for weak lensing surveys - I. Methodology and a fast-fitting algorithm," *MNRAS*, vol. 382, pp. 315–324, Nov. 2007.

[111] T. D. Kitching, L. Miller, C. E. Heymans, L. van Waerbeke, and A. F. Heavens, "Bayesian galaxy shape measurement for weak lensing surveys - II. Application to simulations," *MNRAS*, vol. 390, pp. 149–167, Oct. 2008.

[112] J. P. Dietrich, A. Biviano, P. Popesso, Y. Zhang, M. Lombardi, and H. Böhringer, "Weak lensing observations of potentially X-ray underluminous galaxy clusters," *A&A*, vol. 499, pp. 669–677, June 2009.

[113] L. Miller and CFHTLenS Collaboration, "Weak Lensing Shape Measurement in CFHTLenS," in *American Astronomical Society Meeting Abstracts*, vol. 219 of *American Astronomical Society Meeting Abstracts*, Jan. 2012.

[114] M. Bartelmann and P. Schneider, "Power spectrum from weak-shear data," *A&A*, vol. 345, pp. 17–21, May 1999.

[115] M. Loverde and N. Afshordi, "Extended Limber approximation," *Phys. Rev. D*, vol. 78, p. 123506, Dec. 2008.

[116] P. Schneider, L. van Waerbeke, M. Kilbinger, and Y. Mellier, "Analysis of two-point statistics of cosmic shear. I. Estimators and covariances," *A&A*, vol. 396, pp. 1–19, Dec. 2002.

[117] F. Bernardeau, L. van Waerbeke, and Y. Mellier, "Weak lensing statistics as a probe of {OMEGA} and power spectrum.," *A&A*, vol. 322, pp. 1–18, June 1997.

[118] P. Schneider, L. van Waerbeke, and Y. Mellier, "B-modes in cosmic shear from source redshift clustering," *A&A*, vol. 389, pp. 729–741, July 2002.

[119] C. M. Hirata and U. Seljak, "Intrinsic alignment-lensing interference as a contaminant of cosmic shear," *Phys. Rev. D*, vol. 70, pp. 063526–+, Sept. 2004.

[120] G. Bernstein and D. Huterer, "Catastrophic photometric redshift errors: weak-lensing survey requirements," *MNRAS*, vol. 401, pp. 1399–1408, Jan. 2010.

[121] R. Feldmann, C. M. Carollo, C. Porciani, S. J. Lilly, P. Capak, Y. Taniguchi, O. Le Fèvre, A. Renzini, N. Scoville, M. Ajiki, H. Aussel, T. Contini, H. McCracken, B. Mobasher, T. Murayama, D. Sanders, S. Sasaki, C. Scarlata, M. Scodeggio, Y. Shioya, J. Silverman, M. Takahashi, D. Thompson, and G. Zamorani, "The Zurich Extragalactic Bayesian Redshift Analyzer and its first application: COSMOS," *MNRAS*, vol. 372, pp. 565–577, Oct. 2006.

[122] L. van Waerbeke, M. White, H. Hoekstra, and C. Heymans, "Redshift and shear calibration: Impact on cosmic shear studies and survey design," *Astroparticle Physics*, vol. 26, pp. 91–101, Sept. 2006.

[123] I. Tereno, C. Schimd, J. Uzan, M. Kilbinger, F. H. Vincent, and L. Fu, "CFHTLS weak-lensing constraints on the neutrino masses," *A&A*, vol. 500, pp. 657–665, June 2009.

[124] H. Hoekstra, Y. Mellier, L. van Waerbeke, E. Semboloni, L. Fu, M. J. Hudson, L. C. Parker, I. Tereno, and K. Benabed, "First Cosmic Shear Results from the Canada-France-Hawaii Telescope Wide Synoptic Legacy Survey," *ApJ*, vol. 647, pp. 116–127, Aug. 2006.

[125] W. Hu, "Dark energy and matter evolution from lensing tomography," *Phys. Rev. D*, vol. 66, pp. 083515–+, Oct. 2002.

[126] W. Hu, "Power Spectrum Tomography with Weak Lensing," *ApJ Lett.*, vol. 522, pp. L21–L24, Sept. 1999.

[127] D. Huterer, "Weak lensing and dark energy," *Phys. Rev. D*, vol. 65, p. 063001, Mar. 2002.

[128] A. Heavens, "3D weak lensing," *MNRAS*, vol. 343, pp. 1327–1334, Aug. 2003.

[129] G. Bernstein and B. Jain, "Dark Energy Constraints from Weak-Lensing Cross-Correlation Cosmography," *ApJ*, vol. 600, pp. 17–25, Jan. 2004.

[130] M. Takada and B. Jain, "Cosmological parameters from lensing power spectrum and bispectrum tomography," *MNRAS*, vol. 348, pp. 897–915, Mar. 2004.

[131] B. Jain and A. Taylor, "Cross-Correlation Tomography: Measuring Dark Energy Evolution with Weak Lensing," *Physical Review Letters*, vol. 91, p. 141302, Oct. 2003.

[132] W. Hu and B. Jain, "Joint galaxy-lensing observables and the dark energy," *Phys. Rev. D*, vol. 70, p. 043009, Aug. 2004.

[133] A. F. Heavens, T. D. Kitching, and A. N. Taylor, "Measuring dark energy properties with 3D cosmic shear," *MNRAS*, vol. 373, pp. 105–120, Nov. 2006.

[134] A. N. Taylor, T. D. Kitching, D. J. Bacon, and A. F. Heavens, "Probing dark energy with the shear-ratio geometric test," *MNRAS*, vol. 374, pp. 1377–1403, Feb. 2007.

[135] T. Bayes and Price, "An Essay towards solving a Problem in the Doctrine of Chance. By the late Rev. Mr. Bayes, communicated by Mr. Price, in a letter to John Canton, M. A. and F. R. S.," *Philosophical Transactions of the Royal Society of London*, vol. 53, pp. 370–418, 1763.

[136] M. Hobson, A. Jaffe, A. Liddle, P. Mukherjee, and D. Parkinson, *Bayesian Methods in Cosmology*. Cambridge University Press, 2010.

[137] H. Jeffreys, *The Theory of Probability*. Oxford University Press, 1961.

[138] R. Kass and R. A., "Bayes Factors," *Journal of the American Statistical Association*, vol. 90, pp. 773–795, 1995.

[139] S. Devinderjit and J. Skilling, *Data Analysis: A Bayesian Tutorial*. Oxford University Press, 2006.

[140] H. Lancaster and S. E., *Chi-Square Distribution. Encyclopedia of Biostatistics*. John Wiley & Sons, 1969.

[141] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipes in Fortran*. Cambridge University Press, 1992.

[142] A. Gelman and D. Rubin, "Inference from Iterative Simulation Using Multiple Sequences," *Statistical Science*, vol. 7, pp. 457–472, 1992.

[143] A. Heavens, "Statistical techniques in cosmology," *ArXiv e-prints*, June 2009.

[144] C. Geyer, "Burn-in is unneccesary," 1993. `http://www.stat.umn.edu/~charlie/mcmc/burn.html`.

[145] C. Robert and C. G., *Monte Carlo statistical methods*. New York: Springer, 2004.

[146] J. Shlens, "A tutorial on principal component analysis," 2009. `http://www.snl.salk.edu/~shlens/pca.pdf`.

[147] A. Hajian, "Efficient cosmological parameter estimation with Hamiltonian MonteCarlo technique," *Phys. Rev. D*, vol. 75, pp. 083525–+, Apr. 2007.

[148] O. Cappé and E. Moulines, "Online EM Algorithm for Latent Data Models," *ArXiv e-prints*, Dec. 2007.

[149] D. Wraith, M. Kilbinger, K. Benabed, O. Cappé, J.-F. Cardoso, G. Fort, S. Prunet, and C. P. Robert, "Estimation of cosmological parameters using adaptive importance sampling," *Phys. Rev. D*, vol. 80, p. 023507, July 2009.

[150] M. Kilbinger, D. Wraith, C. P. Robert, K. Benabed, O. Cappé, J.-F. Cardoso, G. Fort, S. Prunet, and F. R. Bouchet, "Bayesian model comparison in cosmology with Population Monte Carlo," *MNRAS*, vol. 405, pp. 2381–2390, July 2010.

[151] B. Brewer, "Nested sampling," 2009. `http://www.physics.ucsb.edu/~brewer/nested.pdf`.

[152] P. Simon, L. J. King, and P. Schneider, "The covariance of cosmic shear correlation functions and cosmological parameter estimates using redshift information," *A&A*, vol. 417, pp. 873–885, Apr. 2004.

[153] D. J. Bacon, A. N. Taylor, M. L. Brown, M. E. Gray, C. Wolf, K. Meisenheimer, S. Dye, L. Wisotzki, A. Borch, and M. Kleinheinrich, "Evolution of the dark matter distribution with three-dimensional weak lensing," *MNRAS*, vol. 363, pp. 723–733, Nov. 2005.

[154] E. Semboloni, Y. Mellier, L. van Waerbeke, H. Hoekstra, I. Tereno, K. Benabed, S. D. J. Gwyn, L. Fu, M. J. Hudson, R. Maoli, and L. C. Parker, "Cosmic shear analysis with CFHTLS deep data," *A&A*, vol. 452, pp. 51–61, June 2006.

[155] R. Massey, J. Rhodes, A. Leauthaud, P. Capak, R. Ellis, A. Koekemoer, A. Réfrégier, N. Scoville, J. E. Taylor, J. Albert, J. Bergé, C. Heymans, D. Johnston, J.-P. Kneib, Y. Mellier, B. Mobasher, E. Semboloni, P. Shopbell, L. Tasca, and L. Van Waerbeke, "COSMOS: Three-dimensional Weak Lensing and the Growth of Structure," *ApJ Supp*, vol. 172, pp. 239–253, Sept. 2007.

[156] T. Schrabback, J. Hartlap, B. Joachimi, M. Kilbinger, P. Simon, K. Benabed, M. Bradač, T. Eifler, T. Erben, C. D. Fassnacht, F. W. High, S. Hilbert, H. Hildebrandt, H. Hoekstra, K. Kuijken, P. J. Marshall, Y. Mellier, E. Morganson, P. Schneider, E. Semboloni, L. van Waerbeke, and M. Velander, "Evidence of the accelerated expansion of the Universe from weak lensing tomography with COSMOS," *A&A*, vol. 516, p. A63, June 2010.

[157] S. Bridle and L. King, "Dark energy constraints from cosmic shear power spectra: impact of intrinsic alignments on photometric redshift requirements," *New Journal of Physics*, vol. 9, pp. 444–+, Dec. 2007.

[158] C. E. Cunha, D. Huterer, M. T. Busha, and R. H. Wechsler, "Sample variance in photometric redshift calibration: cosmological biases and survey requirements," *ArXiv e-prints*, Sept. 2011.

[159] H. Oyaizu, M. Lima, C. E. Cunha, H. Lin, and J. Frieman, "Photometric Redshift Error Estimators," *ApJ*, vol. 689, pp. 709–720, Dec. 2008.

[160] Z. Ma, W. Hu, and D. Huterer, "Effects of Photometric Redshift Uncertainties on Weak-Lensing Tomography," *ApJ*, vol. 636, pp. 21–29, Jan. 2006.

[161] L. Sun, Z.-H. Fan, C. Tao, J.-P. Kneib, S. Jouvel, and A. Tilquin, "Catastrophic Photo-z Errors and the Dark Energy Parameter Estimates with Cosmic Shear," *ApJ*, vol. 699, pp. 958–967, July 2009.

[162] E. Semboloni, H. Hoekstra, J. Schaye, M. P. van Daalen, and I. G. McCarthy, "Quantifying the effect of baryon physics on weak lensing tomography," *MNRAS*, vol. 417, pp. 2020–2035, Nov. 2011.

[163] J. Harnois-Deraps, S. Vafaei, and L. Van Waerbeke, "Gravitational Lensing Simulations I : Covariance Matrices and Halo Catalogues," *ArXiv e-prints*, Feb. 2012.

[164] M. Kilbinger and P. Schneider, "Analysis of two-point statistics of cosmic shear. II. Optimizing the survey geometry," *A&A*, vol. 413, pp. 465–476, Jan. 2004.

[165] N. Kaiser, "Weak gravitational lensing of distant galaxies," *ApJ*, vol. 388, pp. 272–286, Apr. 1992.

[166] P. Schneider, L. van Waerbeke, B. Jain, and G. Kruse, "A new measure for cosmic shear," *MNRAS*, vol. 296, pp. 873–892, June 1998.

[167] R. J. Barlow, *Statistics*. John Wiley & Sons, 1991.

[168] J. Pan and I. Szapudi, "The monopole moment of the three-point correlation function of the two-degree Field Galaxy Redshift Survey," *MNRAS*, vol. 362, pp. 1363–1370, Oct. 2005.

[169] J. Hartlap, P. Simon, and P. Schneider, "Why your model parameter confidences might be too optimistic. Unbiased estimation of the inverse covariance matrix," *A&A*, vol. 464, pp. 399–404, Mar. 2007.

[170] T. W. Anderson, *An introduction to multivariate statistical analysis, 3rd Edn.* Wiley-Interscience, 2003.

[171] M. Sato, T. Hamana, R. Takahashi, M. Takada, N. Yoshida, T. Matsubara, and N. Sugiyama, "Simulations of Wide-Field Weak Lensing Surveys. I. Basic Statistics and Non-Gaussian Effects," *ApJ*, vol. 701, pp. 945–954, Aug. 2009.

[172] M. White and W. Hu, "A New Algorithm for Computing Statistics of Weak Lensing by Large-Scale Structure," *ApJ*, vol. 537, pp. 1–11, July 2000.

[173] V. Springel, S. D. M. White, A. Jenkins, C. S. Frenk, N. Yoshida, L. Gao, J. Navarro, R. Thacker, D. Croton, J. Helly, J. A. Peacock, S. Cole, P. Thomas, H. Couchman, A. Evrard, J. Colberg, and F. Pearce, "Simulations of the formation, evolution and clustering of galaxies and quasars," *Nature*, vol. 435, pp. 629–636, June 2005.

[174] J. Kim, C. Park, G. Rossi, S. M. Lee, and J. R. Gott, III, "The New Horizon Run Cosmological N-Body Simulations," *Journal of Korean Astronomical Society*, vol. 44, pp. 217–234, Dec. 2011.

[175] P. Schneider, L. van Waerbeke, Y. Mellier, B. Jain, S. Seitz, and B. Fort, "Detection of shear due to weak lensing by large-scale structure," *A&A*, vol. 333, pp. 767–778, May 1998.

[176] M. L. Brown and R. A. Battye, "Polarization as an indicator of intrinsic alignment in radio weak lensing," *MNRAS*, vol. 410, pp. 2057–2074, Jan. 2011.

[177] H. Hildebrandt, T. Erben, K. Kuijken, L. van Waerbeke, C. Heymans, J. Coupon, J. Benjamin, C. Bonnett, L. Fu, H. Hoekstra, T. D. Kitching, Y. Mellier, L. Miller, M. Velander, M. J. Hudson, B. T. P. Rowe, T. Schrabback, E. Semboloni, and N. Benítez, "CFHTLenS: improving the quality of photometric redshifts with precision photometry," *MNRAS*, p. 2386, Feb. 2012.

[178] N. Benítez, "Bayesian Photometric Redshift Estimation," *ApJ*, vol. 536, pp. 571–583, June 2000.

[179] C. Power and A. Knebe, "The impact of box size on the properties of dark matter haloes in cosmological simulations," *MNRAS*, vol. 370, pp. 691–701, Aug. 2006.

[180] M. Sato, M. Takada, T. Hamana, and T. Matsubara, "Simulations of Wide-field Weak-lensing Surveys. II. Covariance Matrix of Real-space Correlation Functions," *ApJ*, vol. 734, p. 76, June 2011.

[181] S. Hilbert, J. Hartlap, and P. Schneider, "Cosmic shear covariance: the log-normal approximation," *A&A*, vol. 536, p. A85, Dec. 2011.

[182] T. Eifler, P. Schneider, and J. Hartlap, "Dependence of cosmic shear covariances on cosmology. Impact on parameter estimation," *A&A*, vol. 502, pp. 721–731, Aug. 2009.

[183] I. Smail, R. S. Ellis, and M. J. Fitchett, "Gravitational Lensing of Distant Field Galaxies by Rich Clusters - Part One - Faint Galaxy Redshift Distributions," *MNRAS*, vol. 270, p. 245, Sept. 1994.

[184] D. N. Spergel, R. Bean, O. Doré, M. R. Nolta, C. L. Bennett, J. Dunkley, G. Hinshaw, N. Jarosik, E. Komatsu, L. Page, H. V. Peiris, L. Verde, M. Halpern, R. S. Hill, A. Kogut, M. Limon, S. S. Meyer, N. Odegard, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright, "Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Implications for Cosmology," *ApJ Supp.*, vol. 170, pp. 377–408, June 2007.

[185] M. Takada and M. White, "Tomography of Lensing Cross-Power Spectra," *ApJ Lett.*, vol. 601, pp. L1–L4, Jan. 2004.

[186] R. Mandelbaum, C. M. Hirata, M. Ishak, U. Seljak, and J. Brinkmann, "Detection of large-scale intrinsic ellipticity-density correlation from the Sloan Digital Sky Survey and implications for weak lensing surveys," *MNRAS*, vol. 367, pp. 611–626, Apr. 2006.

[187] C. M. Hirata, R. Mandelbaum, M. Ishak, U. Seljak, R. Nichol, K. A. Pimbblet, N. P. Ross, and D. Wake, "Intrinsic galaxy alignments from the 2SLAQ and SDSS surveys: luminosity and redshift scalings and implications for weak lensing surveys," *MNRAS*, vol. 381, pp. 1197–1218, Nov. 2007.

[188] T. D. Kitching, A. F. Heavens, A. N. Taylor, M. L. Brown, K. Meisenheimer, C. Wolf, M. E. Gray, and D. J. Bacon, "Cosmological constraints from COMBO-17 using 3D weak lensing," *MNRAS*, vol. 376, pp. 771–778, Apr. 2007.

[189] T. D. Kitching, A. F. Heavens, and L. Miller, "3D photometric cosmic shear," *MNRAS*, vol. 413, pp. 2923–2934, June 2011.

[190] A. Heavens, A. Refregier, and C. Heymans, "Intrinsic correlation of galaxy shapes: implications for weak lensing measurements," *MNRAS*, vol. 319, pp. 649–656, Dec. 2000.

[191] C. Heymans and A. Heavens, "Reducing and constraining the intrinsic galaxy alignment contamination to weak lensing measurements," *ArXiv Astrophysics e-prints*, Oct. 2003.

[192] B. Joachimi and P. Schneider, "The removal of shear-ellipticity correlations from the cosmic shear signal. Influence of photometric redshift errors on the nulling technique," *A&A*, vol. 507, pp. 105–129, Nov. 2009.

[193] L. King and P. Schneider, "Suppressing the contribution of intrinsic galaxy alignments to the shear two-point correlation function," *A&A*, vol. 396, pp. 411–418, Dec. 2002.

[194] C. Heymans, M. Brown, A. Heavens, K. Meisenheimer, A. Taylor, and C. Wolf, "Weak lensing with COMBO-17: estimation and removal of intrinsic alignments," *MNRAS*, vol. 347, pp. 895–908, Jan. 2004.

[195] B. Joachimi and P. Schneider, "The removal of shear-ellipticity correlations from the cosmic shear signal via nulling techniques," *A&A*, vol. 488, pp. 829–843, Sept. 2008.

[196] P. Catelan, M. Kamionkowski, and R. D. Blandford, "Intrinsic and extrinsic galaxy alignment," *MNRAS*, vol. 320, pp. L7–L13, Jan. 2001.

[197] M. L. Brown, A. N. Taylor, N. C. Hambly, and S. Dye, "Measurement of intrinsic alignments in galaxy ellipticities," *MNRAS*, vol. 333, pp. 501–509, July 2002.

[198] L. J. King, "Cosmic shear as a tool for precision cosmology: minimising intrinsic galaxy alignment-lensing interference," *A&A*, vol. 441, pp. 47–53, Oct. 2005.

[199] B. Joachimi, R. Mandelbaum, F. B. Abdalla, and S. L. Bridle, "Constraints on intrinsic alignment contamination of weak lensing surveys using the MegaZ-LRG sample," *A&A*, vol. 527, pp. A26+, Mar. 2011.

[200] C. M. Hirata and U. Seljak, "Erratum: Intrinsic alignment-lensing interference as a contaminant of cosmic shear [Phys. Rev. D 70, 063526 (2004)]," *Phys. Rev. D*, vol. 82, p. 049901, Aug. 2010.

[201] R. Mandelbaum, C. Blake, S. Bridle, F. B. Abdalla, S. Brough, M. Colless, W. Couch, S. Croom, T. Davis, M. J. Drinkwater, K. Forster, K. Glazebrook,

B. Jelliffe, R. J. Jurek, I.-H. Li, B. Madore, C. Martin, K. Pimbblet, G. B. Poole, M. Pracy, R. Sharp, E. Wisnioski, D. Woods, and T. Wyder, "The WiggleZ Dark Energy Survey: direct constraints on blue galaxy intrinsic alignments at intermediate redshifts," *MNRAS*, vol. 410, pp. 844–859, Jan. 2011.

[202] R. Bock and W. Krisher, "The data analysis brief book," 1999. `http://rkb.home.cern.ch/rkb/titleA.html`.

[203] S. Vafaei, T. Lu, L. van Waerbeke, E. Semboloni, C. Heymans, and U.-L. Pen, "Breaking the degeneracy: Optimal use of three-point weak lensing statistics," *Astroparticle Physics*, vol. 32, pp. 340–351, Jan. 2010.

[204] "Condition number of a matrix," 1998. `http://teal.gmu.edu/ececourses/ece699/notes/note4.html`.

[205] P. Norberg, C. M. Baugh, E. Gaztañaga, and D. J. Croton, "Statistical analysis of galaxy surveys - I. Robust error estimation for two-point clustering statistics," *MNRAS*, vol. 396, pp. 19–38, June 2009.

[206] I. Tereno, O. Doré, L. van Waerbeke, and Y. Mellier, "Joint cosmological parameters forecast from CFHTLS-cosmic shear and CMB data," *A&A*, vol. 429, pp. 383–398, Jan. 2005.

[207] L. Van Waerbeke, Y. Mellier, R. Pelló, U.-L. Pen, H. J. McCracken, and B. Jain, "Likelihood analysis of cosmic shear on simulated and VIRMOS-DESCART data," *A&A*, vol. 393, pp. 369–379, Oct. 2002.

[208] M. Kilbinger et al., "CFHTLenS: Cosmological constraints from weak cosmological lensing (In Prep.)," 2012.

[209] J. Benjamin et al., "CFHTLenS: Tomographic weak lensing with broad redshift bins (In Prep.)," 2012.

[210] T. Erben, M. Schirmer, J. P. Dietrich, O. Cordes, L. Haberzettl, M. Hetterscheidt, H. Hildebrandt, O. Schmithuesen, P. Schneider, P. Simon, E. Deul, R. N. Hook, N. Kaiser, M. Radovich, C. Benoist, M. Nonino, L. F. Olsen, I. Prandoni, R. Wichmann, S. Zaggia, D. Bomans, R. J. Dettmar, and J. M. Miralles, "GaBoDS: The Garching-Bonn Deep Survey. IV. Methods for the image reduction of multi-chip cameras demonstrated on data from the ESO Wide-Field Imager," *Astronomische Nachrichten*, vol. 326, pp. 432–464, July 2005.

[211] T. Erben, H. Hildebrandt, M. Lerchster, P. Hudelot, J. Benjamin, L. van Waerbeke, T. Schrabback, F. Brimioulle, O. Cordes, J. P. Dietrich, K. Holhjem, M. Schirmer, and P. Schneider, "CARS: the CFHTLS-Archive-Research Survey. I. Five-band multi-colour data from 37 sq. deg. CFHTLS-wide observations," *A&A*, vol. 493, pp. 1197–1222, Jan. 2009.

[212] K. Kuijken, "Shears from shapelets," *A&A*, vol. 456, pp. 827–838, Sept. 2006.

[213] L. Miller et al., "Bayesian Galaxy Shape Measurement for Weak Lensing Surveys -III. Application to the Canada-France-Hawaii Telescope Lensing Survey. (In Prep).," 2012.

[214] H. Hildebrandt, S. Arnouts, P. Capak, L. A. Moustakas, C. Wolf, F. B. Abdalla, R. J. Assef, M. Banerji, N. Benítez, G. B. Brammer, T. Budavári, S. Carliles, D. Coe, T. Dahlen, R. Feldmann, D. Gerdes, B. Gillis, O. Ilbert, R. Kotulla, O. Lahav, I. H. Li, J.-M. Miralles, N. Purger, S. Schmidt, and J. Singal, "PHAT: PHoto-z Accuracy Testing," *A&A*, vol. 523, p. A31, Nov. 2010.

[215] E. Semboloni, T. Schrabback, L. van Waerbeke, S. Vafaei, J. Hartlap, and S. Hilbert, "Weak lensing from space: first cosmological constraints from three-point shear statistics," *MNRAS*, vol. 410, pp. 143–160, Jan. 2011.

[216] F. Schmidt, A. Leauthaud, R. Massey, J. Rhodes, M. R. George, A. M. Koekemoer, A. Finoguenov, and M. Tanaka, "A Detection of Weak-lensing Magnification Using Galaxy Sizes and Magnitudes," *Ap.J.Lett.*, vol. 744, p. L22, Jan. 2012.