WEAK GRAVITATIONAL LENSING UNCERTAINTIES

by

Tingting Lu

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of Astronomy and Astrophysics
University of Toronto

Abstract

Weak gravitational lensing uncertainties

Tingting Lu

Doctor of Philosophy

Graduate Department of Astronomy and Astrophysics

University of Toronto

2009

Dark matter dominates the mass distribution of the universe, and dark energy determines its expansion. The two are the most mysterious and attractive subjects in modern cosmology, because they provide an opportunity to discover new fundamental physics.

Cosmological weak gravitational lensing, which describes the deflection of photons by the gravitational force from large-scale structure in the universe, has been an active area of research in the past decade with many completed, ongoing, and upcoming surveys. Because weak lensing is sensitive to the growth of structure and expansion history of the universe, it is a great tool for improving our understanding of both dark matter and dark energy problems.

Cosmic structures have become non-linear by gravitational clustering. The non-linear structures are important to weak lensing, and cause non-Gaussianity in the lensing maps. In this thesis, I study the influence of non-linearity and non-Gaussianity on the uncertainty of lensing measurements. I develop a new method to robustly measure the covariance matrix of the lensing convergence power spectrum, from simulations. Because 21-cm intensity map may soon cover half sky at redshift 1-4, I build optimal estimators for reconstructing lensing from the 21-cm sources. I develop Gaussian optimal estimators which can be derived analytically, and non-Gaussian optimal estimators which can be constructed numerically from simulation data. I then run a large number of N-body simulations. For both lenses and 21-cm sources, I explore the statistical uncertainties in the simulation data.

We show that the non-Gaussianity nature of lensing decreases the dark energy figure of merit by a factor of 1.3 to 1.6 for a few future surveys. We also find that the non-Gaussianity nature of the 21-cm sources reduces the signal to noise ratio by several orders of magnitude. The reconstruction noise saturates at mildly non-linear scales, where the linear power spectrum of the source is $\Delta^2 \sim 0.2 - 0.5$. For 21-cm sources at $z \sim 2 - 4$, the lensing reconstruction is limited by cosmic variance at $\ell \lesssim 100$, which is in the linear regime of gravitational growth, and robustly predicted by theory. This allows promising constraints to various modified gravity and dark energy models.

Acknowledgements

This thesis is finally done after years of works and struggling. I need to thank my supervisor, Professor Ue-Li Pen, who has provided numerous instructions and discussions. I have been inspired much by his way of focusing on the physical picture of problems and thinking with intuitions, instead of jumping into the mathematical details. I appreciate the suggestions and comments on my research, from my PhD committee members, Professor Howard Yee, Dick Bond and Ray Carlberg, at many (extra-) committee meetings.

I'm extremely grateful to Tzu-Ching Chang. Thank you for always encouraging me when I feel frustrated and depressed. I will always remember you as a kind elder sister. Doing research is not easy, I would not be able to get the works done without the helps and discussions from Olivier Dore, Neal Dalal, Pengjie Zhang, Hy Trac, Pat MacDonald, Brice Menard, Hugh Merz, Renbin Yan, and many many others.

The graduate school would be much more boring without my fellow (previous) graduate students, though I can't name all of them, I would like to mention Brian Lee, Adam Muzzin, Tornado Li, Rodrigo Fernandez, Zhiqi Huang and Joachim Harnois-Deraps. I have also received a lot of help from Marc Goodman and Angie Ho at the graduate office in the department of Astronomy and Astrophysics.

At the end, I would like to thank my dearest friend Lei Zhao, and my parents, for your unselfish supports, both emotionally and financially.

This research is mainly supported by the Connaught Graduate Scholarships and University of Toronto Fellowships. All computations were performed on the Canada Foundation for Innovation funded CITA Sunnyvale cluster.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 The non-Gaussianity problem in cosmological lensing

Different measurements such as Cosmic Microwave Background (CMB) (Efstathiou et al., 1992; Komatsu et al., 2008), Type-I supernovae (Riess et al., 2004; Astier et al., 2006), galaxy clustering (Cole et al., 2005; Tegmark et al., 2006), and weak gravitational lensing (Hoekstra et al., 2006; Massey et al., 2007b), all suggest that our universe is made of dark energy (75%), cold dark matter (21%) and baryon (ordinary matter, 4%) components. This is the standard model of cosmology. Today, at the golden era of precision cosmology, the most important goal of cosmological research is to constrain the parameters for the cosmological models through measuring astrophysical observables in cosmological phenomena.

In the measurement of all these experiments and observations, one has to deal with noises generated by systematics and statistical uncertainties. How to conduct the measurement with maximum signal to noise ratio (S/N), and how to analyze the data without losing information is crucial. Technically speaking, one needs to construct unbiased estimators of the cosmological parameters with minimum noise. For a Gaussian system, such optimization can usually be found analytically. But in most cases, one needs to find the peak of the likelihood numerically, or to approximate it with an optimal quadratic estimator, which we will introduce in section 1.3.

In the scenario of standard cosmology, the primordial density fluctuations from inflation are Gaussian random perturbations. Eventually, the cosmic structures become non-linear because gravitational clustering plays an important role in the evolution. Consequently, non-Gaussianity arises. The statistics of non-Gaussian fields can not be calculated analytically.

Gravitational lensing is a very useful tool in cosmology, to map matter distribution and constrain the evolution of dark energy. This is because lensing only relies on gravity and is sensitive to the geometry of the universe. According to General Relativity, massive objects and their gravitational fields curve the space-time continuum. Since photons follow geodesic trajectories, they will be deflected in the presence of gravitational field. Therefore images from distant sources are distorted after propagating through or close to objects with large masses. Cosmological lensing is an integrated effect of the contribution of the gravitational potential along the line-of-sight. Therefore, non-Gaussianities in the density distribution will leave an imprint on lensing. The prediction of the lensing signal and its noise in observations both involve non-Gaussianity (Hu & White, 2001; Semboloni et al., 2007; Takada & Jain, 2008; Eifler et al., 2008). This thesis is aimed at addressing such problems with the help of large numbers of numerical simulations.

As I will explain in section 1.3, non-Gaussianity usually means larger uncertainty in measurement. We would like to know how much this enlargement is. A useful quantity to compare the uncertainties is the Fisher information content, which is the reciprocal of the variance for the parameter to be measured. Rimes & Hamilton (2005) found an interesting phenomenon in the Fisher information content in the primordial amplitude of the dark matter power spectrum: At trans-linear scales, the cumulative Fisher information stops to increase as the resolution of structures is increased. They called such

Fisher information “saturated”. This seems to be counter intuitive, however halo models can help to give an explanation (Neyrinck et al., 2006). Because of the strong tie between dark matter and gravitational lensing, we will investigate the same phenomena with the lensing convergence power spectrum. The influence of the non-Gaussianity on the prediction of weak lensing constraints on the dark energy figure of merit will also be discussed. This will be the main subject in chapter 2.

Current lensing measurements are mostly from the shape distortion of galaxies, which is called cosmic shear. Another possible lensing source is the 21-cm hyperfine line transition of neutral hydrogen at the epoch of reionization ($z \sim 7-20$), where the neutral universe after recombination starts to be ionized by the formation of the first generation of stars. The detection of the 21-cm emission at the epoch of reionization has recently been a hot topic in astronomical research. With possible mapping of 21-cm emission, scientists started to consider it as a new source for lensing (Pen, 2004; Cooray, 2004; Zahn & Zaldarriaga, 2006; Mandel & Zaldarriaga, 2006; Metcalf & White, 2007; Hilbert et al., 2007; Metcalf & White, 2008). Though the exact mechanism of reionization and evolution is still a developing story, we know that the 21-cm sources, the neutral hydrogen region, are more or less tracing the total matter, therefore they are non-Gaussian distributions. As we will show in chapter 3, the non-Gaussianity of the 21-cm sources will be a dominant effect in the lensing reconstruction. Previous treatment of the 21-cm lensing estimator all assumed the sources follow a Gaussian distribution. We also derive “optimal” quadratic estimators of lensing under the Gaussian assumption. However, we generate non-Gaussian sources from cosmological simulations and discuss the influence of non-Gaussianity on the noise of the reconstructed lensing map using these estimators. We report a saturation effect similar to that of the Fisher information content in dark matter power spectrum.

A a few experiments have been proposed to map post-reionization ($z \sim 1-6$) 21-cm gas with large sky coverage (Peterson et al., 2006; Chang et al., 2008; Wyithe et al., 2008;

Tegmark & Zaldarriaga, 2008). In particular, Cosmological Hydrogen Intensity Mapping initiative Experiment (CHIME) [1] (Chang et al., 2008) proposed to map the intensity of 21-cm sources at $z \sim 1-3$ for half of the sky in a two year operation. By measuring lensing from these 21-cm sources, cosmic structures at $z \sim 0.5-1.5$ can be investigated. Although the dark energy in the standard model of cosmology has successfully explained the observed accelerating expansion of the universe, the nature of dark energy is not clear. We cannot distinguish the dark energy model with a smooth stress-energy component from alternative models such as modified gravity (Schmidt, 2008). Cosmic structures at $z \sim 1$ are sensitive to dark energy in most of these models. In chapter 4, we will extend the lensing reconstruction calculation to 21-cm sources at $z = 1-6$. Because the non-linearity at these redshifts are a more serious issue, the optimal estimator derived from a Gaussian assumption is significantly non-optimal. We develop a real optimal estimator for the non-Gaussian sources, constructed numerically from a large number of N-body simulations. We find that the saturation effect appears at all redshifts, and that the saturation scale depends on the shape of the power spectrum. Although the non-Gaussianity of 21-cm sources increases the noise, the noise from sources at $z \sim 3$ is about a few percent of the lensing power spectrum at $\ell \sim 100$, which can be a powerful tool to distinguish alternative and the smooth dark energy models at $z \sim 1$.

We will first review the cosmology background necessary for all the works within the thesis in section 1.2, and introduce the statistical and analysis methods in cosmology in section 1.3. In section 1.4 we will review the theory of gravitational lensing and state the problems we attempt to solve in the thesis. At last, we will talk about cosmological simulations and the simulation codes we used in 1.5.

---

[1] http://pirsa.org/08100036

1.2 Cosmology background

The thesis focuses on the discussion of weak lensing uncertainties in the frame of standard cosmology. It is necessary to understand the basis of cosmological models. First I will briefly introduce the cosmological background of the thesis, following the introduction and notation in Bartelmann & Schneider (2001). More comprehensive introductions to cosmology can be found in Peacock (1999) and Dodelson (2003).

1.2.1 The standard cosmology

The standard cosmology is sometimes called the big bang theory, which states that the universe is created in a big bang about 13.7 billion years ago. Our current universe is the relic of the big bang.

A very important fact in cosmology is that our universe is expanding, which was first discovered by Hubble in 1929. Imagine the cosmic structure is scaled by factor $a(t)$ at a cosmic time t. If we define $a(t_0) = 1$, where t_0 is the cosmic time at present, $a(t)$ increases from a value close to zero to unity when t increases from zero to t_0. Hereafter, we will use $[...]_0$ to represent a quantity at the present cosmic time. The speed of the expansion is defined as the Hubble parameter $H \equiv \dot{a}/a$, where $\dot{a}$ means the derivative of a with respect to t. Because all parts of the universe expand with the same speed at a certain cosmic time, H is also called Hubble constant. $H_0 = 100h\,\mathrm{km/s\,Mpc^{-1}}$ is the the Hubble constant at present. The recent five-year Wilkinson Microwave Anisotropy Probe (WMAP5) experiment results provide $h = 0.719 \pm 0.026$ (Dunkley et al., 2008).

In addition to the cosmological scale factor a, another important concept defined in cosmology is the cosmological redshift z. It is the cosmological application of the Doppler effect which states the frequency of light or sound will be shifted for a source moving away or toward the observer.

$$1 + z = \frac{\nu_{\mathrm{emit}}}{\nu_{\mathrm{obs}}} = \frac{a(t_0)}{a(t)}\,, \tag{1.1}$$

where $\nu_{\rm emit}$ and $\nu_{\rm obs}$ are frequency of light emitted at the cosmological distance denoted by z and the frequency observed at present. Higher redshift represents the universe at an earlier epoch, or farther away from us.

There are mainly four components in the universe: matter, radiation, vacuum, and curvature, whose densities are normally described by dimensionless density parameters: $\Omega_{\rm m}$, $\Omega_{\rm r}$, Ω_Λ, Ω_K respectively. $\Omega_K \equiv 1 - \Omega_{\rm m} - \Omega_\Lambda - \Omega_{\rm r}$, where $\Omega_{\rm r}$ is sometimes neglected because it is much smaller than others except at the very beginning of the expansion. The dimensionless density parameters are defined as $\Omega = \rho/\rho_c$, where $\rho_c = 3H^2/8\pi G^2$ is the critical density of the universe, and G is the gravitational constant. The critical density at present is $\rho_c = 1.88h^2 \times 10^{-29}{\rm g\,cm^{-3}}$. The density evolutions of these four components follow: $\rho_{\rm m} \propto a^{-3}, \rho_{\rm r} \propto a^{-4}, \rho_K \propto a^{-2}$, and ρ_Λ is constant. In this thesis, we will use $\Omega_{\rm r}, \Omega_{\rm m}$ and Ω_Λ to represent the density parameters at present. The evolution of the Hubble constant is given by :

$$H(z)^2 = H_0^2[\Omega_{\rm r}a^{-4} + \Omega_{\rm m}a^{-3} + (1 - \Omega_{\rm m} - \Omega_{\rm r} - \Omega_\Lambda)a^{-2} + \Omega_\Lambda]\,. \tag{1.2}$$

Einstein draw the physical picture of the universe and presented the famous cosmological principle about half a century ago: over large spatial volumes, the universe is homogeneous and isotropic. This is the foundation of all current studies in the cosmological field. Cosmologists solve physics problems at very large scales, where the language of General Relativity needs to be used. Mathematically, the space-time of our universe is described by the Friedmann-Robertson-Walker (FRW) metric:

$$ds^2 = -c^2dt^2 + a^2(t)(d\chi^2 + f_K(\chi)^2(d\theta^2 + \sin^2\theta d\varphi^2))\,, \tag{1.3}$$

where ds is the Lorentzian line element, c is the speed of light, χ is the radial coordinate, θ is the azimuthal angle, and φ is the polar angle. The comoving angular diameter distance $f_K(\chi)$ can be expressed as:

$$f_K(\chi) = \sin_K \chi \equiv \begin{cases} K^{-1/2} \sin K^{1/2}\chi, \ K > 0 \\ \chi, \ K = 0 \\ (-K)^{-1/2} \sinh(-K)^{1/2}\chi, \ K < 0 \, , \end{cases} \tag{1.4}$$

where $K = -\Omega_K H_0^2$ is the curvature term. $K = 0$ indicates that the geometry of the universe is flat. $K < 0$ stands for an open universe and $K > 0$ shows a closed one. Current observations imply our universe is a flat one.

To understand the dynamics of the universe, we need to solve the evolution of the scale factor $a(t)$. This is described by the famous Einstein's field equation of General Relativity, which reveals the underlying relationship between the space-time geometry and mass distribution.

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} \mathcal{R} = 8\pi G T_{\mu\nu} \, , \tag{1.5}$$

where $R_{\mu\nu}$ and R are the Ricci tensor and Ricci scalar, $G_{\mu\nu}$ is the Einstein tensor, and $T_{\mu\nu}$ is the energy-momentum tensor. Both the Ricci tensor and Ricci scalar can be calculated from the metric $g_{\mu\nu}$.

Assume the matter in the universe is a fluid with density $\rho(t)$ and pressure $p(t)$, by inserting the FRW metric into the Einstein's field equation, we obtain

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3\frac{p}{c^2}) + \frac{\Lambda}{3} \, , \tag{1.6}$$

and

$$(\frac{\dot{a}}{a})^2 = \frac{8\pi G}{3}\rho - \frac{K}{a^2} + \frac{\Lambda}{3} \, , \tag{1.7}$$

where $\Lambda = 3H_0^2\Omega_\Lambda$. Eq. (1.7) is called the Friedman equation, which is one of the most important equations in the standard model of cosmology.

1.2.2 Large-scale structures in the universe

The main ingredients of matter in the universe are: photons, dark matter, baryons, neutrinos. They are coupled to each other in a complicated way: dark matter does not interact with other species directly, but dominates the contribution to the gravitational potential, which will influence photons and baryons; photons are affected by gravity and by Compton scattering with free electrons; there is Coulomb scattering between the electrons and protons components of baryons; baryons and neutrinos both contribute to the gravitational potential and are also affected by it. Another component of universe is dark energy (or vacuum energy), which contributes to about two thirds of the total density of the universe, enough to make its geometry flat. Dark energy influences others by a repulsive force but does not perturb.

It is often assumed that the universe is weakly perturbed from the homogeneous and isotropic FRW space-time as shown in Eq. (1.3), therefore first-order perturbation theory can be used to find the approximate solution to the Einstein's equation. Density fluctuations are an example of scalar perturbations, but there also exist vector perturbations and tensor perturbations in some models of structure formation, the latter being caused by primordial gravitational waves. Any perturbation can be decomposed into these three independent classes, each of which evolves independently. In the context of this thesis, we are mainly interested in the density perturbations.

To describe the behavior for an ensemble of particles, Boltzmann's equation can be applied (Peacock, 1999)

$$(p^\mu \frac{\partial}{\partial x^\mu} - \Gamma^\mu_{\alpha\beta} p^\alpha p^\beta \frac{\partial}{\partial p^\mu}) f = C\,, \tag{1.8}$$

where $f(\mathbf{x}, \mathbf{p})$ is the phase space distribution function of the ensemble, x^μ and p^μ are position and momentum respectively, $\Gamma^\mu_{\alpha\beta}$ is the Christoffel symbol, and C is the collision term which includes the interactions between all matter components. The perturbations of the matter and radiation can thus be calculated by solving a set of second-order partial differential equations, often done numerically by packages such as CMBFAST[2] and CAMB[3].

---

[2] http://www.cfa.harvard.edu/ mzaldarr/CMBFAST/cmbfast.html

[3] http://camb.info/

Because we need to study the behavior of ensembles of particles, it is necessary to use the tools and the language of statistics. We will introduce the relevant statistics notations in section 1.3. Here I will simply mention the most important concept, power spectrum $P(k)$, which is the variance of the Fourier modes from an over-density field. Higher power spectrum $P(k)$ means more structures at certain scale k. In cosmology, we normally use the fluctuation power per logarithmic interval to denote a dimensionless power spectrum, i.e.,

$$\Delta^2(k) \equiv \frac{k^3}{2\pi^2} P(k) \,. \tag{1.9}$$

If $\Delta^2 \ll 1$, the fluctuation is in the linear regime, while when Δ^2 is comparable to or larger than 1, the fluctuation is said to be “non-linear”. A very useful quantity in observational cosmology is the normalization factor $\sigma_8 = \sigma(R = 8h^{-1}\mathrm{Mpc})$, which is the variance of mass at a sphere of radius $8h^{-1}$Mpc with a top hat smoothing window W,

$$\sigma^2(R) = \langle \delta_R^2 \rangle = \int \mathrm{d}^3k P(k) W^2(kR) \,, \tag{1.10}$$

where δ_R means the overdensity fluctuation of cosmic structures in the sphere, $\delta = (\rho - \bar{\rho})/\bar{\rho}$, and $\bar{\rho}$ is the mean density of the universe. Higher σ_8 means more clustering.

From inflation theory, we know that the primordial power spectrum of matter is a scale-invariant spectrum (also called the Harrison-Zeldovich-Peebles spectrum). The actual linear power spectrum is modified by a transfer function $\mathcal{T}(k)$.

$$P(k, z) = P_p(k) G^2(z) \mathcal{T}^2(k) \,. \tag{1.11}$$

$P_p(k)$ is the power spectrum of the primordial density fluctuation and $G(z)$ is the growth factor (Carroll et al., 1992; Van Waerbeke & Mellier, 2003):

$$G(z) = \frac{5}{2} \Omega_\mathrm{m} H(z) \int_0^a \frac{\mathrm{d}a'}{a'^3 H^3(a')} \,. \tag{1.12}$$

The transfer function can be calculated by CAMB or CMBFAST, or fitting formulas such as Bardeen et al. (1986); Eisenstein & Hu (1998, 1999).

At later stages of the evolution, cosmic structures become non-linear on small scales. The dynamics of the structure formation becomes complicated, and the calculation of power spectrum can not be treated in a fully analytical way. Thus numerical simulations are needed. From semi-analytical approaches, with the aid of N-body simulations, fitting formulas of non-linear power spectrum were found (Peacock & Dodds, 1996; Smith et al., 2003). However, these fitting formulas can only provide the expectation value of the non-linear power spectrum. In this thesis, we use N-body simulations as a major tool to calculate the higher order statistics such as the covariance matrix of power spectrum.

1.3 Primer of cosmological analysis

In this section, I will give a brief introduction to the analysis used in cosmological measurements and calculations. I will follow the definitions and discussions in Dodelson (2003), which has a very good introduction chapter to the analysis.

1.3.1 Statistical description

Gaussian random field

Gaussian random field is a very important concept in statistical analysis. Given a random field $\delta(\boldsymbol{x})$, the Fourier transform of the field is

$$\delta(\boldsymbol{k}) = \int \mathrm{d}^3 x \delta(\boldsymbol{x}) e^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} \,. \tag{1.13}$$

If all the phases of different Fourier modes are random and uncorrelated with each other, the field is a Gaussian random field (Peacock, 1999). For a Gaussian random field, all the statistical properties can be described by two-point statistics, i.e., correlation function

$$\langle \delta(\boldsymbol{x})\delta(\boldsymbol{x}')\rangle = \xi(|\boldsymbol{x}-\boldsymbol{x}'|) \,, \tag{1.14}$$

and power spectrum

$$\langle \delta(\boldsymbol{k})\delta(\boldsymbol{k}')\rangle = (2\pi)^3 \delta^{3\mathrm{D}}(\boldsymbol{k}+\boldsymbol{k}')P(\boldsymbol{k}) \,, \tag{1.15}$$

where $\delta^{3\mathrm{D}}$ is the 3-D Dirac delta function (we will use δ^{D} and $\delta^{2\mathrm{D}}$ to represent the 1-D and 2-D Dirac delta function). The power spectrum P and correlation function ξ can be related by Fourier transforms. Therefore, all required features of a Gaussian random field is described entirely by the power spectrum.

All the statistics of a Gaussian random field can be represented by the two-point function because the density probability distribution of the field is a joint Gaussian distribution (Peacock, 1999).

$$P[\delta_1, \delta_2, ..., \delta_m] = \frac{|\mathbf{C}^{-1}|^{1/2}}{(2\pi)^{m/2}} exp(-\frac{1}{2}\mathbf{V}^T\mathbf{C}^{-1}\mathbf{V})\,, \tag{1.16}$$

where $\mathbf{V}^T = (\delta_1, \delta_2, ..., \delta_m)$ is a vector composed of density at m different positions. $\mathbf{C}$ is the covariance matrix between all m points, i.e., $\mathbf{C}_{i,j} = \langle\delta_i\delta_j\rangle$. All orders of the correlation functions can be calculated by integrating with this Gaussian functional. In calculations, Wick's theorem is often applied (Bernardeau et al., 2002):

$$\langle\delta_1...\delta_{2p+1}\rangle = 0\,, \tag{1.17}$$

$$\langle\delta_1...\delta_{2p}\rangle = \sum \prod_{\text{p pairs (i,j)}} \langle\delta_i\delta_j\rangle\,, \tag{1.18}$$

where p is an integer.

For non-Gaussian fields, higher order statistics cannot be fully expanded into two-point functions, and some connected terms will remain. These remainders usually cannot be calculated analytically.

1.3.2 Analysis

Likelihood function

The probability of obtaining some data given a theory is called the likelihood. For a theory with parameter λ and its noise σ_λ, the probability of measuring data y in the experiment is given as (Dodelson, 2003)

$$P[y|\lambda, \sigma_\lambda] \equiv \mathcal{L}(y; \lambda, \sigma_\lambda) = \frac{1}{\sqrt{2\pi\sigma_\lambda^2}} exp\frac{-(y-\lambda)^2}{2\sigma_\lambda^2}\,. \tag{1.19}$$

At each measurement, y will be the contribution from both signal λ and noise, which has a Gaussian distribution with variance σ_λ. If there is a set of data $\mathbf{y} = y_i$, the likelihood becomes a multivariate Gaussian distribution, and the term inside the exponential will be $-\mathbf{y}^t\mathbf{C}^{-1}\mathbf{y}/2$. Here $\mathbf{C} = \langle \mathbf{y}^t\mathbf{y}\rangle - \langle \mathbf{y}^t\rangle\langle \mathbf{y}\rangle$. y could also be a complicated function of λ.

When making measurements, we need to estimate the parameters from the data set in an optimal way. According to the Bayesian theorem, the posterior likelihood is

$$P[\lambda, \sigma_\lambda|\mathbf{y}] = \frac{P[\mathbf{y}|\lambda, \sigma_\lambda]P[\lambda, \sigma_\lambda]}{P[\mathbf{y}]}\,. \tag{1.20}$$

where $P[\lambda, \sigma_\lambda]$ is the prior probability. Any knowledge of the prior distribution should be used if possible. However, in the case where we have no information about the prior, a flat distribution is normally chosen. In most cases, it is reasonable to assume that there is no preferred particular value for the parameter of a model. It turns out that even if the actual prior is far from flat, the likelihood function will be steeply peaked; therefore, the choice of the prior should not greatly affect the results (Dodelson, 2003).

The best estimator of parameter λ, given the data set $\mathbf{y}$, should maximize the likelihood, i.e.,

$$\frac{\partial \mathcal{L}}{\partial \lambda} = 0\,. \tag{1.21}$$

This equation is usually non-linear, and needs to be solved numerically. If there are multiple parameters λ_α, solutions to a set of equations must be found. In practice, the derivative of $ln\mathcal{L}$ is normally calculated because the algebra is simpler.

Fisher information matrix and optimal quadratic estimator

Finding the root of several non-linear equations is usually slow, because the sampling space has many dimensions. To accelerate the progress in finding the optimal estimator, we can apply the Taylor expansion to the likelihood function near its peak, (Seljak, 1998b; Dodelson, 2003).

$$ln\mathcal{L}(\lambda + \delta\lambda) = ln\mathcal{L}(\lambda) + \frac{\partial ln\mathcal{L}(\lambda)}{\partial \lambda}\delta\lambda + \frac{1}{2}\frac{\partial^2 ln\mathcal{L}(\lambda)}{\partial\lambda\partial\lambda}(\delta\lambda)^2\,. \tag{1.22}$$

The root of Eq. (1.21) can be found iteratively with the Newton-Raphson method. For multiple parameters λ_α, the expansion is performed for all the parameters up to the second order. If we take the initial estimate of the parameters to be zero, the roots become

$$\hat{\lambda}_\alpha = \frac{1}{2}\mathbf{F}^{-1}_{\alpha\beta}(\mathbf{y}^T\mathbf{C}^{-1}\frac{\partial\mathbf{C}}{\partial\lambda_\beta}\mathbf{C}^{-1}\mathbf{y} - Tr[\mathbf{C}^{-1}\frac{\partial\mathbf{C}}{\partial\lambda_\beta}])\,, \tag{1.23}$$

which is called the optimal quadratic estimator (or the minimum variance estimator). The maximum likelihood solution can be approximated by this quadratic estimator. The estimator is unbiased, i.e., $\langle\hat{\lambda}_\alpha\rangle = \lambda_\alpha$.

$\mathbf{F}_{\alpha\beta}$ represents the Fisher information matrix, which is defined as the expectation of the curvature for the likelihood function,

$$\begin{aligned}\mathbf{F}_{\alpha\beta} &\equiv \langle -\frac{\partial^2(ln\mathcal{L})}{\partial\lambda_\alpha\partial\lambda_\beta}\rangle \\ &= \frac{1}{2}Tr[\frac{\partial\mathbf{C}}{\partial\lambda_\alpha}\mathbf{C}^{-1}\frac{\partial\mathbf{C}}{\partial\lambda_\beta}\mathbf{C}^{-1}]\,.\end{aligned} \tag{1.24}$$

The error bar of the optimal quadratic estimator for each parameter is $\sqrt{(\mathbf{F}^{-1})_{\alpha\alpha}}$, which is the minimum error bar for any parameter estimation method according to the Cramer-Rao inequality.

The optimal quadratic estimator (or minimum variance estimator) is equivalent to the maximum likelihood estimator for a Gaussian likelihood distribution. When the distribution is non-Gaussian, the optimal quadratic estimator is still an unbiased estimator though the variance is no longer minimized. The optimal quadratic estimator will, however, still be a very good substitute for the estimator with the minimum noise (Seljak, 1998a,b). The expected variance of parameters will also be larger than that calculated using the inverse of Fisher matrix. The latter still gives the lowest value of all possible noises.

Fisher information content of dark matter power spectrum

Part of my thesis work was motivated by recent investigations into the information content of the 3-D matter power spectrum as quantified by the projection of the Fisher information on the amplitude of power spectrum (Rimes & Hamilton, 2005, 2006; Hamilton et al., 2006; Neyrinck et al., 2006).

The results obtained in these papers are somewhat surprising. While a Gaussian-like behavior was observed on linear (large) scales, the non-linear growth of structures entails an information saturation at partially non-linear scales. A quasi-Gaussian behavior was recovered once fully in the non-linear regime but at a substantially lower level. Thus, when devising a survey, optimizing the survey to gain sensitivity and resolution in the trans-linear regime (where the power spectrum is currently interpreted cosmologically) would not entail much pay-off if one was only focusing on this statistic, i.e., the Fisher information of the amplitude of the dark matter power spectrum. Despite the fact that the effects of the non-linear growth of structures had been widely studied before these works, formulating it in this way led to this surprising answer, heuristically well understood within the context of the halo model (Neyrinck et al., 2006). Neyrinck & Szapudi (2007) later showed that if we project the Fisher information into other parameters, analogous behaviors are then observed. This validates the insightful value of the amplitude projection.

As we will introduce in section 1.4, the weak lensing effect is caused by dark matter distribution along the line of sight of light rays from distant sources. The non-Gaussianity in the dark matter distribution will thus affect the weak lensing results. We will discuss how much this non-Gaussianity will affect the Fisher information content of lensing power spectrum, and further consider the projected effects on the cosmological constraints.

1.4 Weak gravitational lensing by the large-scale structure

Gravitational lensing happens when the trajectory of photons is deflected by a nearby gravitational field. Here we mainly discuss the gravitational lensing effect caused by the large-scale structure. Lensing can also be caused by local mass distributions such as clusters with large masses and galaxies, but these are not in the scope of this thesis. On cosmological scales, the deflection effect is small, so that the convergence κ, which I will introduce in the following section, is much smaller than unity. In this case, the type of lensing is called weak lensing.

Unlike other astronomical observations, such as the statistical distribution of galaxies or cosmic velocity fields, weak gravitational lensing measures the gravitational potential directly, and thus is not sensitive to the light-mass biasing or dynamics of a complicated system.

The weak lensing signal is mainly determined by the source distribution, the matter density Ω_m, the slope of the power spectrum Γ, and the normalization σ_8. As complementary measurements to the CMB, weak lensing observations could help to break the degeneracy of cosmological parameters.

1.4.1 Formalism of cosmological lensing

Here we follow the notation in Schneider et al. (1998). We only consider the scalar perturbations in the perturbed Robertson-Walker metric. In the conformal Newtonian gauge,

$$ds^2 = -(1+\frac{2\Phi}{c^2})c^2dt^2 + a^2(t)(1-\frac{2\Phi}{c^2})(d\chi^2 + f_K(\chi)^2(d\theta^2+\sin^2\theta d\varphi^2))\,, \tag{1.25}$$

where Φ is the Newtonian gravitational potential.

If all light rays from a source are deflected by the same amount, the image of the object cannot be distinguished from the original. Therefore, the absolute deflection of light rays has no physical meaning. We will therefore consider the deflection of rays relative to a fiducial ray. For a bundle of rays that enclose a fiducial ray in the center and intersect at the observer, we can denote each rays by its initial angular separation $\boldsymbol{\theta}$ from the fiducial ray. The comoving transverse separation of one ray will be $\boldsymbol{x}(\boldsymbol{\theta}, \chi)$ when the photons propagate along the fiducial ray for a comoving distance χ. The geodesic equation is

$$\frac{d^2\boldsymbol{x}}{d\chi^2} + K\boldsymbol{x} = -\frac{2}{c^2}\left[\nabla_\perp \Phi\left(\boldsymbol{x}\left(\boldsymbol{\theta}, \chi\right), \chi\right) - \nabla_\perp \Phi^0(\chi)\right] , \qquad (1.26)$$

where $\Phi(\boldsymbol{x}(\boldsymbol{\theta}, \chi), \chi)$ is the potential on a deflected ray, while $\Phi^0(\chi)$ is the potential on the fiducial ray, and $\boldsymbol{\nabla}_\perp$ is the gradient component on the transverse plane. Note the factor of two difference between Newtonian gravity and Einstein's theory in the weak field approximation ($\Phi \ll c^2$) for massless particles. Based on this simple relation, we can straightforwardly calculate the trajectory of photons. The integrated deflection can be calculated using the following formula:

$$\boldsymbol{x}(\boldsymbol{\theta}, \chi) = f_K(\chi)\boldsymbol{\theta} - \frac{2}{c^2}\int_0^\chi d\chi' f_K(\chi - \chi')\left[\boldsymbol{\nabla}_\perp \Phi\left(\boldsymbol{x}\left(\boldsymbol{\theta}, \chi'\right), \chi'\right) - \boldsymbol{\nabla}_\perp \Phi^{(0)}(\chi')\right] . \qquad (1.27)$$

The Jacobian matrix, which describes the mapping between the source plane and the image plane, is given by

$$\mathbf{J}_{ij}(\boldsymbol{\theta}, \chi) = \delta_{ij} - \frac{2}{c^2}\int_0^\chi d\chi' \frac{f_K(\chi - \chi') f_K(\chi')}{f_K(\chi)} \Phi_{,ik}\left(\boldsymbol{x}\left(\boldsymbol{\theta}, \chi'\right), \chi'\right) \mathbf{J}_{kj}(\boldsymbol{\theta}, \chi') . \qquad (1.28)$$

Here i and j refer to the two perpendicular components on the transverse plan, and $\Phi_{,ik}$ is the second-order derivative of the potential with respect to the transverse coordinates. Note that the integral is along the actual perturbed path of each photon. Since it is a weak gravitational field, $\mathbf{J}$ can be expanded into powers of the potential Φ. Normally only the lowest order is considered:

$$\mathbf{J}_{ij}(\boldsymbol{\theta}, \chi) = \delta_{ij} - \frac{2}{c^2}\int_0^\chi d\chi' \frac{f_K(\chi - \chi') f_K(\chi')}{f_K(\chi)} \Phi_{,ij}\left(f_K(\chi')\boldsymbol{\theta}, \chi'\right) . \qquad (1.29)$$

Furthermore, the deflection is integrated along the unperturbed path, and this is called the Born approximation. Those two approximations can greatly simplify the calculation.

The physical quantity frequently used to describe a lensing field is the convergence κ and the shear γ, which are defined by

$$\mathbf{J}(\boldsymbol{\theta},\chi) = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix} .$$

κ describes the isotropic magnification, while γ_1 and γ_2 describe the anisotropic distortion. The convergence and shear can be defined as $\kappa = (\Phi_{,11}+\Phi_{,22})/2$, $\gamma_1 = (\Phi_{,11}-\Phi_{,22})/2$ and $\gamma_2 = \Phi_{,12}$.

Limber's equation

In the small angle approximation (Limber, 1954), $\nabla_\perp^2$ can be replaced by ∇^2 in the integral. It is a very important simplification, when combined with the Poisson equation,

$$\nabla^2\Phi = \frac{3H_0^2}{2}\,\Omega_\mathrm{m}\,\frac{\delta}{a}\,. \tag{1.30}$$

This leads to Limber's equation:

$$\kappa(\boldsymbol{\theta}) = \frac{3}{2}\left(\frac{H_0}{c}\right)^2 \Omega_\mathrm{m}\int_0^{\chi_H} g(\chi)\,\frac{\delta(f_K(\chi)\boldsymbol{\theta},\chi)}{a(\chi)}\,d\chi, \tag{1.31}$$

where

$$g(\chi) = \int_\chi^{\chi_H} d\chi n_\chi(\chi')\frac{f_K(\chi'-\chi)f_K(\chi)}{f_K(\chi')}, \tag{1.32}$$

$n_\chi(\chi)$ is the distribution function of sources, and χ_H is the comoving distance to the horizon. The convergence of the mapping is then related to the matter distribution in the universe: it is the line of sight integrals from the source to redshift zero. Therefore, we can probe the large-scale structures through measurements of the weak lensing effect.

Kaiser (1996) derived the Fourier-space version of Limber's equation:

$$\frac{\ell^2 C_\ell^\kappa}{2\pi} = \frac{9}{4}\left(\frac{H_0}{c}\right)^4 \Omega_\mathrm{m}^2\int_0^{\chi_\mathrm{H}} d\chi\frac{g^2(\chi)}{a^2(\chi)}P\left(\frac{\ell}{f_K(\chi)},\chi\right)\,, \tag{1.33}$$

where $\ell^2 C_\ell^\kappa/2\pi$ is the two-dimensional power spectrum of the κ field, $\langle\kappa(\boldsymbol{\ell})\kappa^\star(\boldsymbol{\ell}')\rangle = (2\pi)^2\delta^{2\mathrm{D}}(\boldsymbol{\ell}-\boldsymbol{\ell}')C_\ell^\kappa$. The equation is valid when the power spectrum P evolves slowly over time corresponding to the scales of fluctuation of interest, and these fluctuation scales are smaller than the curvature scale. It also requires the small angle approximation. Limber's equation can also be written as the integral over all redshifts, as shown in Eq. (2.9). For convenience, in this thesis, comoving distances are used instead of χ:

$$D(z) = f_K(\chi(z))\,. \tag{1.34}$$

A source with size L at redshift z will correspond to an angle $L/a(z)D(z)$.

E-B mode decomposition

Because shear is a spin-2 field, similar to the polarization of light, γ_1 and γ_2 are equivalent to Stokes quantities Q and U. We can define an E-mode component (analogous to an electric field) and a B-mode component (analogous to a magnetic field) from these quantities:

$$E(\ell) = \gamma_1\cos 2\phi_\ell + \gamma_2\sin 2\phi_\ell\ , B(\ell) = -\gamma_1\sin 2\phi_\ell + \gamma_2\cos 2\phi_\ell\ , \tag{1.35}$$

where ϕ_ℓ is the angle between ℓ and the coordinate. Therefore, any lensing field can be decomposed into an E-mode field, and a B-mode field.

On the other hand, a more general treatment of a distortion matrix should also include the rotation ω.

$$\mathbf{J} = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2+\omega \\ -\gamma_2-\omega & 1-\kappa+\gamma_1 \end{pmatrix}\,.$$

An important but elusive fact is that the E-mode is actually equal to κ, and the B-mode is equal to the rotation ω (Hirata & Seljak, 2003). κ can also be calculated from γ using the equation: $\kappa^2(\ell) = \gamma_1(\ell)^2 + \gamma_2(\ell)^2$. Gravity can only generate the E-mode, the B-mode (or rotation) measured from lensing experiment comes from other sources and can be used as a good estimation of the systematic noise.

1.4.2 Cosmic shear

The weak gravitational lensing effect was first studied with distant galaxies. The light from the galaxies is disturbed by the gravitational field, and therefore the apparent images of the galaxies are distorted. When the distortion is small, the effect of lensing field can be measured from the statistical change in the ellipticity of galaxies, which is called cosmic shear. Lensing can also be measured by the magnification effect (Jain, 2002; Zhang & Pen, 2005), however it is difficult to measure with data from current observations.

With various observations successfully completed in about a decade, such as VIRMOS-DESCART (Van Waerbeke et al., 2000, 2002), RCS (Hoekstra et al., 2002), CFHTLS (Hoekstra et al., 2006; Semboloni et al., 2006), COSMOS (Massey et al., 2007a,b), cosmic shear has become an important tool to constrain cosmological models. The consistency of different observations on different telescopes, self consistency of different statistics in single surveys, and the agreement of cosmic shear measurements with other cosmological observation methods, have all shown that cosmic shear signals are from cosmic structures (Van Waerbeke & Mellier, 2003).

To compare observations with theory, two-point statistics of κ or γ are calculated from the shear measurement. In practice, shear correlation functions $\langle\gamma_t\gamma_t\rangle$ and $\langle\gamma_r\gamma_r\rangle$ are calculated, where γ_t and γ_r are the tangential and 45 degree rotated components of shear, respectively. These shear correlation functions can then be separated into E-mode and B-mode components. Because lensing only generates E-mode, the B-mode component can provide a measurement of systematics. Another frequently used two-point statistics is aperture mass variance, which is the variance of the convergence map smoothed with an aperture window. The aperture mass variance can also be calculated from the measured shear correlation function through a finite integral.

Although computationally more complicated, the power spectrum, widely used in other cosmological observations such as CMB and galaxies surveys, is a more convenient choice of two-point statistics for the cosmological interpretation of lensing. Lensing power spectra were also measured from cosmic shear data (Pen et al., 2002; Padmanabhan et al., 2002). Because cosmic shear is an integrated effect of the dark matter distribution along the line of sight, one could reconstruct the dark matter distribution by inverting Limber's equation (Pen et al., 2003), similar to the extraction of the 3-D galaxy power spectrum from the galaxy angular power spectrum (Maddox et al., 1990; Baugh & Efstathiou, 1993; Dodelson & Gaztañaga, 2000; Eisenstein & Zaldarriaga, 2001; Dodelson et al., 2002). Pen et al. (2003) presented the results for the minimum variance inversion using a singular decomposition (SVD) technique with the data from the VIRMOS-DESCART survey. The evolution of the dark matter power spectrum was assumed to be linear, and the power spectra were evolved back to $z = 0$. The inverted 3-D power spectrum agrees well with the CMB and the RCS lens survey.

Higher order statistics, such skewness, have a different dependence on the cosmological parameters (Bernardeau et al., 1997; Van Waerbeke & Mellier, 2003), and can be used to break the degeneracy between Ω_m and σ_8 (Zhang et al., 2003; Vafaei et al., 2008):

$$\begin{aligned} \sigma^\kappa &= \langle \kappa^2 \rangle^{1/2} \propto \sigma_8 {\Omega_m}^{0.8} \,, \\ S_3(\kappa) &= \frac{\langle \kappa^3 \rangle}{\langle \kappa^2 \rangle^2} \propto {\Omega_m}^{-0.8} \,. \end{aligned} \tag{1.36}$$

The three-point correlation in Fourier space (bispectrum) was also measured (Pen et al., 2003). The predictions on the three-point or even higher order statistics can be calculated from hierarchical perturbation theory or measured from ray-tracing simulations of lensing (Schneider et al., 1998; Dodelson & Zhang, 2005).

The estimation of the cosmological parameters can be performed by maximizing the likelihood function. Hoekstra et al. (2006) found that $\sigma_8 = 0.85 \pm 0.06$ with a fiducial matter density $\Omega_m = 0.3$, which agrees well with previous works.

Cosmic shear, or weak lensing in general, is not only useful for the mapping of dark matter, but also helpful for exploring dark energy. Dark energy enters the calculation in two ways. First, the growth of structure depends on $G(z)$, and second, the comoving distances $D(z)$ between the source, lens, and observer are very sensitive to dark energy. Both $G(z)$ and $D(z)$ are sensitive functions of redshift. Hoekstra et al. (2006); Semboloni et al. (2006) tried to investigate the dark energy constraint from CFHTLS deep and wide surveys data assuming a constant equation of state w. They both found $w < -0.8$ at 68% confidence level. In the report of Dark Energy Task Force (DETF)[4] (Albrecht et al., 2006), four methods, including Baryon acoustic oscillations (BAO), galaxy cluster counting, supernova, and weak lensing, were compared using their capacity to constrain dark energy. They reported that weak lensing is the method with the greatest potential, though it is likely to be limited by unknown systematics.

There are a few practical factors which limit the precision of weak lensing measurements, even though the underlying physics of weak lensing is very clean. First of all, because cosmic shear produces only a small percentage of change on the ellipticities of galaxies, measuring it from images of galaxies is a challenging problem (Heymans et al., 2006; Massey et al., 2007c). The images are convolved with the Point Spread Function (PSF) of the atmosphere, telescope, and camera. This is the main concern for ground based surveys, though the images from space based surveys are less affected. Secondly, the lensing effect depends on the geometry between the source, lens, and observer. Measuring the redshift of more than thousands of galaxies accurately is a crucial problem. Current surveys normally use photometric redshifts, which could add uncertainties to the results (Ma et al., 2006). Finally, there is contamination from the unknown intrinsic ellipticity of galaxies and their correlation with neighboring galaxies caused by local tidal forces (Croft & Metzler, 2000; Heavens et al., 2000; Lee & Pen, 2001; Crittenden et al., 2001). One can try to eliminate these effects by not counting the galaxies in the same redshift bins when the two-point functions of shear are calculated.

---

[4] http://home.fnal.gov/ rocky/DETF/

With numerous next generation surveys such as DUNE[5], JDEM[6], PanSTARRS[7], LSST[8] from both space and the ground, cosmic shear seems very promising as a tool to understand dark energy (Albrecht et al., 2006). As we mentioned earlier, the dark matter distribution itself follows an non-Gaussian distribution. The non-Gaussianity degrades the Fisher information content of dark matter power spectrum by order of magnitudes, and causes saturation effect at trans-linear scales (Rimes & Hamilton, 2005). It will be useful to study how much this will affect the dark energy constraints resulting from cosmic shear surveys.

Because the statistics of cosmic shear are non-Gaussian, to correctly estimate the constraints on cosmological parameters, using the likelihood function, requires the input of higher order statistics such as the covariance matrix of the correlation function or the power spectrum. This is often provided by ray-tracing lensing simulations. As we will explain in chapter 2, the four-point statistics of the κ field generated from ray-tracing simulations are biased because of the inevitable recycling of simulation boxes. This issue is less serious with the two-point statistics. In chapter 2, we will present a new method for calculating the covariance matrix from simulations, and probing its constraints on dark energy parameters.

1.4.3 Lensing reconstruction from diffuse background

Galaxies are plentiful on the sky, but their intrinsic properties are not understood from first principles, and must be measured from the data. Future surveys may map as many as 10^{10} source objects. Using galaxies as lensing sources has several potential limits (Hirata & Seljak, 2004), including the need to calibrate redshift space distributions and PSF corrections, to be better than the desired accuracy, say 1%. This will be challenging

---

[5] http://www.dune-mission.net

[6] http://jdem.gsfc.nasa.gov/

[7] http://pan-starrs.ifa.hawaii.edu

[8] http://www.lsst.org

for the next generation of experiments.

Some sources, such as the CMB, are in principle very clean, since its redshift and statistical properties are well understood. Unfortunately, there is only one 2-D CMB sky with an exponential damping at $\ell \gg 1000$, which limits the number of source modes to $\sim 10^6$.

The potential of detecting the 21-cm background from the dark ages will open a new window for cosmological detections. Studying the 21-cm background as high redshifts lensing source, as well as the physics of the 21-cm background itself, will provide rich and valuable information to the evolution of universe. The number of modes on the sky is potentially very large, with numbers of 10^{16} or more. For this reason, 21-cm lensing has recently attracted attention.

The lensing reconstruction method for CMB and 21-cm background are different from galaxies. Galaxies are resolved lensing sources, and thus the lensing effect can be determined by the change in the shape of galaxies. In diffuse sources, individual objects can not be distinguished and only a diffuse background can be mapped. For such sources, quadratic estimators need to be constructed to extract the lensing signal. I first heuristically review quadratic lensing estimation in two dimensions. Then I briefly introduce the works done with CMB Lensing in 2-D. Later I will proceed with a generalization to 3-D with a quantitative derivation in Chapter 3 and 4.

Lensing changes the spatial distribution of a temperature field. Lensing estimation relies on statistical changes to quadratic quantities in the source plane temperature field. We use a tilde to denote a lensed quantity. All estimators work by convolving the temperature field with a window,

$$\tilde{T}_1(\boldsymbol{x}) = \int \mathrm{d}^2x' \tilde{T}(\boldsymbol{x}') W_1(\boldsymbol{x} - \boldsymbol{x}') \ , \tag{1.37}$$

and a second window

$$\tilde{T}_2(\boldsymbol{x}) = \int \mathrm{d}^2x' \tilde{T}(\boldsymbol{x}') W_2(\boldsymbol{x} - \boldsymbol{x}') \ . \tag{1.38}$$

The quadratic estimator is simply the product of the two convolved temperature fields,

$$E(\boldsymbol{x}) \equiv \tilde{T}_1(\boldsymbol{x})\tilde{T}_2(\boldsymbol{x}). \tag{1.39}$$

In the weak lensing case, the estimator is a linear function of the weak lensing parameters (κ, γ). The simplest case is two equal, azimuthally symmetric window functions $W_1 = W_2 = f(r)$. We first consider the limit where κ is a constant value, then the estimator is linearly proportionate to κ:

$$\langle E(\boldsymbol{x})\rangle \propto \kappa + V \ , \tag{1.40}$$

and V is a normalization constant proportional to the mean covariance of the smoothed temperature field when lensing is absent $\langle T_1(\boldsymbol{x})T_2(\boldsymbol{x})\rangle$. Here $\langle ... \rangle$ means ensemble average. For a stochastic random field, the ensemble average can be calculated by the volume average if the volume is big enough. We can absorb V as well as the normalization coefficient into E for convenience, i.e., $E(\boldsymbol{x}) \equiv \tilde{T}_1(\boldsymbol{x})\tilde{T}_2(\boldsymbol{x}) - V$. Therefore $E(\boldsymbol{x})$ is an unbiased estimator of κ, i.e., $\langle E(\boldsymbol{x})\rangle = \langle \hat{\kappa}(\boldsymbol{x})\rangle = \kappa$. When κ is spatially variable, E needs to be normalized by a scale dependent factor $b(k)$. This corresponds to a convolution of κ with a kernel:

$$\langle E(\boldsymbol{x})\rangle = \int \mathrm{d}^2x' \kappa(\boldsymbol{x}')b(\boldsymbol{x} - \boldsymbol{x}') \ , \tag{1.41}$$

where kernel b is the Fourier transform of the normalization factor. Therefore $E(\boldsymbol{k}) = b(\boldsymbol{k})\hat{\kappa}(\boldsymbol{k})$, here $\hat{\kappa}(\boldsymbol{k})$ is the Fourier transform of the estimator $\hat{\kappa}(\boldsymbol{x})$.

One can optimize the functions to minimize the error on the lensing variables. In this thesis we will compare various forms of the smoothing windows, which include as special case the traditional Optimal Quadratic Deflection Estimator (OQDE). The simplest case is a constant value of κ, for which one can compute its variance

$$\langle \hat{\kappa}^2 \rangle = \langle (\tilde{T}_1(\boldsymbol{x})^2)^2 \rangle. \tag{1.42}$$

Lensing is a small perturbation of the variance, therefore we can calculate the variance from the unlensed source field, i.e., $\langle (T_1(\boldsymbol{x})^2)^2 \rangle \approx \langle (\tilde{T}_1(\boldsymbol{x})^2)^2 \rangle$. Performing a variation to minimize the variance, one can find the optimal window function. Generally, one would think a fixed window function might not be optimal for all scales, i.e., $\langle\hat{\kappa}(\boldsymbol{k})^2\rangle$ might not be minimized at the same time for all $\boldsymbol{k}$ modes. Fortunately, it turns out that the optimal window functions do not depend on the spatial structure of the lensing field. Only the normalization factor b in Eq. (1.41) is scale dependent. We solve the optimal window function at scales $k = |\boldsymbol{k}| \lesssim k_a$, where the constant κ approximation works well. Here the characteristic scale k_a is determined by the smallest scale k_c of 21-cm sources resolved in a given experiment, as well as the non-Gaussianity of sources. We will explain the details later in section 3.3. We expect the resulting window function to be optimal for all scales, and will verify this at section 3.4.

Shear and deflection angles are tensorial and vectorial quantities and require anisotropic or vectorial choices of the window function

$$\mathbf{E}_\gamma = \tilde{\boldsymbol{T}}_1\tilde{\boldsymbol{T}}_2,\, \tilde{\boldsymbol{T}}_1 = \int \mathrm{d}^2\theta'\tilde{T}(\boldsymbol{\theta}')\boldsymbol{W_1}(\boldsymbol{\theta}-\boldsymbol{\theta}'),\, \tilde{\boldsymbol{T}}_2 = \int \mathrm{d}^2\theta'\tilde{T}(\boldsymbol{\theta}')\boldsymbol{W_2}(\boldsymbol{\theta}-\boldsymbol{\theta}')\ , \tag{1.43}$$

$$\boldsymbol{E}_\mathrm{d} = \tilde{\boldsymbol{T}}_1\tilde{T}_2,\, \tilde{\boldsymbol{T}}_1 = \int \mathrm{d}^2\theta'\tilde{T}(\boldsymbol{\theta}')\boldsymbol{W_1}(\boldsymbol{\theta}-\boldsymbol{\theta}'),\, \tilde{T}_2 = \int \mathrm{d}^2\theta'\tilde{T}(\boldsymbol{\theta}')W_2(\boldsymbol{\theta}-\boldsymbol{\theta}')\ . \tag{1.44}$$

This will be explained in detail in section 3.2 and section 3.2.1.

The source is usually treated as a Gaussian random field in the literature on reconstruction methods. While this is valid for CMB on large angular scales, 21-cm background sources are not necessarily Gaussian. In this thesis we attempt to understand the influence of this non-Gaussianity. Optimal estimators for Gaussian sources are not necessarily optimal for non-Gaussian sources. Here, we will construct the convergence and shear field directly, instead of following the deflection angles or potential field reconstruction in CMB lensing. There are three reasons to do this: Firstly, the strength of lensing is evident through the magnitude of κ or γ since they are dimensionless quantities. The rms deflection angle of photons from 21-cm emission is at the magnitude of a few arcmin, which is comparable to the lensing scales we are resolving. Some authors argued that perturbation theory on the deflection angle will break down at these scales

(Cooray, 2004; Mandel & Zaldarriaga, 2006). However, κ and γ are still small and can still work with perturbation calculations without ambiguity. Secondly, κ and γ have well defined limits as they approach a constant, while only spatially variable deflection angles or potentials can be measured. This significantly simplifies the derivations. Finally, κ and γ are standard variables to use in broader lensing studies, such as strong lensing and cosmic shear. Using the same convention in different subfields will help to generalize the underlying physics of lensing.

In chapter 3, we will derive the optimal quadratic estimators for κ and γ separately assuming a Gaussian source distribution. These estimators are unbiased, as shown in appendix A, though they are non-optimal for non-Gaussian sources. Furthermore, we confirm that our combined estimators for κ and γ have the same optimality as the OQDE for Gaussian sources. We generate non-Gaussian sources from cosmological simulations, and apply both OQDE and our estimators. We find our new estimators are superior.

In chapter 4, we construct the optimal estimator for non-Gaussian sources. The estimator is composed by the inverse covariance matrix of source power spectrum, which is measured from a large number of independent N-body simulations. We find the optimal estimator have a few times fewer noise than the Gaussian optimal estimator of chapter 3.

2-D case: CMB Lensing

Though my thesis mainly focuses on lensing reconstruction with 21-cm sources, the lensing estimators we will use in chapter 3 and 4 use the same approach as CMB lensing, and therefore it is helpful to explain what CMB lensing is.

The CMB, which originates from the last scattering surface, is the farthest source which can be observed. The CMB is an approximately uniformly distributed background radiation at a temperature of 2.725K today, with a fluctuation level of 10^{-5}. According to inflation theory, the CMB fluctuation follows a Gaussian distribution. The CMB photons are distorted by the large-scale structure as they propagate from $z \sim 1100$ to present, and can be used to explore the large scale structure from present time to the last scattering surface at redshift 1100. This will change the statistical distribution of the background radiation. We will discuss the lensing of CMB by large-scale structure. The CMB is also lensed by clusters (Seljak & Zaldarriaga, 2000), which will not be discussed here. The deflection is at about the 2 arcmin level (Lewis & Challinor, 2006). The CMB power spectrum is suppressed by a few percent at $\ell \sim 1000$, and has an added tail at $\ell > 6000$. This makes it hard to distinguish the lensing from CMB power spectrum directly.

Seljak and Zaldarriaga (Zaldarriaga & Seljak, 1999; Seljak & Zaldarriaga, 1999) first started to extract lensing information by constructing quadratic estimators which are two-point function of the lensed temperature field. The estimators are the natural results of perturbation theory. They constructed the estimators with

$$\begin{aligned} E_\kappa &= \frac{-\sigma_S^{-1}[\tilde{T}_{,x}^2 + \tilde{T}_{,y}^2](\boldsymbol{\theta}) + 1}{2}\,, \\ E_{\gamma_1} &= \frac{-\sigma_S^{-1}[\tilde{T}_{,x}^2 - \tilde{T}_{,y}^2](\boldsymbol{\theta})}{2}\,, \\ E_{\gamma_2} &= -\sigma_S^{-1}[\tilde{T}_{,x}\tilde{T}_{,y}](\theta)\,, \end{aligned} \tag{1.45}$$

where $\tilde{T}_{,x}, \tilde{T}_{,y}$ are the gradients of the lensed temperature field at the two perpendicular directions, and σ_S is the normalization constant. These quadratic estimators can give unbiased results, i.e., $\langle E_\kappa \rangle = \kappa$, $\langle E_{\gamma_1} \rangle = \gamma_1$, $\langle E_{\gamma_2} \rangle = \gamma_2$, however they do not provide the minimum noise. A widely used estimator is the minimum variance quadratic estimator for the deflection angle or potential (Hu, 2001). The original form of their estimator is

$$\boldsymbol{E}_{\rm d}(\boldsymbol{\ell}) = {\rm i}\ell N(\boldsymbol{\ell}) \int \frac{d^2\ell}{(2\pi)^2} \tilde{T}(\boldsymbol{\ell}')\tilde{T}(\boldsymbol{\ell}' - \boldsymbol{\ell}) g(\boldsymbol{\ell}, \boldsymbol{\ell}')\,, \tag{1.46}$$

$$g(\boldsymbol{\ell}, \boldsymbol{\ell}') = \frac{(\boldsymbol{\ell} - \boldsymbol{\ell}') \cdot \boldsymbol{\ell} C_{|\boldsymbol{\ell} - \boldsymbol{\ell}'|} + \boldsymbol{\ell}' \cdot \boldsymbol{\ell} C_{\boldsymbol{\ell}}}{2\tilde{C}^{\rm tot}_{\boldsymbol{\ell}'}\tilde{C}^{\rm tot}_{\boldsymbol{\ell} - \boldsymbol{\ell}'}}\,, \tag{1.47}$$

where $N(\boldsymbol{\ell})$ is a normalization factor, which happens to be the reciprocal of the reconstruction noise; $\tilde{T}(\boldsymbol{\ell})\tilde{T}(\boldsymbol{\ell}') = \delta^{\rm 2D}(\boldsymbol{\ell} - \boldsymbol{\ell}')\tilde{C}_\ell$, $\tilde{C}^{\rm tot}_\ell = \tilde{C}_\ell + C^N_\ell$, $\tilde{C}_\ell$ is the power spectrum of the lensed CMB temperature map, and C_ℓ^N is the experimental noise contribution. We could rewrite the estimator as the product of two smoothed temperatures:

$$\boldsymbol{E}_{\mathrm{d}}(\boldsymbol{\theta}) = \boldsymbol{T}_1(\boldsymbol{\theta}) T_2(\boldsymbol{\theta}) \ , \tag{1.48}$$

and

$$\boldsymbol{E}_{\mathrm{d}}(\boldsymbol{\ell}) = b_d(\boldsymbol{\ell}) \boldsymbol{d}(\boldsymbol{\ell}) \, . \tag{1.49}$$

b_{d} is a normalization factor, $\boldsymbol{T}_1 = \int \mathrm{d}^2\theta' T(\boldsymbol{\theta}') \boldsymbol{W_1}(\boldsymbol{\theta} - \boldsymbol{\theta}')$ and $T_2 = \int \mathrm{d}^2\theta' T(\boldsymbol{\theta}') W_2(\boldsymbol{\theta} - \boldsymbol{\theta}')$ are convolved temperature fields, where the window functions are Fourier transforms of (Lewis & Challinor, 2006):

$$\begin{aligned} \boldsymbol{W}_1(\boldsymbol{\ell}) &= \frac{-\mathrm{i}\boldsymbol{\ell} C_{\boldsymbol{\ell}}}{\tilde{C}_{\boldsymbol{\ell}}^{\mathrm{tot}}}, \\ W_2(\boldsymbol{\ell}) &= \frac{1}{\tilde{C}_{\boldsymbol{\ell}}^{\mathrm{tot}}} \ . \end{aligned} \tag{1.50}$$

The CMB is a 2-D distribution, so that there is one CMB sky behind one patch of lensing field. The CMB lensing reconstruction itself can be very noisy. However, this can be improved by cross correlating a lensed CMB map with other cosmic structure tracers such as galaxies distributions. The lensing signal was first detected in the CMB by Smith et al. (2007): They found a 3.4σ detection in the cross correlation of the all sky WMAP CMB map and the radio galaxies in the NRAO VLA Sky Survey (NVSS).

3-D case: 21-cm Lensing

In this thesis, I will mainly discuss the lensing of 21-cm emission from neutral Hydrogen gas, from before and during the epoch of ionization ($z \sim 7-20$), and at lower redshifts ($1 < z < 6$).

The 21-cm line emission is the transition between the two spin-flip states status of ground state neutral hydrogen. In the Rayleigh-Jeans limit, the brightness temperature T_b of 21-cm radiation in a patch of the sky is equal to

$$T_b = T_{\mathrm{CMB}} e^{-\tau} + T_s (1 - e^{-\tau}) \, , \tag{1.51}$$

where $T_{\rm CMB}$ is the temperature of CMB photons, T_s is the spin temperature, and the optical depth (Iliev et al., 2002) can be calculated as

$$\tau = \frac{3\lambda_0^3 A_{10} T_\star n_{\rm HI}(z)}{32\pi T_s H(z)} = \frac{0.28}{T_s}(\frac{1+z}{10})^{3/2}\frac{\rho_{\rm HI}}{\langle\rho_H\rangle}\,, \tag{1.52}$$

$T_\star = 0.068$ K came from the the energy difference E_{10}, A_{10} is the Einstein coefficient, $\lambda_0 = 21.16$ cm is the rest frame wavelength of the emission, $n_{\rm HI}$ and $\rho_{\rm HI}$ are the number density and density of neutral hydrogen, and $\langle\rho_H\rangle$ is the mean density of total hydrogen. Therefore, the increment of brightness temperature on top of the CMB background $\delta T_b \approx (T_s - T_{\rm CMB})\tau/(1+z)$ is also proportional to the ionization ratio and the density at that region. The spin temperature of 21-cm radiation is defined by

$$\frac{n_1}{n_0} = \frac{g_1}{g_0}e^{-E_{10}/k_B T_s}\,, \tag{1.53}$$

where n_1, n_0 are the number densities of electrons at the two spin level, $g_1/g_0 = 3$ is the ratio of spin degeneracy factors, and k_B is the Boltzmann constant. When the spin temperature is higher than the CMB temperature, emission will be observed; while absorption will be observed when T_s is lower than $T_{\rm CMB}$. For $z \approx 140-1000$, $T_K \approx T_{\rm CMB}$ because of Thompson scattering of CMB photons. For $z \approx 20-140$, the gas decouples from the CMB photons and cools adiabatically, and T_s is coupled to T_K by collisions, which is below $T_{\rm CMB}$. Later, when the first generation of stars form and ionize the surrounding gas, the kinetic temperature of the gas become higher than $T_{\rm CMB}$ from heating by X-rays, Lyα photons, and shocks. T_s is coupled to T_K by Lyα pumping (or Wouthuysen-Field effect) (Wouthuysen, 1952), $T_s \approx T_K \gg T_{\rm CMB}$ (Ciardi & Madau, 2003). We can generally observe 21-cm emissions from re-ionization.

Because the 21-cm line is a forbidden emission line, the optical depth τ is much smaller than 1. Therefore, the 21-cm sources up to very high distances ($z \sim 1-140$) can be observed. Furthermore, Hydrogen is the most abundant element in the universe (75% of total), and 21-cm line emission is a rich source for cosmological detections. We can use 21-cm line to probe a large part of the Hubble volume. Most importantly, since

21-cm line is a line emission, the redshifts of sources are determined straightforwardly by comparing the frequency of observation with the intrinsic wavelength.

The ongoing 21-cm surveys, such as the Giant Metre Radio Telescope (GMRT) experiment (Pen et al., 2008), 21CMA[9], and many upcoming experiments such as Low Frequency Array (LOFAR[10]) (Kassim et al., 2000), Murchison Widefield Array (MWA[11]), and Square Kilometer Array (SKA[12]) (van Haarlem, 2000), make it promising to study 21-cm for both cosmological and astrophysical interests.

There are a few advantages to use the 21-cm line as lensing sources. The 21-cm sources have structures on small scales and exist between a broad range of redshifts, which indicates a high source number density and thus less noise. the 21-cm sources are located at high redshift, so that the lensing signal is strong. Furthermore, the 21-cm map naturally contains three-dimensional information. The redshift distribution of sources can be determined with small errors.

21-cm emission is similar to the CMB since both are diffuse background. Therefore, the techniques used in CMB lensing can be applied to the 21-cm lensing study. On the other hand, 21-cm is unlike CMB, which has an intrinsic Gaussian distribution. Non-linear gravitational clustering leads to non-Gaussianity, and ultimately to reionization. In contrast to CMB lensing, where the Gaussian assumption works well, non-Gaussianity in 21-cm lensing may affect the results. However, most of the reconstruction methods are based on a Gaussian source distribution assumption, and their results are also discussed for Gaussian sources. (Pen, 2004; Cooray, 2004; Zahn & Zaldarriaga, 2006; Metcalf & White, 2007; Hilbert et al., 2007).

In chapter 3, we will address this non-Gaussianity problem of the lensing of 21-cm sources at the epoch of reionization. We develop optimal estimators for κ and γ

---

[9] http://web.phys.cmu.edu/ past/, http://21cma.bao.ac.cn/

[10] http://www.lofar.nl/

[11] http://www.mwatelescope.org/

[12] http://www.skatelescope.org/

under a Gaussian assumption. Non-Gaussian sources are generated from cosmological simulations, and we examine the influence of non-Gaussianity with the estimators using the simulated sources. We find that non-Gaussianity could increase the reconstruction noise by orders of magnitude.

21-cm radiation from post reionization redshifts ($z < 6$) might be easier to detect, and will allow data to be obtained sooner compared with those from the epoch of reionization, because their distance is closer. Interests in mapping these 21-cm sources are shown by multiple groups (Peterson et al., 2006; Tegmark & Zaldarriaga, 2008; Wyithe et al., 2008; Chang et al., 2008). The most recent one, CHIME (Chang et al., 2008) proposed to map the intensity of 21-cm sources at $z \sim 1-3$ for half of the sky in two years of operation. The original goals of these surveys are to measure the BAO, where the sound horizon scale, $r_s \sim 150$Mpc, works as a standard ruler for cosmological distance. If we use these 21-cm intensity maps at $z \sim 1-3$ as lensing sources, the lenses at $z \sim 0.5-1$ can be explored, where the cosmic structures are very sensitive to different models of dark energy. In Chapter 4, we numerically construct the optimal estimator for non-Gaussian sources by running hundreds of N-body simulations. We will show that the reconstruction noise from the optimal estimator is a few times lower than the optimal Gaussian estimator, and reduced dramatically by a large sky coverage experiment like CHIME.

1.5 Cosmological simulations and the application in lensing

The large-scale structure of the universe can be well described by linear perturbation theory in many research fields. However, because lensing phenomena is sensitive to non-linear effects at $\ell \geq 1000$ (Jain & Seljak, 1997). We have to rely on N-body cosmological simulations rather than linear perturbation theory. In addition, the covariance of shear correlations (or power spectra) is a four-point function of the lensing field. Wick's theorem can not be applied, because the lensing field is not Gaussian. The covariance should be measured directly from simulations.

1.5.1 Cosmological simulations

Many cosmological simulations are using N-body dynamics (Efstathiou et al., 1985; Bertschinger, 1998). The universe are assumed to be homogeneous so that it is sufficient to set periodic boxes, usually about a few hundred Mpc at each dimension. The box needs to be big enough to be a fair sample of the universe. The dark matter and gas distribution are approximately represented by millions or even billions of particles in a cosmic volume. Because simulations are conducted at very large scale and motion of the particles are much smaller compared to the speed of light, Newtonian gravity law can be used. Particles are evolved forward in time with Newtonian force interaction. The forces are calculated from density distribution by solving Poisson equation through Fast Fourier Transform (FFT).

There are two representative schemes in cosmological simulations: particle-mesh (PM) and Tree method. Several improvements on these methods are achieved such as particle-particle/particle-mesh(P3M), adaptive mesh algorithm and some hybrid methods. The Tree algorithm could achieve high resolution, and is good at studying local physical processes such as galaxy formation. On the other hand, the featured advantage of PM algorithm is its faster speed than Tree. Because we need a large number of simulations to understand the uncertainties in lensing statistics, we used a PM algorithm.

The initial power spectrum of density is calculated from

$$P(k) = Ak^n|\mathcal{T}(k)|^2 \,, \tag{1.54}$$

where $\mathcal{T}(k)$ is the transfer function, and n is the spectral index. The amplitude A can be determined by the normalization of σ_8. We normalize it from WMAP measurement in this thesis.

The initial particle distribution is a Gaussian random field. To generate the particle distribution, we generally apply the Zeldovich approximation. The initial displacement of particles can be written as

$$\boldsymbol{x}(t) = \mathbf{q} + G(t)\mathbf{f}(\mathbf{q}), \tag{1.55}$$

$$\mathbf{v} = a\frac{d\boldsymbol{x}}{dt} = a\frac{\mathrm{d}G}{\mathrm{d}t}\mathbf{f}\,, \tag{1.56}$$

where $\boldsymbol{x}$ is the Eulerian position, $\mathbf{q}$ is the Lagrangian position, $\mathbf{f}(\mathbf{q})$ is the displacement field as a function of coordinate. If there is a perturbation, the gravitational collapse first happens along the shortest axis and the structure becomes flattened. The displacement field can be solved by linearized continuity equation,

$$\nabla \cdot \mathbf{f} = -\frac{\delta}{G(t)}\,, \tag{1.57}$$

where $\mathbf{f} = \nabla\psi$ is the gradient of a potential. The Zeldovich approximation is the first order perturbation theory in the Lagrangian coordinate.

1.5.2 Multiple-plane ray-tracing lensing simulation

In the cosmological scale, the dark matter distribution is approximated by a series of mass sheet, called "multiple lens-plane approximation". The Jacobin matrix is calculated by the sum of Jacobin at individual lens plane (thin lens approximation).

$$\mathbf{J} = \sum \mathbf{J}(z_i) \tag{1.58}$$

The higher order terms including the correlation between lens planes are neglected.

In simulations, each output volume is treated as the projected mass sheet at the mid-plane. Rays are shot through these planes, and deflections are calculated on each plane. This treatment is valid under the Born approximation. There are also many modifications to the ray-tracing method to examine the effect of various approximations, e.g., Born approximation, Limber approximation, secondary effect such as Lens-Lens coupling (White & Hu, 2000; Jain et al., 2000; Hamana & Mellier, 2001; Hamana et al., 2002; Semboloni et al., 2007). As we will present in Chapter 2, a large number of independent simulations are much more important than the correction of the Born and Limber approximations in our study of lensing statistics uncertainties. We also find a way to calculate the Fisher matrix of lensing power spectra from a finite number of simulations without worrying about the redundancy of simulation data.

An alternative approach of lensing simulations is the ray bundle method. Barber et al. (1999) use bundles of rays to represent a circular images and calculate the change of the shapes and area. However, their approaches are more complicated, and not suitable for lensing statistics studies with a large number of simulations.

1.5.3 Simulation code

Our numerical simulations in this thesis are performed using the publicly available N-body codes PMFAST (Merz et al., 2005) and CUBEPM[13].

PMFAST is a two-level mesh PM code highly parallized and optimized for speed. PMFAST is supported by distributed memory systems through MPI and shared memory via OpenMP. The gravitational forces are decomposed into a long-range part and a short-range one. The forces are then computed on the two level of meshes. The coarse mesh is global, and four times coarser than the local fine mesh. Most of the force calculations are done locally on the fine mesh. The global coarse forces are updated after several time steps, while the fine forces are calculated every time step. The simulation volume is decomposed along one dimension (slab decomposition). Computation of each part is performed on one node. The communications between these nodes are realized by MPI. Further shared-memory parallelization is done via OpenMP to optimize memory usage on shared memory nodes. PMFAST was run on the lobsters cluster of CITA. The largest run is on a 3712^3 mesh with 6.4×10^9 particles, which take nearly a week to finish at

---

[13] http://www.cita.utoronto.ca/mediawiki/index.php/CubePM

$z = 0$ with about 300 time steps. We will use PMFAST in chapter 3.

CUBEPM is the successor to PMFAST: the decomposition of the simulation volume in CUBEPM is done cubically, which allows more nodes to be used, making the code much faster. In addition to the features provided by PMFAST, CUBEPM contains support for gas evolution through use of a TVD MHD module, and has particle-particle forces implemented at sub-grid scales. To improve the force resolution, CUBEPM has the option to use a P3M function. The short-range gravitational force is further decomposed to a super-grid and sub-grid parts. The super-grid force is done by FFT while the sub-grid force is calculated by direct sum of forces between particles (particle-particle or PP calculation). The PP is usually performed at distances of a few grid sizes. The simulation volume (which is also called simulation box) is cubically decomposed to n^3 sub-volumes, and the calculation of each sub-volume is performed on one node of cluster, i.e., the total number of nodes used in simulation is n^3. The code has been optimal scaling up to (and hopefully beyond) 1000's of nodes. A typical 512^3 mesh simulation with 8 nodes, utilized in this thesis (chapter 4), takes 2-3 hours, and a 1024^3 mesh one with the same number of nodes (chapter 2) takes about 5 hours to finish. This makes it possible for us to run hundreds of simulations for many different redshifts and resolutions, which is crucial to our study of statistical uncertainties.

My thesis is organized as follows: In chapter 2, I will present the Fisher information and dark energy constraints with a new method of generating the covariance matrix of the lensing power spectrum from simulations. The content is based on Dore et al. (2009). In chapter 3, I will present our work on the lensing reconstruction using 21-cm at redshift 10 to 20, which is based on Lu & Pen (2008) (©2008. Royal Astronomical Society). In chapter 4, I will discuss the lensing of 21-cm source at post reionization era and show its utilization in the constraint of dark energy models at $\ell < 100$. This chapter is based on Lu et al. (2009). I conclude in chapter 5.

Chapter 2

The information content of cosmic shear surveys

2.1 Overview

What is the information content of cosmic shear surveys? Although this question has been addressed many times, given the current interest in cosmic shear surveys, to answer it accurately turns out to be a non-trivial task. It is the purpose of this work to offer a precise answer to this question.

The measure of cosmic shear (Bartelmann & Schneider, 2001; Van Waerbeke & Mellier, 2003) is considered to be one of the most promising observational tools to understand the origin of the accelerating expansion of the universe (Albrecht et al., 2006). Commonly attributed to the existence of some extra unknown physics loosely labelled Dark Energy (DE), its exact nature has become a salient question in contemporary cosmology (Riess et al., 2004; Astier et al., 2006; Eisenstein et al., 2005; Cole et al., 2005; Tegmark et al., 2006; Komatsu et al., 2008; Kilbinger et al., 2008). Characterising the physical properties of dark energy is the main scientific drive for the development of new and ambitious surveys.

Whereas new cosmic shear surveys are being advocated (Albrecht et al., 2006) and designed (Amara & Réfrégier, 2007), we want to study in this work how the 3-D Fisher information saturation in the dark matter power spectrum translates into the cosmic shear observables. In particular, one question we would like to answer is whether there exists a scale above which the Fisher Information for the amplitude of the power spectrum (projected onto the dark energy figure of merit) saturates. Since above a given angular scale ($\ell \geq 2000$), our lack of precise modeling of the physics of baryons might require enormous efforts (White, 2004; Zhan & Knox, 2004; Rudd et al., 2008), it would be interesting to know whether such a saturation happens and in particular how it does compared to this theoretical uncertainty scale. To tackle this question will require computing the cosmic shear errors in the fully non-linear regime. While this question has already been investigated in the literature (White & Hu, 2000; Cooray & Hu, 2001; Semboloni et al., 2007; Takada & Jain, 2008), we will address it using numerical N-body simulations to probe accurately the full non-linear regime (neglecting baryons), and a new way to build the covariance matrix from those quantities. This will lead to the improvement of an order of magnitude in accuracy compared to previous numerical works in the literature. We will quantify this statement by measuring the errors on the errors using bootstrap techniques.

In this chapter, we first introduce the methodology of our work in section 2.2, and then revisit the 3-D matter power spectrum results in section 2.3, which works as an introduction to cosmic shear case developed in section 2.4. We discuss in section 2.5 the practical consequences of these results for current and coming optical surveys, as well as for the CMB lensing.

2.2 Methodology

If we want to infer a set of parameters λ_i from observables $\mathbf{y}$ of dimension n_x following a multi-variate Gaussian distribution with a covariance matrix $\mathbf{C} \equiv \langle \mathbf{y}\mathbf{y}^t \rangle - \langle \mathbf{y} \rangle \langle \mathbf{y} \rangle$, the Fisher Information matrix is defined as: (Fisher, 1936; Tegmark et al., 1997, 1998)

$$\mathbf{F}_{ij} \equiv \frac{1}{2}\mathrm{tr}\left[\mathbf{C}^{-1}\frac{\partial \mathbf{C}}{\partial \lambda_i}\mathbf{C}^{-1}\frac{\partial \mathbf{C}}{\partial \lambda_j}\right] + \frac{\partial \langle \mathbf{y} \rangle}{\partial \lambda_i}\mathbf{C}^{-1}\frac{\partial \langle \mathbf{y} \rangle}{\partial \lambda_j}. \tag{2.1}$$

Its relevance for parameter estimation can be seen from the Cramér-Rao inequality stating that the Fisher matrix sets a lower bound on how well a parameter λ_i can be measured, i.e., $\sigma^2(\lambda_i) \geq 1/\mathbf{F}_{ii}$. We assume from now on that the covariance matrix $\mathbf{C}$ does not depend on λ_i (for a discussion in the context of cosmic shear, see Eifler et al. (2008)). Following Rimes & Hamilton (2005) (hereafter RH05), we define the Fisher information content of $\mathbf{y}$ as

$$Inf \equiv \sum_{ij} \mathbf{F}_{ij} = \sum_{ij} \frac{\partial \langle \mathbf{y} \rangle}{\partial \lambda_i}\mathbf{C}^{-1}\frac{\partial \langle \mathbf{y} \rangle}{\partial \lambda_j}. \tag{2.2}$$

In this chapter, the observables we will consider will be either the 3-D matter power spectra, $\mathbf{y} = P(\mathbf{k})$, or the 2-D convergence power spectra $\mathbf{y} = C_\ell^\kappa$ defined in Eq. (2.9). The parameters we will focus on will be the standard cosmological parameters, $\lambda = (w_0, w_a, \Omega_m, \Omega_b, n_S, \sigma_8)$ (Albrecht et al., 2006). We consider a flat cosmological model where the density is dominated at late times by dark energy whose equation of state evolves as $w = w_0 + w_a(1-a)$. The nominal values for these parameters corresponding to the currently favored model are $\lambda = (1., 0., 0.1334, 0.0228, 0.963, 0.796)$ (Komatsu et al., 2008).

For pedagogical reasons, we will also consider a dimensionless version of Eq. (2.2)

$$\bar{Inf} \equiv \sum_{ij} \bar{\mathbf{C}}_{ij}^{-1} \quad , \quad \bar{\mathbf{C}}_{ij} = \frac{\langle \mathbf{y}_i \mathbf{y}_j \rangle}{\langle \mathbf{y}_i \rangle \langle \mathbf{y}_j \rangle}. \tag{2.3}$$

This form would be obtained from Eq. (2.2), if we were measuring the amplitude of a template $P(\mathbf{k})$ (or C_ℓ^κ), so that the partial derivatives were unity. Note that because of non-linear effects, this amplitude does not correspond to σ_8 or the curvature perturbation amplitude A_S. What we will define as the Fisher information in section 2.3 and 2.4 is related to the variance on the amplitude parameter, λ, if the observables are modeled as $\mathbf{y} = \lambda\bar{\mathbf{y}}$. We find this is a convenient way to visualize the property of the covariance matrix. Given this definition, since we are interested in quantifying the effects on non-linearities on the Fisher information content of the measured convergence angular power spectrum (or 3-D matter power spectrum), it will be particularly insightful to investigate the scaling of $I\bar{n}f$ with a cut-off scale ℓ_{max} (k_{max}), i.e., the cumulative Fisher information content as a function of the smallest (angular) modes measured. The comparison between the scaling on large scale (low ℓ and k) where the convergence (matter) field is expected to be Gaussian to the one in the non-linear regime (high ℓ and k), will thus be of particular relevance. For this purpose, we define in the matter power spectrum case (at wavenumber k_b),

$$\bar{\mathbf{C}}_{k_b k_{b'} < k_{max}} = \frac{\langle P_{k_b} P_{k_{b'}} \rangle}{\langle P_{k_b} \rangle \langle P_{k_{b'}} \rangle}\,, \quad I\bar{n}f(k_{max}) = \textstyle\sum_{k_b, k_{b'} < k_{max}} \bar{\mathbf{C}}^{-1}_{k_b k_{b'}}\,, \tag{2.4}$$

where P_{k_b} will be defined in Eq. (2.8), which is the matter power spectrum in a k bin. Similarly, we define in the angular power spectrum case (at multipole ℓ),

$$\bar{\mathbf{C}}_{\ell_1 \ell_2 < \ell_{max}} = \frac{\langle C^{\kappa z_1 z_2}_{\ell_1} C^{\kappa z_3 z_4}_{\ell_2} \rangle}{\sqrt{\langle C^{\kappa z_1 z_1}_{\ell_1} \rangle \langle C^{\kappa z_2 z_2}_{\ell_1} \rangle \langle C^{\kappa z_3 z_3}_{\ell_2} \rangle \langle C^{\kappa z_4 z_4}_{\ell_2} \rangle}}\,, \quad I\bar{n}f(\ell_{max}) = \textstyle\sum_{\ell_1, \ell_2 < \ell_{max}} \bar{\mathbf{C}}^{-1}_{\ell_1 \ell_2}\,, \tag{2.5}$$

where $C^{\kappa z_1 z_2}_{\ell}$ will be defined in Eq. (2.9), which is the convergence (cross-) power spectrum in a ℓ bin. The definition of $I\bar{n}f$ in Eq. (2.3) is easy to interpret since in the Gaussian case, where $\mathbf{C}_{ij} \propto \mathbf{y}_i \mathbf{y}_j \delta_{ij}$, it directly reduces to half the number of measured modes. We thus have a simple analytical prediction for the expected scaling on large scales. Note that we choose to define $I\bar{n}f(\ell_{max})$ by imposing a sharp cut-off in Fourier space. An alternative definition consists in marginalizing over all the modes above k_{max} (ℓ_{max}) by adding a white noise level so that, for example, the S/N equals to 1 at $k = k_{max}$ ($\ell = \ell_{max}$). We found that both approaches give identical results.

As visible from Eq. (2.3), the key quantity to evaluate the Fisher information content of our observables is the covariance matrix. For this purpose, we will use a Monte-

Carlo approach and generate n_{sim} realizations of the observables, $\mathbf{y}_k$, through N-body simulations. We make use of the fast, publicly available CUBEPM code.

Given a set of n_{sim} realizations for $\mathbf{y}_k$, we define an estimator for $\mathbf{C}$ as (see Hamilton et al. (2006) for a thorough discussion on how to measure $\mathbf{C}$ from one simulation only)

$$\begin{aligned} \tilde{\mathbf{C}}_{ij} &= \frac{1}{n_{sim}} \sum_{s=1}^{n_{sim}} \left(\mathbf{y}_i^s - \mu_i\right)\left(\mathbf{y}_j^s - \mu_j\right) , \\ \mu_i &= \frac{1}{n_{sim}} \sum_{s=1}^{n_{sim}} \mathbf{y}_i^s , \end{aligned} \tag{2.6}$$

where $\mathbf{y}_k^s\,(s = 1, \ldots, n_{sim})$ are the values from individual simulations. At this point, it is often missed that the inverse of a maximum-likelihood estimator for a variable y is in general not an unbiased estimator of the inverse y^{-1} (Hartlap et al., 2007). To remedy this fact, a correction factor is required. Since we also evaluate μ from our simulations, it can be shown that the following estimator for $\mathbf{C}^{-1}$ is unbiased:

$$\hat{\mathbf{C}}^{-1} = \frac{n_{sim} - n_y - 2}{n_{sim} - 1}\tilde{\mathbf{C}}^{-1}. \tag{2.7}$$

In the convergence angular (cross-) power spectrum covariance matrix $\langle C_{\ell_1}^{\kappa z_1 z_2} C_{\ell_2}^{\kappa z_3 z_4} \rangle$, we found that the number of independent modes, n_y, is not easy to define. As thus we dropped this correction factor. However, we checked carefully the convergence of our results using an increasing n_{sim} as shown in Fig. 2.1 and Fig. 2.3.

Furthermore, to quantify the error on our statements, we will evaluate the errors on $\hat{\mathbf{C}}^{-1}$ by using the bootstrap method (Efron & Tibshirani, 1993). We will consider 1000 sets of n_{sim} simulations randomly drawn from our n_{sim} simulations and apply to each the formalism defined above. This procedure allows us to weight our initial set of N-body simulations in a random manner. Even though we do not fulfil all the conditions required for the bootstrap method to be reliable (in particular our random variables are correlated), it still gives us a valuable glimpse at the reliability of our statements, i.e., the errors on the errors.

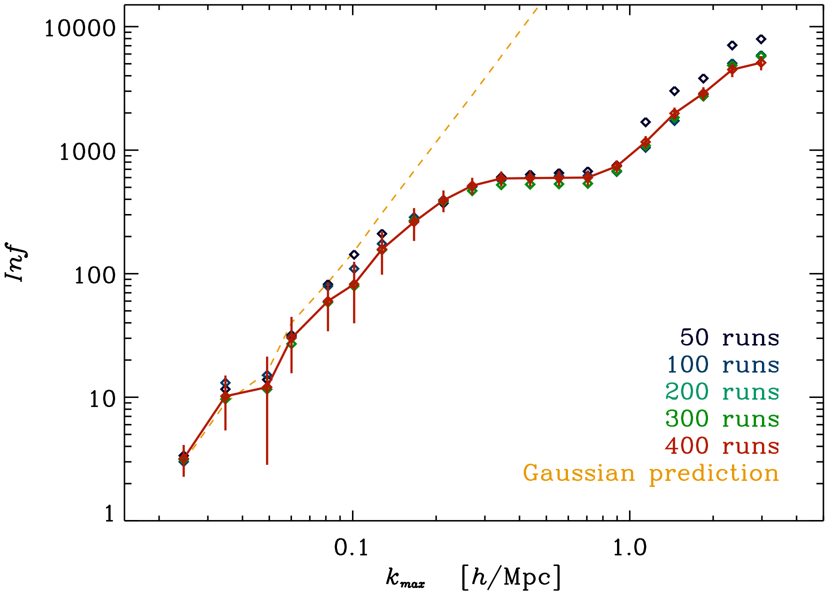

Figure 2.1 Cumulative Fisher information for the matter power spectrum at $z = 1$ as defined in Eq. (2.4). We use either $n_{sim} = 50$, 100, 200, 300 or 400 simulations. The bootstrap error bars are obtained using an analysis of 1000 sets of 400 simulations. The dashed orange line corresponds to the Gaussian prediction, *i.e.*, the number of k modes present in the simulations below k_{max}. Although the convergence does not appear to be perfect, the difference between the measurements using 300 or 400 simulations is smaller than the error bars everywhere. As such we can trust the results obtained with 400 simulations. This plot reproduces the results of RH05.

2.3 Matter power spectrum Fisher information content

We first focus on the matter power spectrum and revisit the results of RH05. For this purpose, we ran 400 N-body simulations with their choice of cosmological model, i.e., a flat ΛCDM model with $\Omega_m = 0.29$, Ω_Λ=0.71, $\Omega_b = 0.046$, $\sigma_8 = 0.97$ and h=0.71. Since we are interested in trans-linear scales, which is the transition regime between the fully linear regime and the fully non-linear regime, simulations with the comoving box-size $L = 256h^{-1}$Mpc and 256^3 grid points are appropriate to be used. This gives us roughly a mass resolution of $9.2{\times}10^{11}M_\odot$ and a force resolution of $1h^{-1}$Mpc. The initial conditions of the particles distribution were generated at $z = 200$.

To improve the convergence, increase the rank of $\tilde{\mathbf{C}}$, and alleviate numerical issues when performing the inversion, we define $n_x = 20$ bins logarithmically spaced in k space and we measure the average power spectrum within a k bin, $\mathbf{y} = P_b$, as

$$\frac{(2\pi)^3}{V} \langle \delta_{\boldsymbol{k}} \delta^*_{\boldsymbol{k}} \rangle_{\boldsymbol{k} \in b} = k_b^3 P_b \,, \tag{2.8}$$

where $\delta_{\boldsymbol{k}}$ is the Fourier transform of the matter over-density, $\boldsymbol{k} \in b$ means all modes in the bin, and $V = L^3$ is the volume of the simulation box. We then use Eq. (2.4) and (2.7) with $n_{sim} = 50, 100$, 200, 300 and 400 to compute the cumulative Fisher information content for various k_{max}. The results at $z = 1$ are plotted in Fig. 2.1 as well as the Gaussian prediction, $I\bar{n}f^G(k_{max}) = \sum_{b<k_{max}} n_{k_b} \propto k^3_{max}$. Here n_{k_b} is the number of modes in a k bin. Qualitatively similar results are obtained at different z.

First, we notice that the convergence of $I\bar{n}f$ in terms of the number of simulation used seems satisfying although not perfect. We found that the lack of convergence adds extra uncertainties of the same order as the bootstrap errors. As discovered by Rimes & Hamilton (2006), the two remarkable features of this cumulative Fisher information are the following. On linear scales ($k < 0.1h\mathrm{Mpc}^{-1}$), the Fisher information content follows the scaling expected from a Gaussian random field; on trans-linear scales, a sharp transition to a plateau is observed; in the fully non-linear regime, the Fisher information returns to a quasi-Gaussian scaling (but with a lower amplitude). These features mean that the primordial amplitude could be measured with an accuracy directly inverse proportional to the square root of the number of modes on linear scales, and not much more is learned on trans-linear scales. In other words, the Fisher information is redundant with the one contained in linear scales. On the other hand, a rising scaling appears in the fully non-linear regime – at less than 1% of the Gaussian information. The sharp transition from the linear to the non-linear regime can qualitatively be understood in the halo model framework, where it corresponds to the transition from the 2-halo term to the 1-halo term (Neyrinck et al., 2006). Whereas on large scales, the Fisher information is contained in the 2-halo term and scales as the number of modes measured (or the number of halos in a given volume), on trans-linear scales, the 1-halo term starts to dominate but with a large variance since most of its contribution comes from rare massive halos. This large variance explains why it is hard to extract any Fisher information from the trans-linear regime supplemental to the one obtained in the linear regime. On smaller scales though, the 1-halo term contribution comes mostly from numerous smaller mass halos whose number is much more constant, *i.e.*, fluctuates with much less variance, and the Fisher information increases again roughly with the number of modes probed, but about two orders of magnitude less.

Now that we have reproduced and introduced the key results regarding the 3-D matter power spectrum, we will move to the original goal of this chapter: how does this Fisher information saturation effect in the matter power spectrum translate into cosmic shear observables?

2.4 Cosmic shear Fisher information content

We now use $n_{sim} = 300$ N-body simulations with our nominal cosmology to investigate the Fisher information content of cosmic shear surveys. To quantify the Fisher information, we choose the convergence cross-power spectra between two redshift bins z_i and z_j as our observable (Bartelmann & Schneider, 2001; Van Waerbeke & Mellier, 2003):

$$\tilde{n}_{z_i}\tilde{n}_{z_j}\frac{l^2 C_\ell^{\kappa z_i z_j}}{2\pi} = \int_0^\infty dz\; W^{z_i}(z) W^{z_j}(z) \frac{H(z)}{D^2(z)} P(\ell/D(z), z) \;, \tag{2.9}$$

where $H(z)$ is the Hubble parameter, $D(z)$ is the angular diameter distance, $P(k,z)$ is the 3-dimensional matter power spectrum at redshift z. The lensing kernel is defined as

$$\begin{aligned} W^{z_i}(z) &= \frac{3}{2}\Omega_m \frac{H_0^2 D(z)}{H(z)}(1+z)\int_z^\infty dz'\; n_{z_i}(z')[1-\frac{D(z)}{D(z')}] \,, & (2.10)\\ \tilde{n}_{z_i} &= \int_0^\infty dz\; n_{z_i}(z) \;. & (2.11) \end{aligned}$$

where $n_{z_i}(z)$ is the galaxy distribution in the redshift bin i, and $\tilde{n}_{z_i}$ is the total number of galaxies in this redshift bin.

In this section, for the sake of simplicity, we will consider a uniform galaxy distribution in $n_z = 4$ redshift bins each of width $\Delta z = 0.5$, i.e., $1.0 < z < 1.5$, $1.5 < z < 2.0$, $2.0 < z < 2.5$ and $2.5 < z < 3.0$. Each distribution is normalized to unity, *i.e.*, $n_{z_i} = 1/\Delta z$ and $\tilde{n}_{z_i} = 1$. This choice is motivated by our interest in low redshift lensing of the diffuse radiation originating from the 21-cm line emission of the neutral hydrogen gas (Lu & Pen, 2008; Pen et al., 2008; Lu et al., 2009). We will consider more realistic galaxy distribution functions when discussing specific surveys in section 2.5. As discussed in chapter 4, we found that simulations with $L = 200h^{-1}$Mpc, 1024^3 grids, and 512^3 particles are close to optimal for our needs. Each one of these simulations takes about 4.5 hours using 8 nodes (64 CPU) on CITA's Sunnyvale cluster. We checked that finite resolution effects do not affect the convergence power spectra up to $\ell \simeq 10000$, which will be the smallest angular scale we consider in this work. The box-size corresponds roughly to an area of 56 square degrees. As a consequence, when considering the dimensionless cumulative Fisher information, the sum is performed for $\ell \geq 50$.

To compute the covariance matrix using $n_{sim} = 300$ simulations, we employ an original method to avoid the artifacts presented in current methods. The now standard approach to simulate cosmic shear has been pioneered in White & Hu (2000); Jain et al. (2000). The main procedure is to ray-trace through a light cone build out of a collection of dark matter density distributions at various redshifts, where the density distributions are outputted from N-body simulations. This method has been widely tested and its limitations (angular resolution, periodicity, mass resolutions, etc.) have been quantified (White & Hu, 2000; Jain et al., 2000; Hamana et al., 2002; Hamana & Mellier, 2001; Semboloni et al., 2007). It provides great reliability, e.g., to produce κ maps in the regime of observational interests. However, it is important to notice that given our box length, from the observer at $z = 0$ to the most distant lens plane at $z = 3$, $n_{box} = 24$

boxes are required. In principle, one could build only 12 fully independent light cones out of our $n_{sim} = 300$, a number far from enough to compute $\mathbf{C}^{-1}$. A common fix is using the same simulations more than once in a given light cone after random translations and rotations of the original boxs. While this trick does increase the number of light cone realisations that can be generated with a given number of N-body simulations, it introduces spurious correlations – density fields from the same simulation at different redshifts are not independent – which are hard to control safely. In fact, from the Limber approximation we know that the combination of shifting, stacking and recycling will lead to the correct convergence power spectrum since it is a projection of linear functions of the density field at each redshift. However, this does not hold anymore when considering covariance matrices. To remedy this problem, we follow an original approach described below.

We first compute the covariance matrix of the matter power spectra for each output redshift by averaging over all the n_{sim} simulations. The final convergence covariance matrix is then an appropriately weighted sum of the matter covariance matrices computed at all redshifts. More formally, this procedure can be written this way. To compute the convergence power spectra, for each output boxes at a given redshift z_{s_1}, we project the over-density field on a randomly chosen side, $\delta(\boldsymbol{x}, z_{s_1})$, Fourier transform it, $\tilde{\delta}_{2d}(\mathbf{k}, z_{s_1})$, and measure its 2-D power spectrum. After converting the comoving wavenumber k to an angular multipole $\ell = kD(z_{s_1})$, we weight the matter power spectra by the lensing kernel and project them into an angular convergence power spectrum at z_{s_1}:

$$\hat{C}_\ell^{z_{s1},z_{s2}} = \langle \tilde{\kappa}_{\ell'}^{z_{s1}} \tilde{\kappa}_{\ell'}^{z_{s2}\,\star} \rangle_{\ell \pm \Delta\ell}\,, \tag{2.12}$$

where

$$\tilde{\kappa}_\ell^{z_{s1}} = \sum_i \tilde{\delta}_{2d}(\ell = kD(z_{s_1}), z_i) W^{z_{s1}}(z_i)\,, \tag{2.13}$$

where $W^{z_{s1}}$ is the lensing kernel for each slice if the sources are distributed in redshift bin z_{s1}, and i denotes one of the n_{box} redshifts. We divide the convergence power spectrum in $n_\ell = 12$ bins, and the average in Eq. (2.12) is taken over the N_ℓ modes contributing to this l bin, $\ell - \Delta\ell < \ell' = kD(z_{s_1}) < \ell + \Delta\ell$. For clarity, in this section, we will use C_ℓ to represent C_ℓ^κ. It follows that to compute the convergence cross power spectra covariance matrix,

$$\mathrm{Cov}\,(C_\ell^{z_{s1},z_{s2}}, C_{\ell'}^{z_{s3},z_{s4}}) = \langle C_\ell^{z_{s1},z_{s2}} C_{\ell'}^{z_{s3},z_{s4}} \rangle - \langle C_\ell^{z_{s1},z_{s2}} \rangle \langle C_{\ell'}^{z_{s3},z_{s4}} \rangle \quad , \tag{2.14}$$

we need to compute both

$$\begin{aligned} \langle C_\ell^{z_{s1},z_{s2}} C_{\ell'}^{z_{s3},z_{s4}\star} \rangle_{sim} &= \langle \tilde{\kappa}_\ell^{z_{s1}} \tilde{\kappa}_\ell^{z_{s2}\,\star} (\tilde{\kappa}_{\ell'}^{z_{s3}} \tilde{\kappa}_{\ell'}^{z_{s4}\,\star})^\star \rangle_{sim} & (2.15) \\ &= \sum_{i_1,i_2,i_3,i_4} W^{z_{s1}}(z_{i_1}) W^{z_{s2}}(z_{i_2}) W^{z_{s3}}(z_{i_3}) W^{z_{s4}}(z_{i_4}) \\ &\quad \langle \tilde{\delta}_{2d}(\ell, z_{i_1}) \tilde{\delta}_{2d}^\star(\ell, z_{i_2}) \tilde{\delta}_{2d}^\star(\ell', z_{i_3}) \tilde{\delta}_{2d}(\ell', z_{i_4}) \rangle_{sim} \,, \\ & & (2.16) \end{aligned}$$

and

$$\begin{aligned} \langle C_\ell^{z_{s1},z_{s2}} \rangle_{sim} \langle C_{\ell'}^{z_{s3},z_{s4}} \rangle_{sim} &= \langle \tilde{\kappa}_\ell^{z_{s1}} \tilde{\kappa}_\ell^{z_{s2}\,\star} \rangle_{sim} \langle \tilde{\kappa}_{\ell'}^{z_{s3}} \tilde{\kappa}_{\ell'}^{z_{s4}\,\star} \rangle_{sim} & (2.17) \\ &= \left[\sum_{i_1} W^{z_{s1}}(z_{i_1}) W^{z_{s2}}(z_{i_1}) \langle \tilde{\delta}_{2d}(\ell, z_{i_1}) \tilde{\delta}_{2d}^\star(\ell, z_{i_1}) \rangle_{sim} \right] \\ &\times \left[\sum_{i_2} W^{z_{s3}}(z_{i_2}) W^{z_{s4}}(z_{i_2}) \langle \tilde{\delta}_{2d}(\ell', z_{i_2}) \tilde{\delta}_{2d}^\star(\ell', z_{i_2}) \rangle_{sim} \right] & (2.18) \end{aligned}$$

Note that one need to be careful with the complex conjugates. Whereas the expectation values for the cross-power spectra are real, the estimator of the cross power spectra are complex.

Note that if $\ell \neq \ell'$, the cross terms $\langle \tilde{\delta}_{2d}(\ell, z) \tilde{\delta}_{2d}^\star(\ell', z) \rangle_{sim}$ do not contribute. Because density distributions at different redshifts are independent, $\langle \tilde{\delta}_{2d}(\ell, z_{i_1}) \tilde{\delta}_{2d}^\star(\ell, z_{i2}) \rangle_{sim}$ are

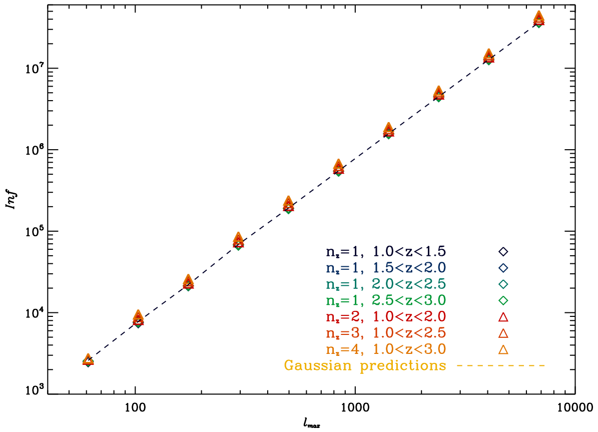

Figure 2.2 Dimensionless cumulative Fisher information for the convergence cross-power spectra defined in Eq. (2.5). We here replace the projected density field in each output box by a Gaussian field with the same power spectrum as the one measured in the N-body box. Each color corresponds to a different subset of redshift source bins of size 1, 2, 3 and 4. The dashed line corresponds to the Gaussian predictions calculated from the number of modes measured in the simulations.

also zero when $i_1 \neq i_2$. After some simple arithmetic, the calculation simplifies to

$$\begin{aligned}
\mathrm{Cov}(C_\ell^{z_{s1},z_{s2}}, C_{\ell'}^{z_{s3},z_{s4}}) \;=\; & \sum_{i_1} W^{z_{s1}}(z_{i_1}) W^{z_{s2}}(z_{i_1}) W^{z_{s3}}(z_{i_1}) W^{z_{s4}}(z_{i_1}) \\
& \Big[\langle \tilde{\delta}_{2d}(\ell, z_{i_1}) \tilde{\delta}^{\star}_{2d}(\ell, z_{i_1}) \tilde{\delta}_{2d}(\ell', z_{i_1}) \tilde{\delta}^{\star}_{2d}(\ell', z_{i_1}) \rangle_{sim} \\
& - \langle \tilde{\delta}_{2d}(\ell, z_{i_1}) \tilde{\delta}^{\star}_{2d}(\ell, z_{i_1}) \rangle_{sim} \langle \tilde{\delta}_{2d}(\ell', z_{i_1}) \tilde{\delta}^{\star}_{2d}(\ell', z_{i_1}) \rangle_{sim} \Big] \\
+ & \sum_{i_1, i_2, i_1 \neq i_2} W^{z_{s1}}(z_{i_1}) W^{z_{s2}}(z_{i_2}) W^{z_{s3}}(z_{i_1}) W^{z_{s4}}(z_{i_2}) \\
& \langle \tilde{\delta}_{2d}(\ell, z_{i_1}) \tilde{\delta}^{\star}_{2d}(\ell, z_{i_1}) \rangle_{sim} \langle \tilde{\delta}_{2d}(\ell', z_{i_2}) \tilde{\delta}^{\star}_{2d}(\ell', z_{i_2}) \rangle_{sim} \,. \qquad (2.19)
\end{aligned}$$

For the sake of clarity, we omitted the sum over ℓ and ℓ' modes within the same ℓ bin. Compared to the common ray-tracing approach, from which kappa maps and the associated angular (cross-) power spectra are build, the advantages of our method are two-folds. First, we ensure that there is no contamination due to the recycling of boxes from the same simulation. While it has been verified that this contamination is a small effect at the power spectrum level (Hamana et al., 2002), it does introduce biases when computing the four-point functions of interest to us; those biases have not been properly quantified yet. In our scheme, most of those terms are not presented although some correlations will remain since the output of the same n_{sim} simulations are used to compute the covariance matrices of the matter power spectra at each redshift. This contamination is however much smaller. Second, the rate of convergence with the number of simulation is much faster than that in the usual ray-tracing approach. Because we need n_{box} to build a full light cone ($n_{box} = 24$ in our case) in the ray-tracing method, the number of independent light cones is thus n_{sim}/n_{box} and the rate of convergence for the covariance matrix goes as $\sqrt{n_{box}/n_{sim}}$. In our case, the convergence rate for each covariance matrix at each output redshift goes as $1/\sqrt{n_{sim}}$. Furthermore, on small angular scales, because the non-linear evolution makes the (high) k modes evolution quite independent from one redshift to another, we gain an additional factor $1/\sqrt{n_{box}}$ so that the convergence rate is about $1/\sqrt{n_{sim}n_{box}}$. This effect is less important on large angular scales. Nevertheless, on all scales, the $1/\sqrt{n_{box}}$ scaling still maintains, because various k modes will appropriate to various ℓ bin so that the errors will still average down as $1/\sqrt{n_{box}}$ when computing the final covariance matrix. We thus claim an improvement of convergence rate close to a factor of n_{box} (= 24 in our case) compared to the standard ray-tracing method, *i.e.*, one order of magnitude improvement (see Fig. 2.3 and discussion in section 2.5).

For cross-checking purposes, we compare all our results with the usual ray-tracing method. We observe a satisfying agreement between the two methods. We also clearly observe the different convergence speed as n_{sim} grows.

Once the covariance matrix has been computed, we apply Eq. (2.7) to obtain an estimate of $\mathbf{C^{-1}}$. As a first check, we consider a set of Gaussian simulations in which the 2-D projected density field at each redshift is replaced by a realisation of a Gaussian field with a spectrum identical to the one resulting from N-body simulations. In that particular case, given the exact number of modes in each ℓ bin, we can exactly predict the scaling of the dimensionless cumulative Fisher information and compare it with the measurements. We consider different numbers of source bins and all the associated angular cross-power

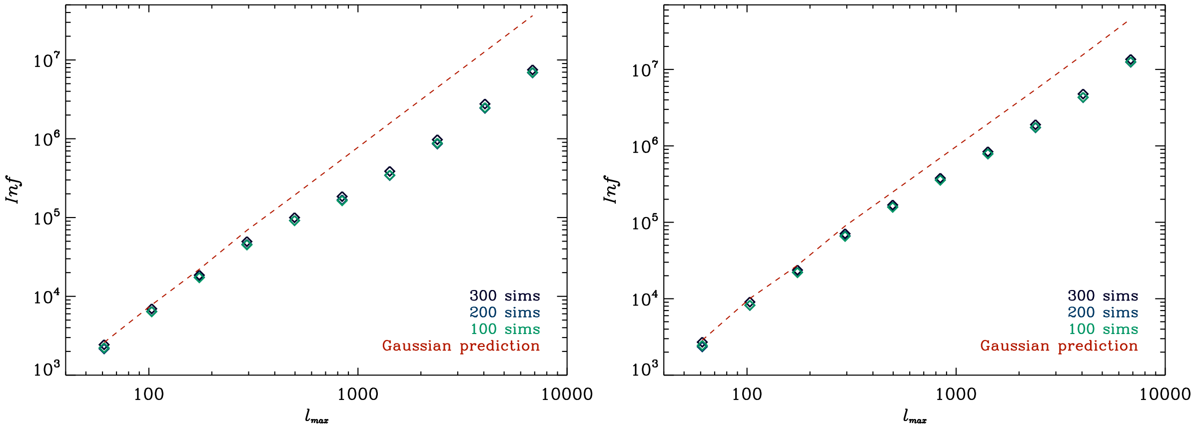

Figure 2.3 Dimensionless cumulative Fisher information for the convergence cross-power spectra defined in Eq. (2.5) when considering either one source redshift bin ($1 < z < 1.5$, the left plot) or four ($1 < z < 1.5, 1.5 < z < 2, 2 < z < 2.5, 2.5 < z < 3.0$, the right plot). Each color corresponds to a different number of simulations (100, 200 or 300) used to compute the covariance matrix. The weak scatter among those different points allows us to assess the satisfying level of convergence we can obtain with 300 simulations.

spectra. The results are displayed in Fig. 2.2, where an excellent agreement is observed between the analytical prediction (dashed line) and the measured quantities (symbols).

We then move to the genuine N-body simulations and look at the dimensionless Fisher information for either one source redshift bin (the left panel of Fig. 2.3) and four redshift source bins (the right panel of Fig. 2.3). In both plots we check the convergence of our results by varying the number of simulations and compare them with the Gaussian predictions (red dashed lines). We considered 100, 200 and 300 simulations respectively. The convergences is satisfying, and the deviance between the symbols are much smaller than the effects we are interested in, *i.e.*, the differences between the red dashed line and the symbols. As will be discussed in chapter 4 (Lu et al., 2009), an important criterion when assessing the convergence of a Monte-Carlo estimator of $\hat{\mathbf{C}}^{-1}$ is not the convergence of the diagonal of $\hat{\mathbf{C}}$ as in Takahashi et al. (2009) but of $\hat{\mathbf{C}}^{-1}$ (or its norm). The relative difference between the 200 and 300 simulations results, is consistent with our convergence rate estimate, which is around 8%.

This satisfying convergence is crucial and gives us confidence in the results displayed in Fig. 2.3. We can now interpret them. First, it is interesting that the saturation effect present in the 3-D matter power spectrum (see Fig. 2.1) also appears in the convergence power spectrum. As expected, whereas N-body results agree with Gaussian predictions on large (linear) angular scales, a departure from the Gaussian behaviour (red dashed line) appears at sub-degree scales ($\ell > 300$). Not surprisingly, for sources at $1 < z < 1.5$, this corresponds to $k \simeq 0.2h\mathrm{Mpc}^{-1}$ at $z = 0.5$ where the lensing kernel peaks, which is consistent with what is observed for the 3-D matter power spectrum (at $z = 1$) in Fig. 2.3. When higher redshift sources are included, we expect the departure from Gaussianity to be milder (non-linear evolution decreases with increasing redshift) and at smaller angular scales (higher ℓ). This corresponds to what is observed in the right panel of Fig. 2.3.

This is illustrated further in Fig. 2.4, where we consider other source redshift distributions. The agreement on large scales with the Gaussian prediction, and the shift of saturation scale to higher ℓ when the source redshift increases, are both clearly visible. The saturation effect also decreases as the source redshift increases, and when tomography is included. Note that the saturation effect is less dramatic here than in the 3-D case since the projection inherent to lensing introduces an extra Gaussianization.

2.5 Discussion

Now we have highlighted the dimensionless cumulative Fisher information for some perfect idealized survey which has the same angular area as our simulations, next we will discuss its implications for current and future optical surveys. In particular, we will quantify the consequences of the Fisher information saturation effects discussed above on the cosmological parameters. We focus on dark energy which is parametrized by an evolving equation of state $w(a) = w_0 + (1 - a)w_a$. We use the dark energy figure of merit (DE FoM) as our main statistics (Albrecht et al., 2006), which is defined as the area

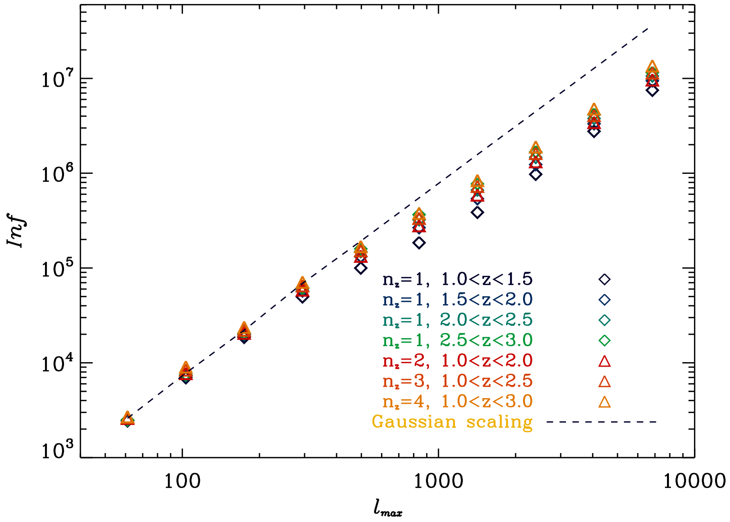

Figure 2.4 Dimensionless cumulative Fisher information for the convergence cross-power spectra defined in Eq. (2.5). Each color correspond to a different sub-set of source redshift band of size $n_z = 1$, 2, 3 or 4. The dashed color line corresponds to the Gaussian prediction for the same number of modes. Similar to the 3-D case, the effect of non-Gaussianity is clearly visible as a drop in the dimensionless Fisher information content in the slightly non-linear regime. The scaling becomes quasi-Gaussian again in the fully non-linear regime. As expected, since non-linearities increase with redshift, the effect is more severe if the source redshift is lower.

of the 95% contour ellipse in the $w_0 - w_a$ plane. Following the definition of the Fisher matrix in Eq. (2.1), if we consider the following observable

$$\mathcal{P}_{a=i(i-1)/2+j,\ell} = \tilde{n}_{z_i}\tilde{n}_{z_j}C_\ell^{\kappa z_i z_j}\,, \quad (i \geq j) \tag{2.20}$$

and a set of six cosmological parameters λ_μ as mentioned in section 2.2, the Fisher information matrix is written as

$$\mathbf{F}_{\mu\nu} = \sum_{\ell=2}^{\ell_{\max}}\sum_{ab}\frac{\partial\mathcal{P}_a}{\partial\lambda_\mu}[\mathbf{C}^{-1}]_{ab}\frac{\partial\mathcal{P}_b}{\partial\lambda_\nu}\,. \tag{2.21}$$

The standard deviance of λ_μ, $\sigma(\lambda_\mu) = \sqrt{\mathbf{F}^{-1}_{\mu\mu}}$.

We will make use of the 300 simulations introduced in section 2.4 and customized for the diffuse 21-cm lensing in chapter 4 to discuss optical cosmic shear. We do so by re-weighting the previous results obtained using the wide uniform redshift bins with the

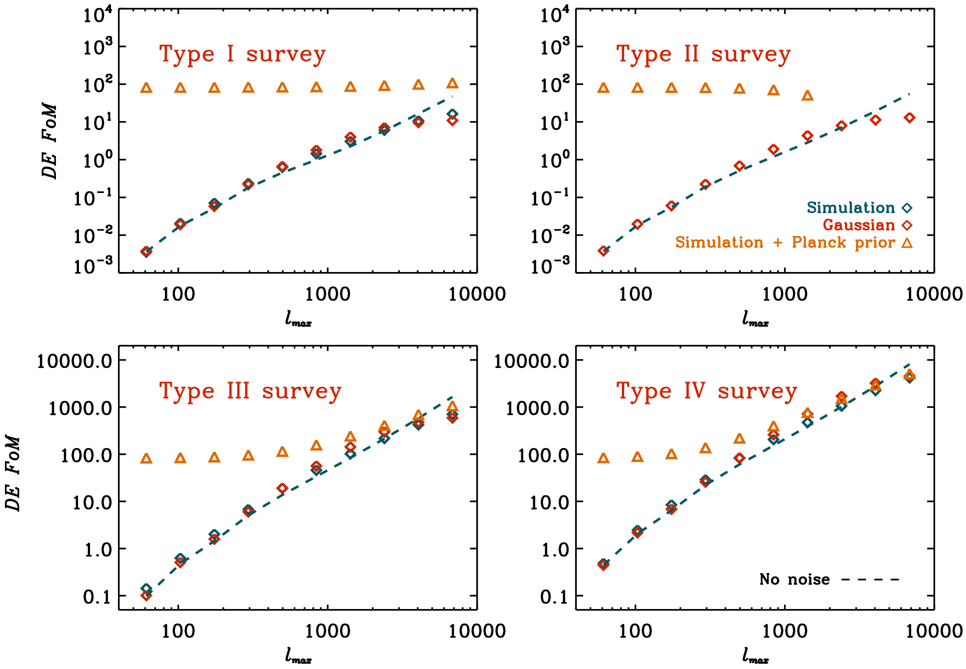

Figure 2.5 Dark energy figure of merit as a function of scales for 4 different surveys whose parameters are given in table 2.5. The blue diamonds correspond to the predictions using non-Gaussian covariance matrices and shot noise contributions. The red diamonds correspond to the Gaussian approximations to the covariance matrices with shot noises. The orange diamonds correspond to the non-Gaussian case with the prior expected from the Planck satellite[1]. The blue and red dashed curves correspond respectively to the non-Gaussian and Gaussian cases, without shot noise. Whereas we can see that the inclusion of non-Gaussian error bars is important for perfect noiseless experiments (blue dashed curves) compared to the perfect noiseless Gaussian errors (red dashed curves), it is less critical when adding the shot noise coming from the dispersion of intrinsic ellipticities (comparison between blue and red diamonds).

Survey	I	II	III	IV
Area (deg^2)	200	200	5 000	20 000
n_{gal} (arcmin^2)	30	30	50	100
σ_e	0.25	0.25	0.25	0.25
number of redshift bins	1	3	3	4

Table 2.1 Optical surveys considered. They roughly correspond to a current survey like CFHTLS (type I,II), the soon on-line DES survey (type III), and a wide and deep space survey like Euclid or JDEM (type IV), respectively.

relative weights appropriate coming from a realistic optical galaxy number density

$$\begin{aligned} n(z) \quad &\propto \quad z^{\alpha} e^{-(z/z_0)^{\beta}} \,, \\ \text{with} \quad \alpha = 2 \quad &\text{and} \quad \beta = 1.5 \,. \end{aligned} \tag{2.22}$$

We also rescale the signal covariance matrix by the survey area considered. In this thesis, we will ignore the uncertainties in the number density, though they are important (Benjamin et al., 2007). We will consider four surveys whose parameters are given in table 2.5. They roughly correspond to a current survey like CFHTLS [2], the soon on-line DES [3] survey, and a wide and deep space survey like Euclid [4] or JDEM, respectively. We will normalize $n(z)$ so that the galaxy density matches the one given on the second line of table 2.5 and we will consider the shot noise coming from the intrinsic ellipticities of objects. Note that our predictions here are somewhat inaccurate as we are re-weighting our 21-cm sources simulations which will be used in chapter 4, and our ideal survey is also non-optimal since we consider only 4 redshift bins to perform tomography, which has been shown to be slightly sub-optimal (Ma et al., 2006). However, our treatment is accurate enough to discuss the effects of non-Gaussianities that are the focus of our

---

[2] http://www.cfht.hawaii.edu/Science/CFHLS/

[3] https://www.darkenergysurvey.org/

[4] http://sci.esa.int/science-e/www/area/index.cfm?fareaid=102

study.

As a reference point, we will also compute the Fisher matrix with a Gaussian approximation to $\mathbf{C}^{-1}$. Doing so, we follow the formalism in Ma et al. (2006). In the Gaussian case, assuming a shot noise level and a simple Gaussian sample variance, the covariance matrix is defined as

$$\mathbf{C}_{ab\ \ell} = \tilde{n}_i \tilde{n}_j \tilde{n}_k \tilde{n}_l \left(C_\ell^{\mathrm{tot\ ik}} C_\ell^{\mathrm{tot\ jl}} + C_\ell^{\mathrm{tot\ il}} C_\ell^{\mathrm{tot\ jk}} \right) , \tag{2.23}$$

where $a = i(i-1)/2 + j$, $b = k(k-1)/2 + l$, and the total power spectrum is

$$C_\ell^{\mathrm{tot\ ij}} = C_\ell^{\kappa\ ij} + \delta_{ij} \frac{{\sigma_e}^2}{\tilde{n}_i} , \tag{2.24}$$

where σ_e is the rms shear error per galaxy contributed by intrinsic ellipticity and measurement errors.

Fig. 2.5 compares the Gaussian and non-Gaussian cases for the 4 different surveys whose parameters are defined in table 2.5. From the Type I survey to the Type IV survey, we simultaneously increase the number of galaxies, the depth, the number of source redshift bins, and the survey area, which results in an increase in the FoM. As we previously did for the dimensionless Fisher information content, we will now study the evolution of the FoM as we increase the number of modes, from the linear regime to the non-linear regime. This is a proxy to quantify the cosmological interpretation in this survey as we increase angular sensitivity. The red dashed lines correspond to the noise free Gaussian approximations while the blue dashed lines correspond to the non-Gaussian noise free evaluations. As we can see, when comparing these two type of curves, the saturation effects discovered earlier in the cumulative Fisher information translate naturally into the evolution of the FoM with ℓ_{max}. In the case of a noiseless survey, the difference in the FoM at high ℓ can be as high as a factor of 4, even when we consider only four tomographic bins. This effect is thus important and in stark contrast to the scaling of the FoM with ℓ_{max} usually assumed in the literature (Amara & Réfrégier, 2007). However, when introducing realistic levels of noises, the effect is somewhat mitigated as

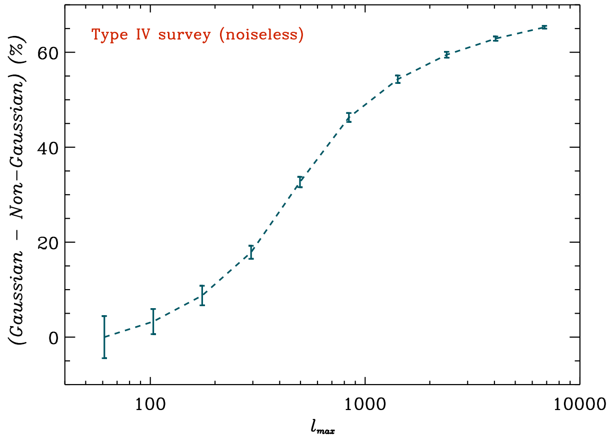

Figure 2.6 Bootstrap errors on the relative difference between the noiseless FoM curves for the Type IV survey. We are thus plotting the relative errors between the red and blue dashed curves in the lower right panel of Fig. 2.5.

is visible when comparing the red diamonds (Gaussian approximation, with noise) to the blue diamonds (non-Gaussian covariance matrix, with noise). It is still non-negligible since the ratio between the Gaussian and non-Gaussian cases at higher ℓ_{max} becomes close to 1.6. This point constitutes the key result from our study. While potentially very damaging to the ideal performances of weak lensing surveys, the effect of non-Gaussianity is tampered by the estimated level of noises expected for current and future surveys. The inclusion of the Planck prior does not affect those conclusions at high ℓ.

Note also that despite the fact that our plot hint at the ability to measure the convergence power spectra up to $\ell < 10000$, in practice, it will most likely be limited by theoretical uncertainties at $\ell \leq 3000$, at least by our inability to model the details of the baryon physics (White, 2004; Zhan & Knox, 2004; Rudd et al., 2008).

Consistent conclusions were reached in the halo model based analytical approach followed in Cooray & Hu (2001); Takada & Jain (2008). We will discuss the latter, which is the most recent one, and refer to its discussion of the earlier results of Cooray & Hu (2001). Our results seem consistent with their evaluation of the S/N (somewhat comparable to our dimensionless Fisher information). So does the effects on parameters when

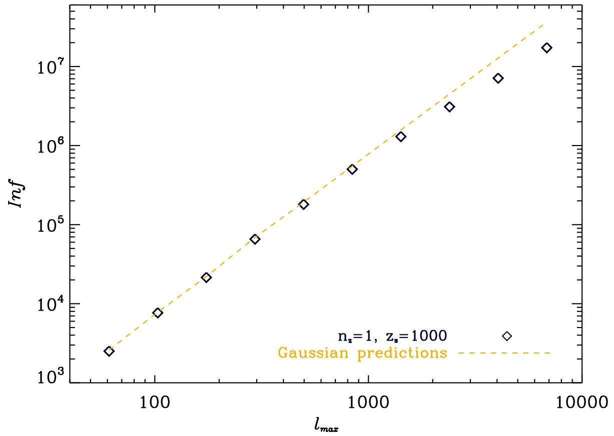

Figure 2.7 Dimensionless cumulative Fisher information content as a function of the maximum multipole for a source plane at $z = 1000$. A mild saturation effect is still visible. Since the CMB lensing is sensitive mostly to the growth of structures at $z \simeq 2.5$, the saturation effect is milder and shifted to smaller angular scales (higher ℓ_{max}).

considering realistic surveys, although our factor of 1.6 for FoM seems a bit higher. Note that the comparison is not immediate since the set of parameters we consider is slightly different. Whereas we consider a flat Universe with 6 parameters including an evolving dark energy equation of state, they consider non-flat models and allow the spectral index to run. We do not translate our evaluation of the cross-power spectrum covariance matrices to real space statistics, and a comparison with real space evaluation of this effect (Semboloni et al., 2007), can be found at Takada & Jain (2008). A more detailed comparison between analytical estimates and numerical works would be instructive and has to be performed, but we leave it for future work. We also do not investigate the beat-coupling effect (Hamilton et al., 2006; Takada & Jain, 2008). As we understand, it denotes the extra mode-coupling induced by the finiteness of the observed volume, and therefore such an effect would be non-existent if the full sky was considered. Since the exact form of this coupling will depend on the exact mask of any given survey, we decide to ignore this effect here. In practice, as it is usually performed, for example, while analysing CMB data, we would start from a mask description in real space and propagate the induced mode-coupling in Fourier space throughout all our calculations. With their own prescription, Takada & Jain (2008) found that it does not affect qualitatively the effect of the non-linear growth of structures.

The comparison with the numerical work of Hu & White (2001) is also not obvious. First, a different cosmology with a higher σ_8 is considered there, which should enhance the non-linear effect. Second, they consider 200 simulations and use a standard tiling technique in ray-tracing approach that should give rise to an accuracy of at most 7% on the errors. All these reasons make a direct comparison a bit difficult.

To quantify further the error on the errors issue, we evaluate the errors of the FoM for the Type IV survey introduced before using the bootstrap method again. We display the results in Fig. 2.6 where we plotted the relative difference between the Gaussian and non-Gaussian FoM as a function of ℓ_{max}. Due to the strong convergence of our covariance matrix estimation technique, we find the error on the errors to be around 25% on large scales and sub-percent on the smaller scales we consider. This fact certainly constitutes an improvement over previous results in the literature and is consistent with our estimate of a $1/\sqrt{n_{sim}n_{box}}$ convergence rate. If we use the scaling of $12/\sqrt{n_{sim}}$ measured by Takahashi et al. (2009), we claim an overall uncertainty due to the limited number of simulations to be around 14%.

Besides the optical observations of weak lensing, the cosmic microwave background (CMB) constitutes another source plane where lensing can be observed (see Lewis & Challinor (2006) for a review). As both the resolution and the sensitivity of detectors are improved, the CMB lensing can now be measured and it defines the next frontier for the CMB temperature and polarization measurement (Smith et al., 2008). Using cross correlation between WMAP data and other tracers of large-scale structures to increase the S/N, a detection of gravitational lensing in the CMB temperature has been achieved with marginal significance, i.e. around 2.4 σ (Hirata et al., 2004; Smith et al., 2007; Hirata et al., 2008). A direct detection in temperature is expected to be achieved soon with high significance thanks to the on-going high angular resolution temperature surveys (ACT[5], SPT[6], Planck). It is thus interesting to evaluate the effects of the non-linear growth of structures on the CMB lensing signal. To do so, we compute the cumulative Fisher information of the amplitude of convergence power spectrum as a function of the maximum angular scale ℓ_{max}, when considering a single redshift source plane at $z = 1000$. The result is displayed in Fig. 2.7. Interestingly, the saturation effect is still visible, and not surprisingly, since the CMB lensing kernel peaks around z$\simeq$ 2.5, the effect is smaller and shifted to smaller scales. However, in practice, given that the CMB reconstruction is most likely going to be limited by secondary anisotropies (kinetic SZ in particular and patchy reionization) around a few $\ell \simeq 3000$, it is unlikely that this effect of non-linear growth will be critical.

To conclude, non-Gaussian effects are potentially very important for weak lensing surveys and might significantly alter the forecasts done so far. When considering realistic noise estimate for the coming optical surveys, the impact of non-Gaussian error bars is much milder.

2.6 Summary

We quantify the Fisher information content of lensing convergence (cross-) power spectrum in cosmic shear survey as a function of resolution. In particular, we investigate whether the Fisher information saturates above a given angular scale. To do so, we rely heavily on N-body simulations in order to accurately probe the non-linear regime. We find that while even in a perfect survey, there is no clear saturation scale, that non-linear growth induced non-Gaussianity can lead to a factor of 4 reduction for the dark energy figure of merit. This effect is however mitigated by realistic levels of shot noise and we find that for future surveys, the effect is about a factor of 1.5. To do so, we develop

---

[5] http://www.physics.princeton.edu/act/about.html

[6] http://pole.uchicago.edu/spt/index.php

a new scheme to compute the relevant covariance matrix of lensing convergence power spectrum. It leads us to claim an order of magnitude improvement in accuracy using only twice as many simulations as were previously used. Finally, we evaluate the errors on the errors using bootstrap methods.

Chapter 3

Lensing of 21-cm line at epoch of reionization

3.1 Overview

Driven by various surveys, in this chapter and chapter 3, we will discuss the lensing of the 21-cm line sources. 21-cm emission is similar to CMB: both are diffuse backgrounds. It is natural to apply the techniques used in CMB lensing. Hu & Okamoto (2002) expand the CMB lensing field in terms of the gravitational potential (or deflection angles), and construct a trispectrum based quadratic estimator of potential with maximum S/N. However, unlike CMB, the 21-cm background has a 3-D distribution and is intrinsically non-Gaussian. A fully 3-D analysis is explored in Zahn & Zaldarriaga (2006), where they generalize the 2-D quadratic estimator of CMB lensing (Hu & Okamoto, 2002) to the 3-D Optimal Quadratic Deflection Estimator (OQDE).

A local estimator was proposed in Pen (2004), which assumed a power-law density power spectrum. In this chapter, we will design localized estimators for the lensing fields under the Gaussian assumption, and apply the derived reconstruction technique to Gaussian and non-Gaussian sources. The influence of non-Gaussianity can be measured by comparing the numerical results between the Gaussian sources and non-Gaussian sources.

Quadratic lensing reconstruction is a two-point function of the lensed brightness temperature field of the 21-cm emission. In this chapter, 3-D quadratic estimators are constructed for the convergence (κ), as well as the shear (γ). Our method recovers the κ and γ directly instead of gravitational potential or deflection angles. Our estimators have in principle the same form as the OQDE, consisting of the covariance of two filtered temperature maps. The OQDE reconstructs the deflection angle, while our estimators reconstruct the kappa and shear fields. Our filtering process can be written as a convolution of the observed fields. As presented in appendix and section 4, our combined estimator is unbiased, and equally optimal as the OQDE for Gaussian sources. Our estimator has better performance for non-Gaussian sources, and recovers three extra (constant) modes for finite fields.

Other authors also developed reconstruction methods from alternative approaches. Metcalf & White (2007) give an estimator for shear. They choose 2-D slices at different redshifts, and then treat these slices as independent source samples for the same lensing structure. They neglect the information between these slices, so the efficiency of reconstruction depends on the width of the slices. Cooray (2004) expands the lensed field to higher order of the gravitational potential, and investigates the resulting corrections to the lensed power spectrum.

This chapter is organized as follows: The basic framework of lensing and the reconstruction method is introduced in section 3.2. The numerical methods are presented in section 3.3. The results are discussed in section 3.4. We conclude in section 3.5.

3.2 Lensing and reconstruction

Photons are deflected by the gravitational pull of clumpy matter when they propagate from the source to the observer. This effect can be used to map the mass distribution if we can measure the distortion of an image. In this section, we will first develop an optimal quadratic estimator using a maximum likelihood method. The reconstruction depends on the power spectrum of the source. The noise and normalization of the reconstruction are calculated in appendix A.

Maximum likelihood estimator of κ

We now derive the quantitative window functions for 21-cm lensing reconstruction. Due to their similarity, it is helpful to quickly review the reconstruction in CMB lensing again: The early work by Zaldarriaga & Seljak (1999) used the quadratic combination of the derivatives of the CMB field to reconstruct the lens distribution, originally also using a κ and γ formulation. Since the CMB has an intrinsic Gaussian distribution, the optimal quadratic estimator (Hu, 2001) can also be applied to lensing reconstruction with CMB polarization (Hu & Okamoto, 2002). Zahn & Zaldarriaga (2006) generalized the optimal quadratic estimator of CMB lensing to 21-cm lensing.

We will construct estimators for κ and γ with the 21-cm brightness temperature fields, starting from a maximum likelihood derivation. We will show that the OQDE and our approach are the same if the sources are Gaussian, however the problem is simplified in a intuitive way by using the limit that κ and γ vary slowly in small scales. We will show later the optimal window functions have the same form when κ and γ vary rapidly.

The magnification is

$$\mu = \frac{1}{(1-\kappa)^2 - \gamma^2} \sim 1 + 2\kappa \ . \tag{3.1}$$

The last approximation is valid since both κ and γ are much smaller than 1 in the weak lensing regime.

We use Bayesian statistics and assume the prior distribution of parameter κ to be flat. For a M pixel map on the sky, the posterior likelihood function of the source field has a Gaussian distribution, and can be written as

$$\mathcal{P}(\tilde{T}(\boldsymbol{k})) = (2\pi)^{-M/2}\mathrm{det}(\mathbf{C}_{\tilde{T}\tilde{T}})^{-\frac{1}{2}}e^{-\frac{1}{2}\tilde{T}^{\dagger}\mathbf{C}_{\tilde{T}\tilde{T}}{}^{-1}\tilde{T}} \ . \tag{3.2}$$

Here $\tilde{T} = \tilde{T}_b + n$ is the brightness temperature of the diffusive 21-cm emission lensed by the large-scale structure plus measurement noise. To simplify the algebra, we use the negative logarithm $\mathcal{L}$ of the likelihood function in our calculation,

$$\mathcal{L} = -\ln\mathcal{P} = \frac{1}{2}\tilde{T}^{\dagger}\mathbf{C}_{\tilde{T}\tilde{T}}^{-1}\tilde{T} + \frac{1}{2}\ln\det\mathbf{C}_{\tilde{T}\tilde{T}}. \tag{3.3}$$

Here $\tilde{T}$ is the 3-D discrete Fourier transform of measured temperature. $\mathbf{C}_{\tilde{T}\tilde{T}} = \mathbf{C}_{\mathrm{S}} + \mathbf{C}_{\mathrm{N}}$ is the covariance matrix, and the signal contribution $\mathbf{C}_{\mathrm{S}}$ and noise contribution $\mathbf{C}_{\mathrm{N}}$ are both diagonal in Fourier space and uncorrelated to each other. In the continuum limit, the likelihood function can be written as

$$\mathcal{L} = \frac{1}{4\pi^2}[\int \mathrm{d}^3k \ln\tilde{P}^{\mathrm{tot}}(\boldsymbol{k}) + \int \mathrm{d}^3k\frac{|\tilde{T}(\boldsymbol{k})|^2}{\tilde{P}^{\mathrm{tot}}(\boldsymbol{k})}] \ . \tag{3.4}$$

We use $\tilde{P}^{\mathrm{tot}} = \tilde{P}(\boldsymbol{k}) + P^{\mathrm{N}}(\boldsymbol{k})$ to represent the signal plus noise power spectrum in the following text, where $\tilde{P}(\boldsymbol{k})$ is 3-D power spectrum of the distorted 21-cm field, and $P^{\mathrm{N}}(\boldsymbol{k})$ is the noise power spectrum. The dimensionless power spectrum of the 3-D 21-cm gas can be written as

$$\Delta^2(k) = \frac{k^3}{2\pi^2}P(k) \ , \tag{3.5}$$

where $k = |\boldsymbol{k}|$ since the gas is statistically isotropic.

The geometry of the 21-cm field will be changed by lensing:

$$\begin{aligned}\tilde{T}_b(\boldsymbol{k}_\perp, k_\parallel) &= \int \mathrm{d}^3x\tilde{T}_b(\boldsymbol{x})e^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} = \int \mathrm{d}^2x_\perp\int \mathrm{d}x_\parallel T_b((1-\kappa)\boldsymbol{x}_\perp, x_\parallel)e^{-\mathrm{i}(\boldsymbol{k}_\perp\cdot\boldsymbol{x}_\perp + k_\parallel x_\parallel)} \\ &= \frac{1}{(1-\kappa)^2}T_b(\frac{\boldsymbol{k}_\perp}{1-\kappa}, k_\parallel) \ ,\end{aligned} \tag{3.6}$$

where '$\perp$' and '$\|$' means the perpendicular and parallel direction to the line-of-sight respectively. We ignore the contribution of shear first. Then the length scale is magnified on the transverse plane by a factor κ. Isotropy is broken in 3-D but is still conserved on the 2-D cross section. The statistical properties of the 21-cm field will be changed by lensing, i.e., the power spectrum will also change:

$$\langle \tilde{T}_b^*(\boldsymbol{k}_\perp, k_\parallel)\tilde{T}_b(\boldsymbol{k}'_\perp, k'_\parallel)\rangle = (2\pi)^2\delta^{\rm 2D}(\boldsymbol{k}_\perp - \boldsymbol{k}'_\perp)(2\pi)\delta^{\rm D}(k_\parallel - k'_\parallel)\tilde{P}(\boldsymbol{k}_\perp, k_\parallel)\ . \tag{3.7}$$

The delta function has the property:

$$\delta^{\rm 2D}(\frac{\boldsymbol{k}_\perp}{1-\kappa} - \frac{\boldsymbol{k}'_\perp}{1-\kappa}) = (1-\kappa)^2\delta^{\rm 2D}(\boldsymbol{k}_\perp - \boldsymbol{k}'_\perp)\ . \tag{3.8}$$

Therefore the relationship between the unlensed and lensed power spectrum is

$$\tilde{P}(\boldsymbol{k}_\perp, k_\parallel) = \frac{1}{(1-\kappa)^2}P(\frac{\boldsymbol{k}_\perp}{1-\kappa}, k_\parallel) = \frac{1}{(1-\kappa)^2}P(\sqrt{\frac{k_\perp^2}{(1-\kappa)^2} + k_\parallel^2})\,. \tag{3.9}$$

The second equivalence is due to the statistical isotropy of the unlensed power spectrum.

The Taylor expansion of the lensed power spectrum is

$$\tilde{P}^{\rm tot} = P^{\rm tot} + \left.\frac{\partial \tilde{P}^{\rm tot}}{\partial \kappa}\right|_{\kappa=0}\kappa + \frac{1}{2}\left.\frac{\partial^2 \tilde{P}^{\rm tot}}{\partial \kappa^2}\right|_{\kappa=\kappa_\star}\kappa^2\,, \tag{3.10}$$

where $\kappa_\star$ in the residual term lies between 0 and κ. The first order derivative at $\kappa = 0$ is $2P + \Delta P$, where $\Delta P = P'k(k_\perp^2/k^2)$, and $P'(k) = {\rm d}P(k)/{\rm d}k$. In this chapter, we will only consider the first order perturbation. Now we can verify the accuracy of the approximation $\Delta\kappa_{\rm Taylor}/\kappa$, which is the ratio of the residual term to the second term in Eq. (3.10). If we write the first order derivative of power spectrum to kappa as $G(\kappa) = \partial\tilde{P}^{\rm tot}/\partial\kappa$, and the second order derivative as $R(\kappa) = \partial^2\tilde{P}^{\rm tot}/\partial\kappa^2$, there is

$$\frac{\Delta\kappa_{\rm Taylor}}{\kappa} = \frac{1}{2}\frac{R(\kappa=\kappa_\star)}{R(\kappa=0)}\frac{R(\kappa=0)}{G(\kappa=0)}\kappa\,. \tag{3.11}$$

In most cases, power spectra have an approximately power-law shape $P = P_0k^n$. For dark matter-like power spectrum, $\Delta^2(k) \propto k$, there is $n = -2$. $R(\kappa=0)/G(\kappa=0) = 3 - 4(k_\perp/k)^2 \in [-1, 3]$. $R(\kappa = \kappa_\star)/R(\kappa=0) = 1/(1-\kappa)^4$ when $k_\perp/k = 0$; $R(\kappa = \kappa_\star)/R(\kappa=0) \to 22/(1-\kappa)^2 - 21$ when $k_\perp/k \to 1$. κ to the epoch of reionization is about at the 5% level, therefore

$$\left|\frac{\Delta\kappa_{\rm Taylor}}{\kappa}\right| \leq 1.8\kappa\,. \tag{3.12}$$

The precision of a first order κ reconstruction would not be better than the magnitude of κ itself, i.e., the quadratic reconstruction would not be better than a few percent level. We will calculate to first order accuracy in κ throughout this chapter, however we need to keep in mind that the first order perturbation approximation will contribute several percent error in our calculation.

The maximum likelihood condition requires

$$\frac{\delta \mathcal{L}}{\delta \kappa} \approx \frac{1}{2} L^3 \int \frac{\mathrm{d}^3 k}{(2\pi)^3} \frac{(\tilde{P}^{\mathrm{tot}} - |\tilde{T}|^2 L^{-3})}{\tilde{P}^{\mathrm{tot}2}} \frac{\delta \tilde{P}}{\delta \kappa} = 0 \,. \tag{3.13}$$

Since we calculate to first order accuracy in κ, a further simplification is $1/\tilde{P}^{\mathrm{tot}2} \approx [1 - 2\kappa(2P + \Delta P)/P]/P^2$. The first order solution is

$$E_\kappa = \int \frac{\mathrm{d}^3 k}{(2\pi)^3} (|\tilde{T}|^2 L^{-3}) \mathcal{F}^\kappa(\boldsymbol{k}) - V_\kappa \,. \tag{3.14}$$

To simplify the problem, we assume the source is a cube with physical length L in each dimension. The offset constant $V_\kappa = \langle \sigma^2 \rangle = \int \mathrm{d}^3 k/(2\pi)^3 P^{\mathrm{tot}}(\boldsymbol{k}) \mathcal{F}^\kappa(\boldsymbol{k})$, and the optimal filter $\mathcal{F}^\kappa$ is

$$\mathcal{F}^\kappa(\boldsymbol{k}) = \frac{2P(\boldsymbol{k}) + \Delta P(\boldsymbol{k})}{P^{\mathrm{tot}^2}(\boldsymbol{k}) Q_\kappa} \,, \tag{3.15}$$

with $Q_\kappa = \int \mathrm{d}^3 k/(2\pi)^3 (2P + \Delta P)(\boldsymbol{k}) \mathcal{F}^\kappa(\boldsymbol{k})$.

From Parseval's theorem, we can rewrite Eq. (3.14) in the form of a convolution of the density field and a window function in real space

$$\int \frac{\mathrm{d}^3 k}{(2\pi)^3} \tilde{T}^*(\boldsymbol{k}) \tilde{T}(\boldsymbol{k}) \mathcal{F}^\kappa(\boldsymbol{k}) = \int \mathrm{d}^3 x \tilde{T}^\kappa_{\mathrm{w}_1}(\boldsymbol{x}) \tilde{T}^\kappa_{\mathrm{w}_2}(\boldsymbol{x}) = L^2 \int \mathrm{d}x_\parallel \tilde{T}^\kappa_{\mathrm{w}_1}(\boldsymbol{x}_\perp, x_\parallel) \tilde{T}^\kappa_{\mathrm{w}_2}(\boldsymbol{x}_\perp, x_\parallel) \,. \tag{3.16}$$

In Eq. (3.16) the two window functions are the decomposition of the optimal filter $W_1^\kappa(\boldsymbol{k}) W_2^\kappa(\boldsymbol{k}) = \mathcal{F}^\kappa(\boldsymbol{k})$. The last '=' in Eq. (3.16) holds when κ is constant. One can choose $W_1^\kappa(\boldsymbol{k}) = W_2^\kappa(\boldsymbol{k}) = \sqrt{\mathcal{F}_\kappa}$. If $\mathcal{F}_\kappa < 0$, we choose $W_1^\kappa = -W_2^\kappa = \sqrt{|\mathcal{F}_\kappa|}$. The convergence field is equivalent to the covariance of the measured maps smoothed by two windows. In the slowly spatially varying κ limit, all decomposition into two windows are equivalent. As we will show later, the shear construction can also be represented in the form of the covariance of two filtered temperature maps. These maps will have symmetric Probability Density Functions (PDF), which can reduce the non-Gaussianity of the maps so that a better S/N level can be achieved, when the shear window functions are chosen properly. The last two steps in Eq. (3.16) assumes the fluctuation of the convergence field is slow compared to the filter. Then we can apply the estimator to each beam in the map:

$$E_\kappa(\boldsymbol{x}_\perp) = L^{-1} \int \mathrm{d}x_\parallel \tilde{T}^\kappa_{\mathrm{w}_1}(\boldsymbol{x}) \tilde{T}^\kappa_{\mathrm{w}_2}(\boldsymbol{x}) - V_\kappa \,, \tag{3.17}$$

where $\tilde{T}^\kappa_{\mathrm{w}_1}$ and $\tilde{T}^\kappa_{\mathrm{w}_2}$ are the convolution of $\tilde{T}$ and window function $W^\kappa_1(\boldsymbol{x})$ and $W^\kappa_2(\boldsymbol{x})$ respectively, which are the real space version of $W^\kappa_1(\boldsymbol{k})$ and $W^\kappa_2(\boldsymbol{k})$. The reconstruction of κ is dominated by the gradient of the power spectrum $\mathrm{d}\ln\Delta^2/\mathrm{d}\ln k$, which follows from the expression of our estimator in Eq. (3.14).

We can generalize the estimator to a spatially varying lensing field. In the appendix we show

$$\int \mathrm{d}^2 x'_\perp \kappa(\boldsymbol{x}'_\perp) b_\kappa(\boldsymbol{x}_\perp - \boldsymbol{x}'_\perp) = \langle E_\kappa(\boldsymbol{x}_\perp) \rangle \;. \tag{3.18}$$

Equivalently, for smaller scales, we will need to normalize the reconstructed lensing field by a scale dependent factor in Fourier space, which is calculated in the appendix.

$$\hat{\kappa}(\boldsymbol{\ell}) = b^{-1}_\kappa(\boldsymbol{\ell}) E_\kappa(\boldsymbol{\ell}) = \kappa(\boldsymbol{\ell}) + n(\boldsymbol{\ell}) \;, \tag{3.19}$$

where $\boldsymbol{\ell} = \boldsymbol{k}_\perp \chi(z_\mathrm{s})$, and z_s is the redshift of the source. Here $b_\kappa(\ell)$ is the normalization factor ($\lim_{\ell\to 0} b_\kappa(\ell) = 1$), and $n(\ell)$ is the noise, since different Fourier modes are independent. They do not depend on direction because variables related to κ are isotropic on the transverse plane. In the appendix, we show that the normalization factor is unity at small l when Q_κ has the form as $Q_\kappa = \int \mathrm{d}^3k/(2\pi)^3 (2P + \Delta P)(\boldsymbol{k}) \mathcal{F}^\kappa(\boldsymbol{k})$.

Amblard et al. (2004) have pointed out that the reconstructed kappa could be biased due to the non-Gaussianity of lenses. One possible reason is that there is only one CMB sky behind each patch of lensing field. For 21-cm, there are many source planes, and we expect there is no such bias effect by averaging over these planes.

Estimator of shear

When shear is taken into account, not only the scale but the directions of the coordinates are changed. We will start the derivation from the constant shear case. In analogy to κ, the optimal window function will be the same when γ is spatially variable.

$$\begin{aligned}\tilde{T}_b(\boldsymbol{k}_\perp, k_\parallel) &= \int \mathrm{d}^3x \tilde{T}_b(\boldsymbol{x}) e^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} = \int \mathrm{d}^2x_\perp \int \mathrm{d}x_\parallel T_b(\mathbf{J}\boldsymbol{x}_\perp, x_\parallel) e^{-\mathrm{i}(\boldsymbol{k}_\perp\cdot\boldsymbol{x}_\perp + k_\parallel x_\parallel)} \\ &= |\mathbf{J}|^{-1} \int \mathrm{d}^2x'_\perp \int \mathrm{d}x_\parallel T_b(\boldsymbol{x}'_\perp, x_\parallel) e^{-\mathrm{i}(\boldsymbol{k}'_\perp\cdot\boldsymbol{x}'_\perp + k_\parallel x_\parallel)} = |\mathbf{J}|^{-1} T_b(\mathbf{J}^{-1}\boldsymbol{k}_\perp, k_\parallel)\end{aligned} \tag{3.20}$$

here $\mathrm{d}^2x'_\perp = |\mathbf{J}|\mathrm{d}^2x_\perp, \boldsymbol{k}'_\perp = \mathbf{J}^{-1}\boldsymbol{k}_\perp$. Now the symmetry is broken even on the transverse plane due to the anisotropic distortion caused by the shear.

Since $\delta^{\mathrm{2D}}(\mathbf{J}^{-1}\boldsymbol{k}) = |\mathbf{J}|\delta^{\mathrm{2D}}(\boldsymbol{k})$, Eq. (3.7) implies

$$\tilde{P}(\boldsymbol{k}_\perp, k_\parallel) = |\mathbf{J}|^{-1} P(\mathbf{J}^{-1}\boldsymbol{k}_\perp, k_\parallel) \approx (1+2\kappa)[P(k) + \Delta P(\boldsymbol{k})(\kappa + \gamma_1 \cos 2\theta_{\boldsymbol{k}_\perp} + \gamma_2 \sin 2\theta_{\boldsymbol{k}_\perp})] \,, \tag{3.21}$$

where $\theta_{\boldsymbol{k}_\perp}$ is the angle between $\boldsymbol{k}_\perp$ and the transverse coordinate. Note in the calculation of κ, we have ignored the contribution from shear. The reason is that the overall contribution of shear would be zero to first order magnitude of κ due to the angular dependence of shear on $\theta_{\boldsymbol{k}_\perp}$.

Maximum likelihood requires $\delta\mathcal{L}/\delta\gamma_1 = 0$, and $\delta\mathcal{L}/\delta\gamma_2 = 0$. The maximum likelihood shear estimators can be written as a tensor $\mathbf{E}_\gamma$:

$$E_{\gamma \mathrm{ij}} = L^{-1} \int \mathrm{d}x_\parallel \tilde{T}^\gamma_{\mathrm{w_i}} \tilde{T}^\gamma_{\mathrm{w_j}} \ , \tag{3.22}$$

where $\tilde{T}^\gamma_{\mathrm{w_i}}$ is convolution of the temperature field with W^γ_{i}, and $W^\gamma_{\mathrm{i}}(\boldsymbol{k}) = (2\Delta P/P^2 Q_\gamma)^{1/2}\hat{k}_{\mathrm{i}}$, $\hat{k}_{\mathrm{i}}$ $(i,j = 1,2)$ is one of the two unit vectors on the transverse plane. When $\Delta P < 0$, we can choose $W^\gamma_1 = |2\Delta P/P^2 Q_\gamma|^{1/2}\hat{k}_1, W^\gamma_2 = -|2\Delta P/P^2 Q_\gamma|^{1/2}\hat{k}_2$. The normalization factor $Q_\gamma = \int \mathrm{d}^3k/(2\pi)^3 \Delta P(\boldsymbol{k}) \hat{k}_1 \hat{k}_2 W^\gamma_1(\boldsymbol{k}) W^\gamma_2(\boldsymbol{k})$. The two components of shear are now:

$$\hat{\gamma}_1 = E_{\gamma 12} = E_{\gamma 21}, \ \hat{\gamma}_2 = \frac{E_{\gamma 11} - E_{\gamma 22}}{2} \ . \tag{3.23}$$

Note that there is a difference between the reconstruction for convergence and shear. Shear reconstruction depends on the gradient of $P(k)$, while convergence reconstruction depends on the gradient of $\Delta^2(k)$ in a 2-D analogue. To test our method, we can generate a Gaussian source field with power-law power spectrum $P(k) = k^\beta$. In the 2-D analogue case, the convergence field can not be measured if $\beta = -2$, because the variance is conserved. However in 3-D, when $\beta = -3$, the convergence field can still be measured, which is due to the more complicated shape of the window function in 3-D. When $\beta = 0$, the shear can not be measured in either 2-D or 3-D.

In analogy to κ reconstruction, we can calculate the normalization factors b_{γ_1} and b_{γ_2}. The calculations for the normalization factors and noise are presented in the appendix.

3.2.1 The combined estimator and the OQDE

The combined estimator of κ can be written as

$$\hat{\kappa}_{\rm comb}(\boldsymbol{\ell}) = c_1\hat{\kappa}(\boldsymbol{\ell}) + c_2\hat{\gamma}_{\rm E}(\boldsymbol{\ell})\,, \tag{3.24}$$

where $\hat{\gamma}_{\rm E}$ is the convergence constructed from shear field,

$$\hat{\gamma}_{\rm E}(\boldsymbol{\ell}) = \hat{\gamma}_1(\boldsymbol{\ell})\cos 2\boldsymbol{\theta}_\ell + \hat{\gamma}_2(\boldsymbol{\ell})\sin 2\boldsymbol{\theta}_\ell\ , \tag{3.25}$$

and $\boldsymbol{\theta}_\ell$ is the angle of $\boldsymbol{\ell}$. c_1 and c_2 are the weights of two components. The optimal c_1 and c_2 can be calculated from the covariance matrix of the two estimators

$$\mathbf{C}_\kappa = \left(\begin{array}{cc} \langle\hat{\kappa}(\boldsymbol{\ell})^2\rangle - \langle\hat{\kappa}(\boldsymbol{\ell})\rangle^2 & \langle\hat{\kappa}(\boldsymbol{\ell})\hat{\gamma}_E(\boldsymbol{\ell})\rangle - \langle\hat{\kappa}(\boldsymbol{\ell})\rangle\langle\hat{\gamma}_E(\boldsymbol{\ell})\rangle \\ \langle\hat{\kappa}(\boldsymbol{\ell})\hat{\gamma}_E(\boldsymbol{\ell})\rangle - \langle\hat{\kappa}(\boldsymbol{\ell})\rangle\langle\hat{\gamma}_E(\boldsymbol{\ell})\rangle & \langle\gamma_E(\boldsymbol{\ell})^2\rangle - \langle\gamma_E(\boldsymbol{\ell})\rangle^2 \end{array} \right) .$$

To minimize the variance of $\hat{\kappa}_{\rm comb}(\boldsymbol{\ell})$, c_1 and c_2 are sum of the first and second row components respectively in $\mathbf{C}_\kappa^{-1}$, the inverse matrix of $\mathbf{C}_\kappa$. When $\ell \lesssim \ell_a = k_a\chi(z_s)$, $\hat{\kappa}(\boldsymbol{\ell})$ and $\hat{\gamma}_{\rm E}(\boldsymbol{\ell})$ are uncorrelated and the covariance matrix is diagonal. The weights are simply proportional to the reciprocal of the noise of two individual estimators $c_1 \propto N_\kappa(\boldsymbol{\ell})^{-1}$ and $c_2 \propto N_{\gamma_{\rm E}}(\boldsymbol{\ell})^{-1}$.

The lensing power spectrum $C^\kappa_{\boldsymbol{\ell}}$ is measured by taking the observed power spectrum $\hat{C}^\kappa_{\boldsymbol{\ell}}$ and subtracting the computable power spectrum of the noise $N_\kappa(\boldsymbol{\ell})$ (sometime also written as $C^N_{\boldsymbol{\ell}}$)

$$C^{\kappa}_{\boldsymbol{\ell}} = \hat{C}^{\kappa}_{\boldsymbol{\ell}} - C^N_{\boldsymbol{\ell}} \, . \tag{3.26}$$

If the source is Gaussian and its power spectrum is known, the noise power spectrum of estimator $N_\kappa(\boldsymbol{\ell})$ can be calculated using the method described in appendix. If source is non-Gaussian, $N_\kappa(\boldsymbol{\ell})$ is not derivable from the two-point function, the lensing power spectrum could be biased. For scales $\ell \lesssim \ell_a = k_a \chi(z_s)$, $N_{\kappa\gamma_E}(\boldsymbol{\ell}) \approx 0$. The $\hat{\kappa} - \hat{\gamma}_E$ cross power spectrum does not contain the noise term and does not depend on the source four-point function. It is less optimal, but more robust. Therefore the cross power spectrum of convergence and shear will be an unbiased estimator of lensing power spectrum even when we do not know the unlensed source four-point function, and the noise of the two estimators themselves.

The 2-D OQDE in CMB lensing can be written as product of two filtered temperature fields (Hu, 2001; Lewis & Challinor, 2006). Furthermore, the 3-D OQDE can be written in the same form as Eq. (1.44), though it is not explicit (private communication with Oliver Zahn).

$$\boldsymbol{E}_{\rm d}(\boldsymbol{\theta}) = L^{-1} \int {\rm d}x_{\|} \boldsymbol{T}_1(\boldsymbol{\theta}, x_{\|}) T_2(\boldsymbol{\theta}, x_{\|}) \ , \tag{3.27}$$

and

$$\int {\rm d}^2\theta' \boldsymbol{d}(\boldsymbol{\theta}') b_{\rm d}(\boldsymbol{\theta} - \boldsymbol{\theta}') = \langle \boldsymbol{E}_{\rm d}(\boldsymbol{\theta}) \rangle \ . \tag{3.28}$$

$b_{\rm d}$ is a normalization factor, $\boldsymbol{T}_1 = \int {\rm d}^2\theta' {\rm d}x'_{\|} T(\boldsymbol{\theta}') \boldsymbol{W_1}(\boldsymbol{\theta} - \boldsymbol{\theta}', x_{\|} - x'_{\|})$ and $T_2 = \int {\rm d}^2\theta' {\rm d}x'_{\|}$ $T(\boldsymbol{\theta}') W_2(\boldsymbol{\theta} - \boldsymbol{\theta}', x_{\|} - x'_{\|})$ are convolved temperature fields, where the window functions are Fourier transforms of:

$$\begin{aligned} \boldsymbol{W}_1(\boldsymbol{\ell}, k_{\|}) &= \frac{-{\rm i}\boldsymbol{l} P(\boldsymbol{\ell}, k_{\|})}{\tilde{P}^{\rm tot}(\boldsymbol{\ell}, k_{\|})}, \\ W_2(\boldsymbol{\ell}, k_{\|}) &= \frac{1}{\tilde{P}^{\rm tot}(\boldsymbol{\ell}, k_{\|})} \ . \end{aligned} \tag{3.29}$$

We note that the OQDE and our estimators have the same form.

As explained in section 1.4.3, the contribution from lensing in Eq. (1.42) is secondary, and the noise of reconstruction is mainly determined by the unlensed terms. Therefore we can measure the numerical reconstruction noise without lensing the sources.

3.3 Numerical methods

3.3.1 Simulation

The fluctuation in the 21-cm brightness temperature may depend on many factors, such as the gas density, temperature, neutral fraction, radial velocity gradient and Ly α flux (Barkana & Loeb, 2005). In our work, we do not consider the redshift space distortion effect caused by the non-zero radial peculiar velocity gradient, and simply assume the brightness temperature is proportional to the density of the neutral gas.

$$T_b \approx (27\text{mK})\left(\frac{1+z}{10}\right)^{1/2}\frac{T_s - T_{\rm CMB}}{T_s}(1+\delta_{\rm HI}) \ , \tag{3.30}$$

where T_b is the brightness temperature increment respective to CMB, T_s is the spin temperature, which is expected to be much bigger than $T_{\rm CMB}$ once structure are non-linear, and $\delta_{\rm HI}$ is the over-density of the neutral hydrogen.

Our work mainly focuses on the non-Gaussian aspect and 3-D properties of the reconstruction, and these effects also exist in a pure dark matter distribution. The neutral gas will trace the total mass distribution, which is dominated by the dark matter haloes. A simplification is to use the dark matter as the source directly. Even though this will bring some bias at small scales, the approximation is valid at large scales (Trac & Pen, 2004). The dark matter distributions are generated using the PMFAST code (Merz et al., 2005).

The high resolution PMFAST simulation was performed on a 1456^3 fine mesh with 3.9×10^8 particles. The production platform was the IA-64 'lobster' cluster at CITA, which consists of 8 nodes. One of them was upgraded, so we used the remaining 7 nodes.

Each node contains four 733 MHz Itanium-1 processors and 64 GB RAM. The simulation started at an initial redshift $z_i = 100$ and ran for 63 steps with comoving box-size $L = 50h^{-1}\,\mathrm{Mpc}$. The initial condition was generated using the Zeldovich approximation, and the matter transfer function was calculated using CMBFAST (Seljak & Zaldarriaga, 1996). The cosmological parameters were chosen in accordance with the WMAP result (Spergel et al., 2003): $\Omega_\mathrm{m} = 0.27, \Omega_\Lambda = 0.73, \Omega_b = 0.044, n = 1.0, \sigma_8 = 0.84$, and $h_0 = 0.71$. 20 independent boxes were generated. We had 3-D data at $z = 7$ at hand, and used them in our numerical tests for convenience.

3.3.2 Convergence and shear map construction

The dimensionless power spectrum, which is the contribution to the variance of overdensity per logarithmic interval in spatial wave number, can be measured from the source data in the periodic simulation box.

To reduce the computation time, our numeric results on the reconstruction used a re-sampled distribution. We generate 20 independent sources each on 512^3 grids, to investigate the statistics. The total co-moving length along the line-of-sight of 20 simulation boxes is $1h^{-1}\,\mathrm{Gpc}$, which is about the same size as the observable 21-cm region distributed between redshifts $10-20$. The correlation between the boxes can be ignored since the box-size is much larger than the non-linear length scale, and the number of neglected modes is small. The evolution of structure over this redshift is significant. Our simulations were all measured at the same redshift, so we anticipate the real effects of non-Gaussianity to be smaller. On the other hand, large scale power is generated by reionization bubbles, which may contribute to non-Gaussianity as well. In Fig. 3.1, the solid line is the average power spectrum of the re-sampled sources. To measure the dependence of non-Gaussianity on scale, we will compare the results with different scales of experimental noise cut-off.

We simply assume the noise to be zero above a cut-off and infinity below the cut-off scale. This is a reasonable approximation for a filled aperture experiment, which has good brightness sensitivity, and an exponentially growing noise at small scales. Three cut-off where chosen at $k_\mathrm{c} = 1h\,\mathrm{Mpc}^{-1}$, $4h\,\mathrm{Mpc}^{-1}$, $16h\,\mathrm{Mpc}^{-1}$, which represent the linear, quasi-linear and non-linear scales. Three different experimental noise levels are shown as vertical lines in Fig. 3.1.

In principle, the convergence map is the variance (or covariance when the filter $\mathcal{F}^\kappa$ has negative value) of the over-density field after a specified filtering process. Shear is the covariance of two maps, since the anisotropic filter can not be factored in to a perfect square. We need to smooth the maps to extract the lensing signal with maximum S/N. The window function used to smooth the lensed map, which is isotropic in the transverse directions to the line-of-sight, can be calculated with Eq. (3.15). The gradient of the power spectrum becomes negative at small scales; that comes from the limited resolution of the N-body simulation and is unphysical. The experimental noise will put a natural cut-off at small scales.

As mentioned in Section 2.2.2, the reconstruction of κ will depend on $2P + \Delta P$. In 2-D this is equivalent to the gradient of the 2-D version of $\Delta^2_{\mathrm{2D}} = k^2 P_{\mathrm{2D}}(k)/2\pi$. In 3-D, it is more complicated since $\Delta P(\boldsymbol{k})$ is not isotropic. The optimal window functions have two parts W_1 and W_2, the choice of which is not unique. One might expect a symmetric decomposition to have the best S/N. The optimal filter of κ is positive except at a few modes, and can be decomposed into two equivalent parts (one part needs to contain a minus sign for those negative value of the filter). In contrast to κ, the shear construction needs to use the covariance between two different windowed temperature fields, since there is a sin or cos component in the window function. The window is a function of the transverse and parallel components of $\boldsymbol{k}$.

We can calculate the mean covariance of the two smoothed maps along the redshift axis for each pixel. From Eq. (3.17) we can construct the convergence map. Shear maps are reconstructed in the same way, except different optimal window functions are used.

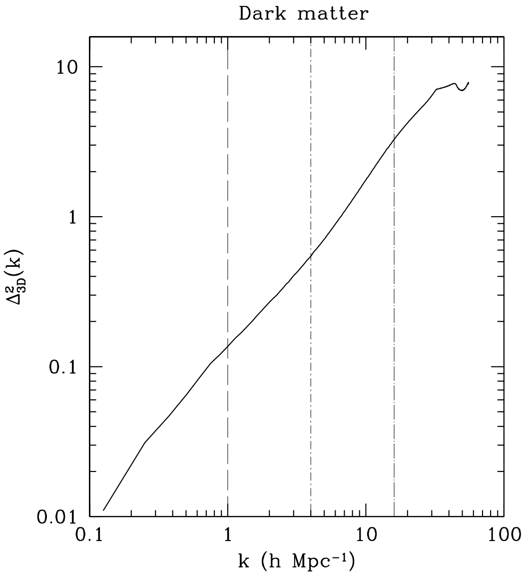

Figure 3.1 The average dimensionless power spectrum of the re-sampled dark matter from the 1456^3 N-body simulation in three dimensions are given. The solid line is the power spectrum on the 512^3 grids. The re-sampled sources keep the non-linearity and the non-Gaussianity of the structures up to $k \sim 30h\,\mathrm{Mpc}^{-1}$. Three different experimental noise cut-offs are shown with $k_\mathrm{c} = 1h\,\mathrm{Mpc}^{-1}$,$4h\,\mathrm{Mpc}^{-1}$, $16h\,\mathrm{Mpc}^{-1}$, which represent the linear, quasi-linear and non-linear scales.

The anisotropic part $\cos 2\theta_{\boldsymbol{k}_\perp}$ can be decomposed into $\cos\theta_{\boldsymbol{k}_\perp} - \sin\theta_{\boldsymbol{k}_\perp}$ and $\cos\theta_{\boldsymbol{k}_\perp} + \sin\theta_{\boldsymbol{k}_\perp}$. Both windows can generate a field with even PDF so that the distribution is less non-Gaussian. This is consistent with the numerical results as shown in Fig. 3.6. Using these two maps, we construct the γ_1 map with their covariance, as shown in the Eq. (3.23). Similarly we can get the γ_2 map.

3.4 Numerical results and discussion

Cooray (2004) claims that the variance will not vary considerably and is not a ideal measurement of the lensing signal. Even though the κ field itself is only a few percent, the integrated effect from the 3-D images will reduce the noise ratio significantly to uncover the signal. Zahn & Zaldarriaga (2006) solve the problem from an alternative

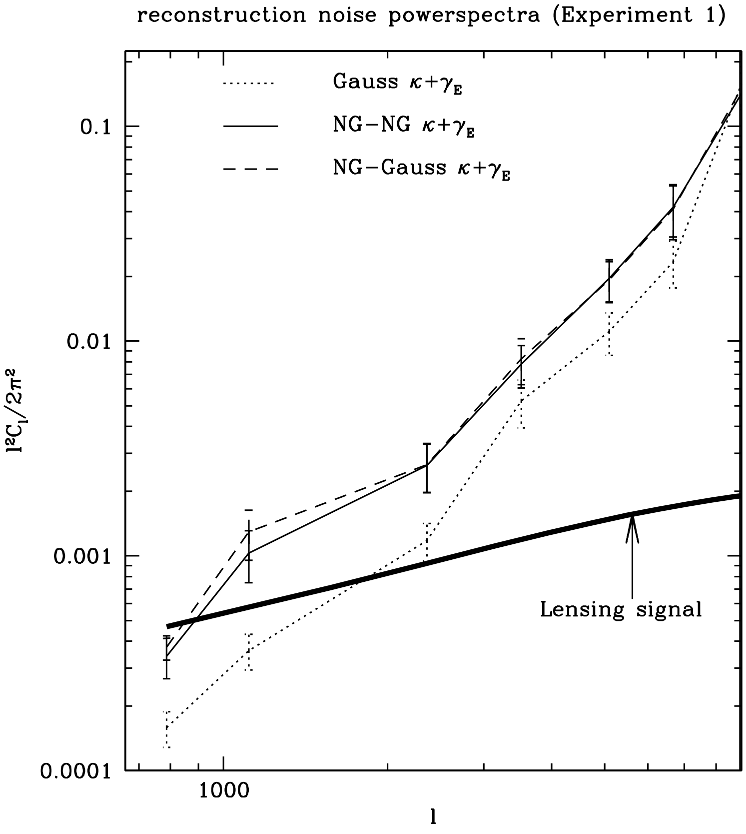

Figure 3.2 The noise of lensing maps from different estimators using experimental noise 1, which cuts off at $k_c = 1h\,\mathrm{Mpc}^{-1}$. We treat the $1h^{-1}$ Gpc rectangle of gas at $z = 10 - 20$ as 20 independent sources each is a $50h^{-1}$ Mpc box-size cube. Structures at these redshifts are similar to those at $z = 7$ used by us, though less non-linear. We can expect to see qualitatively similar non-Gaussianity effects in the reconstruction with the $1h^{-1}$ Gpc space except that the non-Gaussianity of sources may be smaller, but reionization may change that, too. The curves are truncated at $\sqrt{2}k_{\mathrm{c}}$, where the noise goes to infinity. The thick solid line is the expected lensing signal. The dotted line is the lensing reconstruction noise for a simulated Gaussian source with the same power spectrum. The dashed curve is the noise from the N-body simulation using the Gaussian estimator, which increases modestly compared to the Gaussian source. It is identical for the optimal κ, γ reconstruction as it is for the deflection angle. The thin solid line is noise when shear and convergence are re-weighted by their non-Gaussian variances.

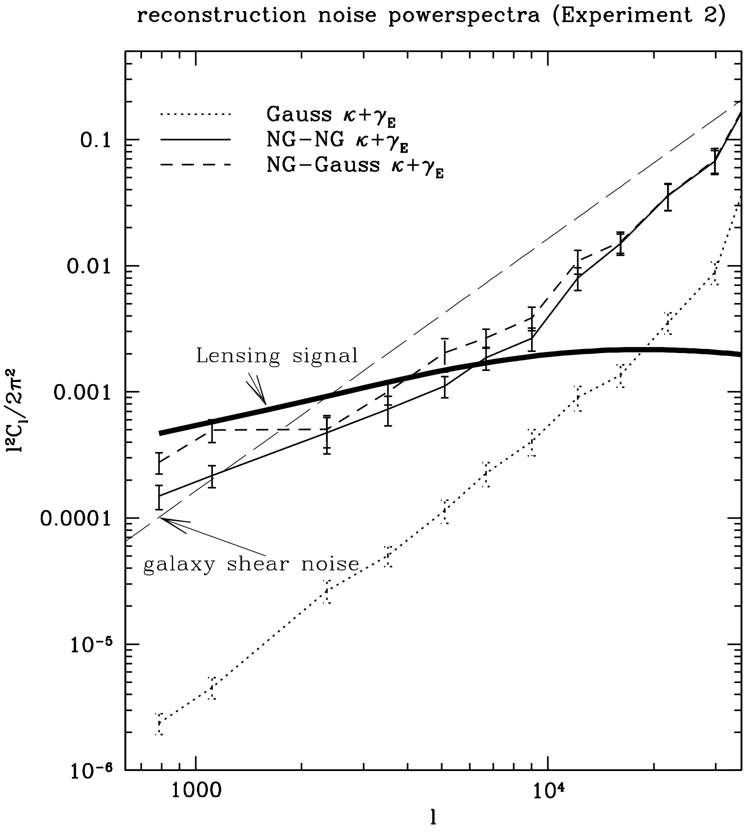

Figure 3.3 Same of Fig. 3.2, but with cut-off at the quasi-linear scales $k_c = 4h\,\mathrm{Mpc}^{-1}$. The effect of non-Gaussianity of sources is more pronounced. We can compare the S/N with a fiducial cosmic shear survey of sources in the same $10 < z < 20$ redshift range, which reconstructs the lensing from the shape of galaxies, with a surface number density of $14\,\mathrm{arcmin}^{-2}$. To map the lensing to the same S/N with redshift $z \sim 1$ sources requires a density of $56\,\mathrm{arcmin}^{-2}$ (Hu & White, 2001) with rms ellipticity of 0.4. We see that proposed optical lensing surveys are unlikely to outperform 21-cm sources.

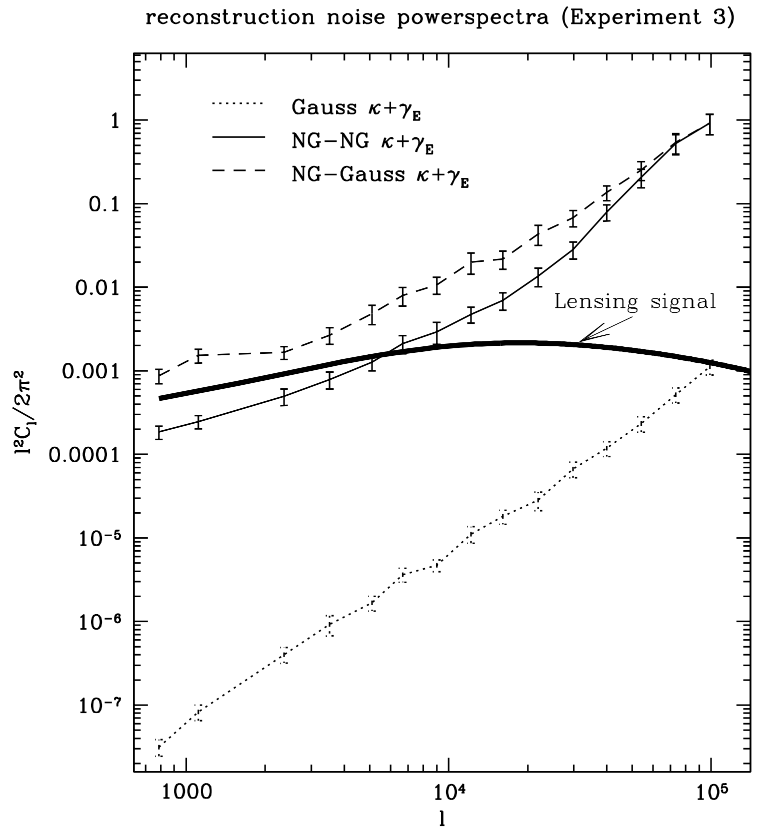

Figure 3.4 Same of Fig. 3.2, but with cut-off at the non-linear scales $k_c = 16h\,\mathrm{Mpc}^{-1}$. At the highly non-linear scales, the non-Gaussian noise variance is about 3 to 4 magnitude higher than the Gaussian noise. The combined re-weighted estimator (NG-NG $\kappa + \gamma_{\mathrm{E}}$) has noise about half an order of magnitude lower than the OQDE.

approach by generalizing the minimum variance quadratic estimator (Hu & Okamoto, 2002) in CMB lensing to 3-D.

Related work was done in Metcalf & White (2007), where they also construct quadratic estimators of shear and convergence in real space, even though they did not include the correlation between the 2-D slices along the line-of-sight and they did not choose the estimator with minimized noise.

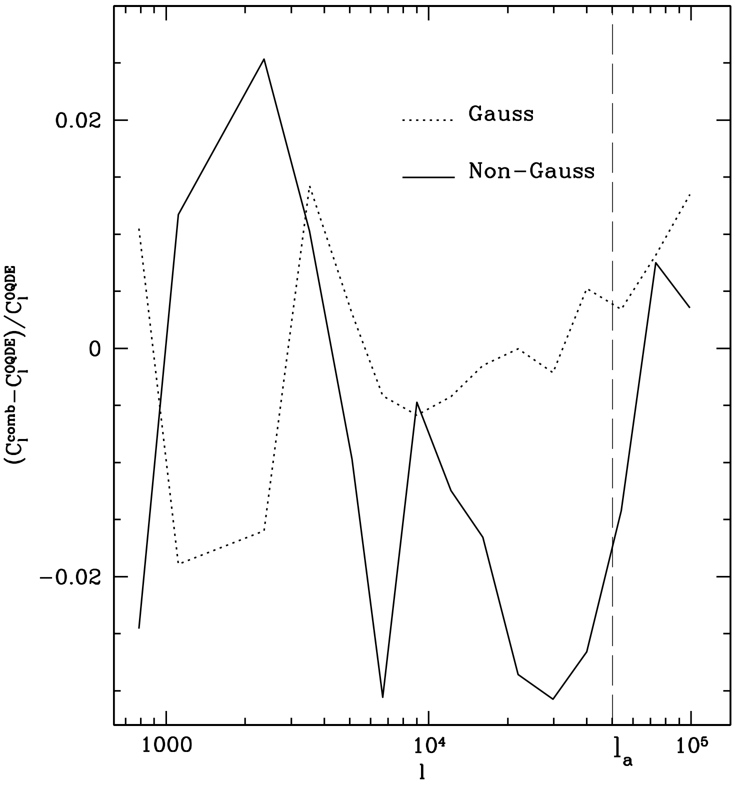

Figure 3.5 The comparison of reconstruction noise from the combined (κ, γ) estimator and the OQDE. As shown in section 3.2.1, the combined estimator becomes the sum of κ and γ_E with weights. For Gaussian sources, the weights are the reciprocal of the noise of the two estimators respectively. To compare with the OQDE, we could still use the same weights for non-Gaussian sources, i.e., using the reciprocal of the noise measured from Gaussian sources. Note the combined estimator with these weights will be less optimal than using weights calculated from the noise of non-Gaussian sources. The vertical line shows the characteristic scale ℓ_c below which the window functions are optimized. While the optimality is only proved at low ℓ, we find the combined estimator and OQDE equally optimal for Gaussian sources at all scales. For non-Gaussian sources, they also have the same results. The scatter is consistent with numerical integration errors from the tabulated power spectrum. Similarly, the noise for non-Gaussian sources is consistent between the two estimators. We conclude that the combined estimator is numerically equivalent to OQDE if the weights are optimized for Gaussian sources.

3.4.1 Non-Gaussianity

The dark matter distribution is linear at large scales, and can be treated as Gaussian. In the non-linear scales, when the amplitude of density fluctuations is big, the structure becomes highly non-Gaussian. Reference Gaussian sources with identical power spectrum to the dark matter are generated.

We treat the $1h^{-1}$ Gpc region at $z = 10 - 20$ as 20 independent sources. Structures at these redshifts are similar to those at $z = 7$ used by us, though less non-linear. We can expect to see similar non-Gaussianity effects in the reconstruction with the $1h^{-1}$ Gpc space except that the non-linear scale is smaller. We compare the reconstruction noise with three different experimental noise as well as the lensing signal in Fig. 3.2, 3.3 and 3.4. The thick solid line in the middle panel is the lensing power spectrum, which is calculated with the Limber integral of the 3-D power spectra of dark matter using Eq. (1.33). We use the publicly available code Halofit.f (Smith et al., 2003) to generate the non-linear dark matter power spectra. The code provides both their fitting results, and the results using the Peacock-Dodds formula (Peacock & Dodds (1996), PD96 hereafter). The halofit code fits the power spectrum at low redshift to Virgo and GIF CDM simulations, which used the transfer function of Efstathiou et al. (1992). At higher redshifts, the code does not operate. We use a combination of the two: Halofit power spectra are used for redshifts lower than $z = 3.0$, and PD96 power spectra are used for higher redshifts.

Since the reconstruction noise of κ is isotropic, one can always choose the direction of the lensing mode $\boldsymbol{\ell}$ to be parallel with a coordinate axis. In this direction, $\gamma_1(\boldsymbol{\ell}) = \kappa(\boldsymbol{\ell}), \gamma_2(\boldsymbol{\ell}) = 0$, and $\gamma_{\mathrm{E}} = \gamma_1$, which simplifies the numerical calculation. As shown in section 3.2.1, the combined estimator becomes the sum of κ and γ_{E} with weights. The optimal weights are calculated from the inverse matrix of $\mathbf{C}_\kappa$, covariance between the two estimators. The matrix is close to diagonal in small ℓ approximation, so for simplicity we take the diagonal terms. For Gaussian sources, the weights are the reciprocal of the noise of the two estimators respectively. For non-Gaussian sources, we could use the same weights as the Gaussian sources, i.e., using the reciprocal of the noise measured from Gaussian sources, or weigh them by calculating their respective measured noise from non-Gaussian sources. We will show that the combined estimator with Gaussian noise derived weights has the same noise as the OQDE for both Gaussian and non-Gaussian sources. Fig. 3.2, 3.3, and 3.4 are results using noise cut-offs from experiment 1, 2 and 3. The curves are truncated at $\sqrt{2}k_{\rm c}$. The non-Gaussianity increased the noise of all estimators. The first cut-off falls in the linear regime, where the non-Gaussianity only has a modest effect on the noise. The second cut-off is at the quasi-linear scale. Here the non-Gaussianity increases the noise of the OQDE by about 1 to 2 orders of magnitude. At the highly non-linear scales, the non-Gaussian noise is about 3 to 4 magnitude higher than the Gaussian noise, and in fact higher than that of the more noisy experiment.

Our estimators were derived in the limit that κ and γ are constant. For spatially variable lens, we solve for the required normalization factors. In the OQDE, the windows do not depend on the scale of the lens, so one might guess the same Ansatz to hold for the (κ, γ) estimators. We verify this numerically in Fig. 3.5. The solid line and dotted line is for Gaussian sources and non-Gaussian sources respectively. The differences are less than a few percent, and consistent with integration errors from the tabulated power spectrum, and most importantly, independent of scale, as we had expected. The vertical line shows the characteristic scale $\ell_a = \chi(z_{\rm s})k_{\rm a} \sim \chi(z_{\rm s})k_{\rm c}/2$, or below which the window functions are optimized. As shown in Fig. 3.7, the reconstruction noise is proportional to k^{-3} if sources are Gaussian, and most of the contribution to the reconstruction is from scales near $k_{\rm c}$. Therefore our approximation holds at scales $\ell < \chi(z_{\rm s})k_{\rm c}/2$ with small deviance. For non-Gaussian sources, $\ell_{\rm a}$ will also be affected by the non-Gaussianity of sources. While the optimality is only proved at low ℓ, we find the combined estimator and OQDE equally optimal for Gaussian sources at all scales. For non-Gaussian sources, they also have the same results. We do note, that for a finite size survey, the (κ, γ) recover the constant mode, which is lost in the OQDE. Three more numbers are recovered.

The combined estimator with κ and γ_{E} weighted by using the non-Gaussian noise is more optimal than weighted by using Gaussian noise, therefore has lower noise than the OQDE. In fact, the non-Gaussian noise of γ_{E} is much smaller than κ. To investigate the origin of this change, we first investigate the cause of the increased noise in non-Gaussian sources for κ. This could be because either the non-Gaussianity leads to a high kurtosis in κ, which boosts the errors; or the non-Gaussianity may lead to correlations between modes, resulting in a smaller number of independent modes, and thus a larger error.

In Fig. 3.6, the PDF of maps smoothed with the κ window are shown. The top, middle and bottom panel show the results with experimental noise cut-offs 1, 2 and 3. The solid line is the PDF for maps smoothed with κ window (T_1^κ, T_2^κ in section 3.2). Because the window functions are almost symmetric, we plot only one PDF. To see the full dynamic range on the x-axis, we plot $\pm|T|^{1/4}$ as x-axis, and $\mathrm{PDF}(|T|^{1/4})|T|^{15/4}$ as the y-axis. The integral of the x-axis weighted by the y-axis will give $\langle T^4\rangle$, which is basically a estimation of the point-wise non-Gaussian reconstruction noise. Here $\mathrm{PDF}(|T|^{1/4})$ is the PDF of $|T|^{1/4}$. To compare with a Gaussian distribution, dotted lines are also plotted. The contributions to the $\langle T^4\rangle$ in experiment 1 mainly come from small fluctuation regions. In experiment 2, the large outliers play a more important role but one can still expect the curve to converge. In experiment 3, most contributions come from rare regions with high fluctuations. Caution should be exercised in the interpretation of the most non-linear scales, since a larger number of source samples may result in a different error. It is clear, however, that the noise has increased dramatically.

The kurtosis of κ is $\langle (T_1^\kappa)^4\rangle / \langle (T_1^\kappa)^2\rangle^2 - 3$, and an analogous quantity can be defined by $\langle (T_1^\gamma T_2^\gamma)^2\rangle / (\langle (T_1^\gamma)^2\rangle \langle (T_2^\gamma)^2\rangle) - 1$ for shear. $T_1^\kappa \approx T_2^\kappa$, and T_1^γ is uncorrelated with T_2^γ. The noise of κ and γ is determined by both kurtosis and number of independent cells. For experimental noise 1, the kurtosis of T^κ and T^γ are 1.2 and 0.29 respectively. The effectively independent cube cells for κ and γ have side length $4.8h^{-1}$ Mpc and $4.6h^{-1}$ Mpc respectively. The corresponding Gaussian sources with the same cut-off have effective cell size $3.0h^{-1}$ Mpc and $3.5h^{-1}$ Mpc. For experimental noise 2, the kurtosis of T^κ and T^γ are 18 and 5.7 respectively. The effective cell size for κ and γ are $1.8h^{-1}$ Mpc and $1.5h^{-1}$ Mpc respectively. The corresponding Gaussian sources with the same cut-off have effective cell size $1.0h^{-1}$ Mpc and $1.1h^{-1}$ Mpc. For experimental noise 3, the kurtosis for T^κ and T^γ are 1.6×10^3 and 3.5×10^2 respectively. The effective cell size for κ and γ are $540h^{-1}$ Kpc and $310h^{-1}$ Kpc respectively. The corresponding Gaussian sources with the same cut-off have effective cell size $240h^{-1}$ Kpc and $290h^{-1}$ Kpc. We conclude that the shear measurements have lower non-Gaussian noise both because of a smaller point-wise kurtosis, and less correlation between modes.

We will see later that experiment 2 has the largest S/N, which is larger than unity for $\ell \lesssim 6000$. We can compare the S/N with cosmic shear surveys, which reconstruct lensing from the shape of galaxies. The noise can be estimated by $\sigma_e^2/n_{\rm g}$ (Hoekstra et al., 2006; Hu & White, 2001), where we use $\sigma_e \approx 0.4$ as the rms intrinsic ellipticity, and $n_{\rm g}$ is the effective number density of galaxies. We plot the shear noise from a survey of sources in the same redshift range $10 < z < 20$ in Fig. 3.3, with a surface number density of $14\,\mathrm{arcmin}^{-2}$. For more realistic source redshifts $z \sim 1$ in proposed optical surveys (Hu & White, 2001), this corresponds to a surface density of $56\,\mathrm{arcmin}^{-2}$ to achieve the same fidelity of dark matter reconstruction. In the CFHTLS wide survey the source galaxies are distributed at redshifts lower than 3, and their effective number density is ~ 12 galaxies arcmin^{-2} (Hoekstra et al., 2006). This noise is larger still. Even though non-Gaussian 21-cm lensing saturates lensing reconstruction, it still measures more modes than current proposed optical surveys.

In Fig. 3.7, we show the reconstruction noise at two different ℓ versus various experimental noise cut-off $k_{\rm c}$. The top panel is for the fundamental mode in the box, $\ell_1 = 2\pi/L = 783$, and the bottom panel is for $\ell_2 = 6\ell_1 = 4715$. As shown in the plot, it is clear that the noise of a Gaussian source decreases as $k_{\rm c}$ increases, because of the increasing number of independent modes. The dotted lines are a least squares fitting power-law $N_0 k_c^{-3}$ to the Gaussian noises, and $N_0 = 3.1 \times 10^{-2}, 1.3 \times 10^{-1}$ for top and bottom panels respectively. This comes from counting the number of available source modes. The dashed lines connect the non-Gaussian noises of the OQDE. The triangles are the reconstruction noise for the combination estimator, which is equal to the OQDE at larger scale k_c and about half an order of magnitude lower at smaller scales of k_c. From this plot, we can see that experiments with lower noise does not necessarily decrease the reconstruction noise of the OQDE for non-Gaussian sources. And the experimental noise has a limit around the quasi-linear scale where the OQDE achieves its best S/N. The S/N achieves its maximum around $k_c^{\rm NG} \approx 4h\,{\rm Mpc}^{-1}$. This cut-off with maximum S/N varies only slowly with ℓ.

If one wants to estimate the effective number of available lensing modes, we can derive an effective cut-off of a Gaussian field which gives the same S/N as the optimal non-Gaussian sources estimator. This is $k_c^{\rm G} \approx 2h\,{\rm Mpc}^{-1}$, where the power spectrum of source is $\Delta^2 \approx 0.2$. The size of the effectively independent cells is $2.0h^{-1}\,{\rm Mpc}$. A simple equivalent Gaussian noise estimate counts all modes up to $\Delta^2(k) < 0.2$, which is perhaps surprisingly low.

For our noise estimates, we stacked simulations all at redshift $z = 7$. While the angular diameter distance does not change much to $z \sim 20$, the structure does evolve. We do not have access to the higher redshift outputs to test this effect, but one would expect a smaller non-linear scale to result in a smaller reconstruction noise.

3.4.2 Future directions

A possible way to find the optimal window functions for non-Gaussian sources is to divide the window into N frequency bins $W_1(\boldsymbol{k}_1, \boldsymbol{k}_2, ..., \boldsymbol{k}_N)$, and apply a numerical variation to those bins. The noise can be measured numerically by applying the estimator to the simulated sources. The process of searching for a optimal filter is equivalent to look for a minimum of reconstruction noise in N dimensional space $\boldsymbol{k}_1, \boldsymbol{k}_2, ..., \boldsymbol{k}_N$. In fact, the optimal window function can be constructed by the inverse of covariance matrix of the source power spectra. We will present details of the method in chapter 4. In this chapter we only considered the class of windows which are identical to the optimal Gaussian estimators with a hard cut-off, as well as two weightings for shear and convergence.

One can also try to Gaussianize the sources by modifying the PDF of all the sources to be Gaussian. The physical explanation and details of Gaussianization can be found in Weinberg (1992). The basic idea is that every pixel should preserve its rank in the whole field during the Gaussianization process. During structure formation, the non-linear evolution at small scales should not destroy most of the information on the peaks and dips of the linear field. However, this Gaussianization process will change the power spectra of sources, and the reconstructed lensing field will be biased. This is not a linear process, and the variation of power spectrum does not have analytical solution, and can only be measured numerically with simulated sources.

Recently it has been proposed that one could economically achieve brightness mapping of 21-cm emission at lower redshifts (Chang et al., 2007), potentially even with existing telescopes. If individual galaxies are not resolved, one can again ask the question of how one could reconstruct a lensing signal. This is very similar to the problem studied in this chapter. We will discuss this with optimal non-Gaussian estimators in chapter 4.

3.5 Summary

In this chapter, we developed the maximum likelihood estimator for the large-scale structure from the 21-cm emission of the neutral gas before the epoch of re-ionization. The convergence and shears can be constructed separately. They are independent for $\ell \lesssim \ell_a$. The cross power spectrum of convergence and shear is more robust unbiased estimator of lensing power spectrum, which does not require knowledge of the unlensed source four-point function. To test the effects of non-Gaussianity, we applied our estimators

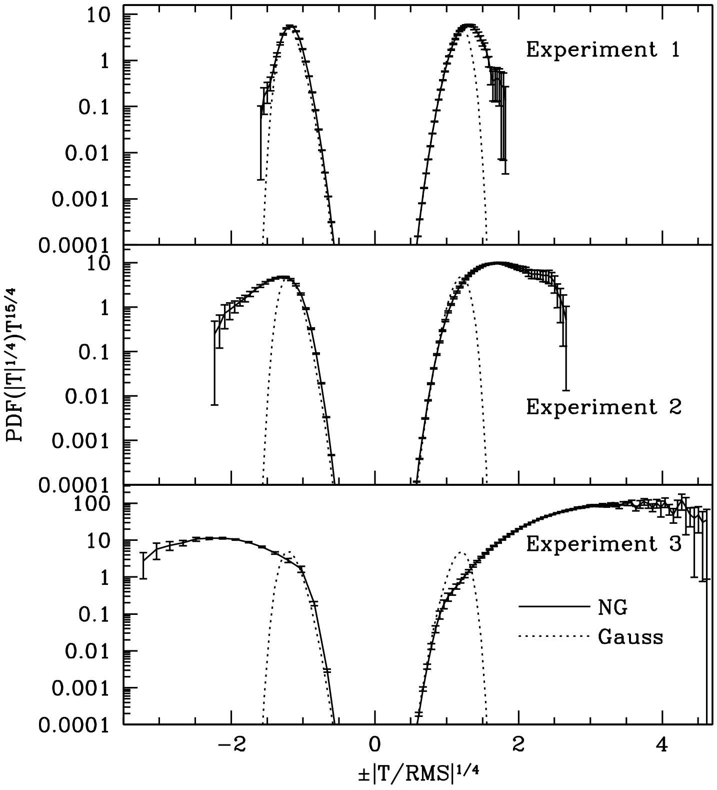

Figure 3.6 The PDF of maps smoothed with a κ window are shown. The top, middle and bottom panel show the results with experimental noise cut-offs 1, 2 and 3. The solid line is the PDF for maps smoothed with the κ window (T^{κ} in section 3.2). To see the full dynamic range on the x-axis, we plot the curve with $\pm|T|^{1/4}$ as x-axis, and $\mathrm{PDF}(|T|^{1/4})|T|^{15/4}$ as the y-axis. The integral of the x-axis weighted by the y-axis will give $\langle T^4\rangle$, which is basically a estimation of the reconstruction noise. The error bars are estimated from the 20 simulations. To compare with a Gaussian distribution, dotted lines are also plotted. The contributions to the $\langle T^4\rangle$ in experiment 1 mainly come from small fluctuation regions. In experiment 2, the large outliers play a more important role but one can still expect the curve to converge. In experiment 3, most contributions come from rare regions with high fluctuations. Caution should be exercised in the interpretation of the most non-linear scales, since a larger number of source samples may result in a different error. It is clear, however, that the noise has increased dramatically.

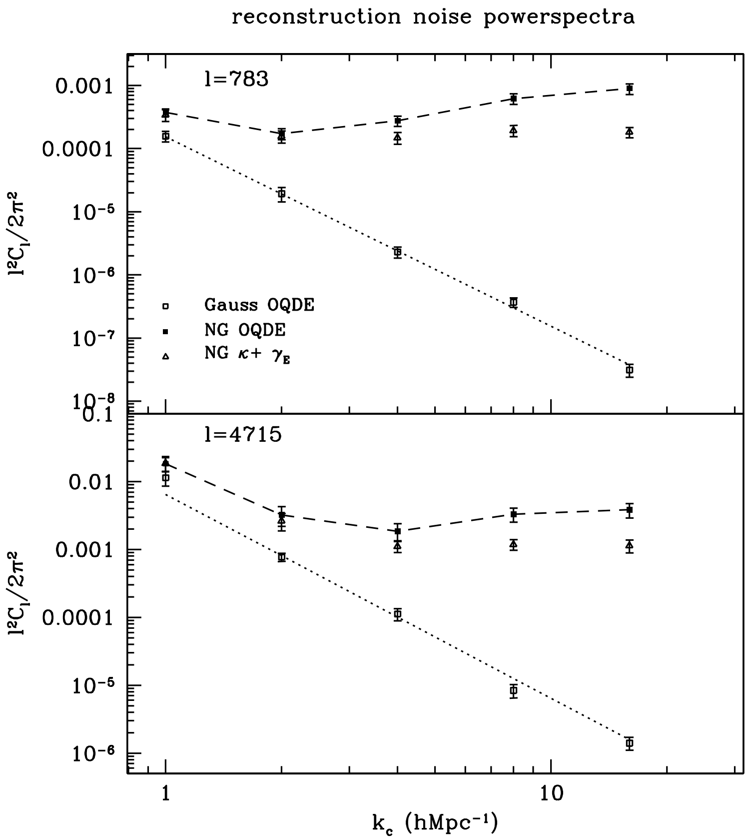

Figure 3.7 The reconstruction noise versus the cut-off in the experimental noise. The top panel is for $\ell_1 = 2\pi/L = 783$, and the bottom panel is for $\ell_2 = 6\ell_1 = 4715$. The noise of Gaussian sources decreases as k_c increases, because of the increasing number of independent modes. The dotted lines are a least squares fitting power-law $N_0 k_\mathrm{c}^{-3}$ to the Gaussian noises, and $N_0 = 3.1 \times 10^{-2}, 1.3 \times 10^{-1}$ for top and bottom panels respectively. The dashed lines connect the non-Gaussian noise of the OQDE. The triangles are the reconstruction noise for the combined estimator, which is equal to the OQDE at larger scale k_c and about half an order of magnitude lower at large k_c. The noise of the non-Gaussian sources changes slowly and saturates or even increases at small scales.

to simulated data. The sources were generated by N-body simulations, because gas is expected to trace the total mass distribution. To investigate the influence of non-Gaussianity, we also use Gaussian sources which have the same power spectra as the simulated sources. We applied our estimator and the OQDE on both the Gaussian and non-Gaussian sources. Though our estimators are derived in the simplified case of a constant convergence, the noise of our combined estimator of convergence and shear are the same as the OQDE for Gaussian sources for spatially variable lenses. For a finite survey area, three extra constant modes can be recovered.

At $z \sim 10 - 20$, the non-Gaussian nature of the source can increase the error bar by orders of magnitude, depending on the experimental cut-off scale. Shear construction is affected less by non-Gaussianity than the convergence field, and the combined estimator with non-Gaussian noise weights is a better choice than reconstructing with the OQDE. S/N can not be boosted infinitely by reducing the experimental noise, and achieves its maximum for a cut-off around $k_{\rm c}^{\rm NG} \approx 4h\,{\rm Mpc}^{-1}$. Below that scale the S/N starts to saturate or even decrease. The maximum S/N for non-Gaussian sources is equal to Gaussian sources with $k_{\rm c}^{\rm G} \approx 2h\,{\rm Mpc}^{-1}$, where the power spectrum of sources is $\Delta^2 \approx 0.2$ and the side length of the effectively independent cells is $2.0h^{-1}\,{\rm Mpc}$. The maximum S/N is greater than unity for $\ell \lesssim 6000$, which makes 21-cm lensing very competitive compared to optical approaches.

Chapter 4

Dark energy and 21-cm intensity lensing

4.1 Overview

The standard cosmological model has achieved substantial quantitative success through precision cosmology. It requires a mysterious and dominant dark energy component, which is not physically understood. Or it may be a hint that Einstein's General Relativity is not the correct theory to describe our universe.

Recently, modified gravity models have been studied quantitatively. In order to pass local precision gravity tests, and match the CMB fluctuations, they tend to differ in their predictions only on very large scales and at low redshift. A direct mapping in this regime is challenging.

The kinematic history of the universe is being mapped out using the standard metre stick of BAO. A range of surveys are underway and have been proposed to produce a coarse image of large-scale structure (LSS) over a substantial sky area and redshift range. While these surveys are not directly suitable for lensing mapping, the maps they produce will be distorted by gravitational lensing (Hui et al., 2007). Since the intrinsic correlations is known to be statistically isotropic, the lensing induced changes can be used to reconstruct the dark matter map, in analogy with lensing of the CMB. The latter is sensitive to the matter at redshifts $z \sim 2-3$, where dark energy is expected to have little impact. On the other hand, BAO surveys map the universe in the redshift ranges 1-4, hence the associated lensing signal will be sensitive to the non-linear evolution of dark matter at $z \sim 0.5-1$, where it is strongly affected by dark energy.

Some of the proposed fast BAO surveys (Chang et al., 2008; Wyithe et al., 2008; Tegmark & Zaldarriaga, 2008) are based on an intensity mapping approach, where the LSS is mapped without detection of individual galaxies over a large fraction of the sky. For such surveys, point source based lensing mappings are not applicable (Zhang & Pen, 2005), and the LSS lensing is the only available tool. Lensing reconstruction using Gaussian Random Fields as sources is now well understood for both 2-D and 3-D structures (Pen, 2004; Cooray, 2004; Zahn & Zaldarriaga, 2006; Metcalf & White, 2007; Lu & Pen, 2008; Metcalf & White, 2008). Metcalf & White (2008) have explored the possibility of applying 21-cm lensing from low and high redshifts to the dark energy constraint. While their method is based on the Gaussian assumption, it was shown that the neglect of non-Gaussianity could generate orders of magnitude difference in the noise estimation (chapter 3, also Lu & Pen (2008)).

In general, lensing reconstruction is a quadratic function of the density field, which is quantified by its power spectrum. RH05 showed that the Fisher information available in the power spectrum saturates at trans-linear scales, and stays several orders of magnitude below the Gaussian Fisher information content. In this chapter, we combine the information from these approaches, and construct the optimal quadratic lensing estimators for non-linear sources fields from N-body simulations.

We find that the reconstruction noise of lensing from the simulated sources decreases with reduced experimental noise at linear scales, saturates at quasi-linear scales, and drops again at highly non-linear scales. There is a plateau region at the trans-linear regime, which is analogous to what was found with the Fisher information content in the initial amplitude of matter power spectrum by Rimes & Hamilton (2005). Both effects are caused by the non-Gaussianity (or non-linearity) introduced by gravitational clustering during structure formation. We found the saturation scale is a steadily increasing function of the shape of the linear power spectrum.

This non-Gaussianity makes the 21-cm lensing less promising than it first appears. However, because the abundance of 21-cm sources, the S/N of 21-cm lensing at high redshift $1-6$ is still competitive compared with other surveys. The effective number density of sources defined by the independent number of cells in the 21-cm source, increase quickly with redshift. At redshift $4-6$, this number density could be $9\,\mathrm{arcmin}^{-2}$.

The chapter is organized as follows: We overview the progress in the LSS information and lensing studies in section 4.2. The optimal lensing estimator is introduced in section 4.3. The numerical methods are presented in section 4.4. The results are discussed in section 4.5. We conclude in section 4.6.

4.2 Information and lensing from the LSS

The lensing of Gaussian Random Fields can be described in several ways. For the purposes of this chapter, we are in a regime where the fluctuation scale of the sources is smaller than the structures of interest in the lenses. In this limit, lensing has two effects on an image: convergence (κ) and shear (γ).

As we will find through detailed calculations, these two quantities carry a comparable amount of information, dependent on the slope of the power spectrum of sources. For pedagogical purposes, we first consider the convergence, that is the stretching of the image changes the amplitude of fluctuations at a fixed apparent angular scale. By comparing the power spectrum in different patches, we can measure the relative convergence.

The accuracy of this procedure is proportionate to the precision with which the power spectrum can be measured. This problem has been studied in terms of Fisher Information by RH05, and their results directly translate into the accuracy of lensing reconstruction. After a technical calculation, we find that the effective number of independent lensing sources is equal to the RH05 Fisher Information, scaled by a coefficient of order unity, which depends on the slope of the power spectrum.

A related problem is the impact of lensing on the sky averaged power spectrum (Hui et al., 2007, 2008; Loverde et al., 2008). There had been concern that the lensing at low redshift may impact the ability to accurately measure the matter power spectrum at high redshift. To first order, there is an equal number of positive and negative convergence patches, so the net effect is second order. Since we can measure the first order effect on each patch, we can actually correct its impact, and cancel off the second order bias.

4.3 Lensing reconstruction

In the lensing of diffuse background, what is observed is the distorted brightness temperature distribution. As we discussed in chapter 3, all lensing estimators are quadratic terms of temperature fields. This is because the underlying lensing field has zero expectation value at the position of any particular pixel, thus the single field terms disappear. In other words, the lensing field (convergence κ, shear γ, or deflection angle $\boldsymbol{d}$) can be reconstructed with the product of two smoothed temperature fields. For different estimators, the window functions could also be scalars or vectors.

How to choose the smoothed window is just a matter of optimization. There are also progressive levels of optimization to be considered: Is the source isotropic? Is the window function isotropic? For Gaussian distributed temperature field, the optimal window function can be solved analytically. Hu & Okamoto (2002) formulated the optimal quadratic deflection angle estimator (OQDE) for 2-D CMB lensing, and later Zahn & Zaldarriaga (2006) generalized it to 3-D 21-cm lensing. The OQDE surprisingly has the same form at all scales, even though an additional scale dependent normalization factor needs to be applied at the last step. In an alternative approach, by optimizing the reconstruction noise in the zero-mode, we derived the optimal Gaussian window for the convergence and shear fields separately in chapter 3.

In reality, 21-cm sources are not Gaussian fields. For simplicity, we use the 3-D dark matter distribution to represent the 21-cm emission at the same redshift. We investigated both OQDE and our estimators by numerical tests with simulated non-Gaussian source fields and mocked Gaussian ones. For Gaussian sources, we found the combined result of our two estimators has exactly the same S/N as the OQDE. This agrees with our expectation because of the scale in-dependency of the optimal window functions. If one window function is optimal at one scale, it is also optimal at other scales. For non-Gaussian sources however, the optimal window function derived from Gaussian assumption is not optimal any more. The lensing reconstruction noise could be underestimated by orders of magnitude if we use the results from Gaussian sources, which could be calculated analytically by applying Wick's theorem. By choosing appropriate weights, our combined estimator has better performance than the OQDE.

In this chapter, we will further explore the optimization of non-Gaussian sources. We will develop the optimal non-Gaussian estimator, which can only be constructed by numerical measurements from a large sample of simulations. The reconstruction noise of optimal lensing estimators for non-Gaussian 21-cm sources are closely related to the Fisher information contained in non-linear matter power spectrum described by RH05. They indicated that the Fisher information for the initial amplitude of matter power spectrum could be written as the sum of the inverse of covariance matrix of power spectrum estimates, multiplied by the partial derivative of power spectrum with respect to the initial amplitude. They have revealed the cumulative Fisher information results up to a maximum scale, which enters the calculation as the upper limit scale of the covariance matrix. They found a very interesting phenomenon: the Fisher information increases at linear scales, and then the growth becomes very slow at the quasi-linear region (which they label information plateau), and the Fisher information starts to rise steeply again in non-linear scales.

In this context, Fisher information means how accurate the amplitude of the power spectrum can be measured. As we will discuss later in section 4.3.1, lensing can be calibrated by the change of the source power spectrum. We will also prove that the optimal window function of non-Gaussian sources actually contains the inverse covariance matrix of the power spectrum. Thus the reconstruction noise of lensing roughly estimates how much information is contained in the source power spectrum: the lower the noise level achieved, the more information is gained. Instrumental noises of 21-cm experiments can be approximately treated as hard cut-offs at some scales, then we could inspect the reconstruction noises with different cut-off scales. When the sources are Gaussian, the covariance matrix is diagonal and could be expressed as a square term of the source power spectrum, and the optimal window function reverts to the form we found in chapter 3. Our numerical results with Gaussian estimators had similar behavior as RH05: the reconstruction S/N increased first and reached a peak at trans-linear scales. At non-linear scales, the S/N dropped again. We will explain later, the decline of S/N is an artificial effect by using the non-optimal window function for non-Gaussian sources. In this chapter, we will show that the optimal non-Gaussian estimator also leads to a plateau in trans-linear scales in the S/N plot, and like Fisher information of matter power spectrum amplitude, the curve falls again in non-linear scales.

The covariance matrix of non-linear power spectrum is measured from an ensemble of N-body simulations. RH05 showed that the covariance of power spectrum can only be measured with at least hundreds of independent simulations. Therefore we generated around 100 or more different sources, depending on the redshift, by running the same number of N-body simulations.

4.3.1 Optimal non-Gaussian estimator

As we explained in chapter 1, for the lensing of a diffuse background, all unbiased estimators are two-point functions of the source temperature field with each points convolved with a window function, no matter which quantity is reconstructed: the convergence, shear, deflection angle or the potential. To optimize the estimator, one minimizes its variance which is a four-point function of the temperature distribution. Except for a Gaussian distribution, four-point functions are in general not analytical, and can not be calculated from two-point statistics.

In chapter 3, we derived optimal estimators of the convergence and shear for a Gaussian source distribution. Since the optimal window function for Gaussian sources (hereafter optimal Gaussian estimator or optimal Gaussian window function) does not depend on the scale, we were able to look for its form in the limit of a slowly varying κ (or γ). We then applied a maximum likelihood method to solve the optimal window. The calculation was done analytically assuming that the covariance matrix of the power spectrum is diagonal and can be obtained from the power spectrum, which is the implication of Wick's theorem. The variance of the optimal Gaussian estimators rise when the resolution of the observation is improved, which shows the optimal Gaussian estimators are far from optimal for non-Gaussian sources distribution.

To find the optimal estimator for non-Gaussian sources, we pursue the calculation in a slightly difference approach. We adopt a minimum variance method, demonstrate it with the Gaussian case and generalize it to the non-Gaussian case. First, we solve the optimal estimator for a constant κ and a Gaussian source, i.e., $\kappa(\boldsymbol{x}) = \kappa_0$. We then generalize the estimator to the optimal non-Gaussian estimator of varying κ and γ.

The lensed power spectrum from observations can be Taylor expanded in first order as:

$$\begin{aligned}\tilde{P}^{\rm tot}(\boldsymbol{k}) &= \tilde{P}(\boldsymbol{k}) + P^N(\boldsymbol{k}) \\ &\approx P(\boldsymbol{k}) + P^N(\boldsymbol{k}) + \left.\frac{\partial \tilde{P}(\boldsymbol{k})}{\partial \kappa}\right|_{\kappa=0} \kappa\,, \end{aligned} \tag{4.1}$$

where the noise $P^N = P^e + P^s$ includes both the instrumental noise, P^e, and the statistical fluctuation of the power spectrum, P^s, which is the sample variance for a Gaussian source. In chapter 3, we have defined the derivative of the lensed power spectrum with respect to κ, i.e., $G(\kappa, \boldsymbol{k}) \equiv \frac{\partial \tilde{P}(\boldsymbol{k})}{\partial \kappa}$. For convenience, we will use G to represent its value at $\kappa = 0$, therefore

$$G(\boldsymbol{k}) \approx 2P(\boldsymbol{k}) + \Delta P(\boldsymbol{k})\,, \tag{4.2}$$

where $\Delta P = P' k (k_\perp^2/k^2)$, $P'(k) = {\rm d}P(k)/{\rm d}k$, $k_\perp$ means the component of k in the transverse plane of the line of sight, as defined in chapter 3. Note that here $P(\boldsymbol{k})$ is the expectation value of source power spectrum estimate, and $\tilde{P}(\boldsymbol{k})$ is the lensed power spectrum of one realization. In the absence of lensing, the power spectrum in one realization is measured by an estimator $\hat{P}(\boldsymbol{k})$, which is equal to the sum of $P(\boldsymbol{k})$ and $P^N(\boldsymbol{k})$. We frequently use the approximation that $\langle(\tilde{P}^{\rm tot}(\boldsymbol{k}) - P(\boldsymbol{k}))^2\rangle \approx \langle(\hat{P}(\boldsymbol{k}) - P(\boldsymbol{k}))^2\rangle$ in the reconstruction noise estimation. To simplify the algebra, we first consider the case $P^e = 0$, and $P^N = P^s$. It will later be straightforward to add the instrumental noise.

Discrete case

Following the convention of Bertschinger (1992), we use subscript $[...]_{\boldsymbol{k}}$ to describe discrete quantities, and parenthesis $[...](\boldsymbol{k})$ to describe continuous quantities. For example,

$$\begin{aligned}\tilde{T}(\boldsymbol{k}) &= \int_{-\infty}^{\infty} {\rm d}^3x \tilde{T}(\boldsymbol{x}) e^{-{\rm i}\boldsymbol{k}\cdot\boldsymbol{x}}\,, \\ \tilde{T}_{\boldsymbol{k}} &= \sum_j \tilde{T}(\boldsymbol{x}_j) e^{-{\rm i}\boldsymbol{k}\cdot\boldsymbol{x}_j}\,. \end{aligned} \tag{4.3}$$

Note that the observed temperature field is from the contribution of lensed brightness temperature field and noise: $\tilde{T} = \tilde{T}_b + n$. The expectation value of the source power spectrum is $P_{\boldsymbol{k}} = \langle T_{b\boldsymbol{k}} T^{\star}_{b\boldsymbol{k}} \rangle$, and $T_{b\boldsymbol{k}}$ is the discrete Fourier transform of $T_b(\boldsymbol{x})$. If the source brightness temperature follows a Gaussian distribution, the estimators of the unlensed power spectrum $\hat{P}_{\boldsymbol{k}}$ at different $\boldsymbol{k}$ are independent, and their variance can be calculated as $\mathrm{Var}(\hat{P}_{\boldsymbol{k}}) = \langle {P^{N}_{\boldsymbol{k}}}^2 \rangle = \langle {P^{s}_{\boldsymbol{k}}}^2 \rangle = P^2_{\boldsymbol{k}}/n_{\boldsymbol{k}_b}$, where $n_{\boldsymbol{k}_b}$ is the number of modes in the $\boldsymbol{k}$ frequency bin.

We first consider the discrete case with an independent set of Fourier frequencies $\boldsymbol{k}$. For every frequency, an estimator of κ can be constructed as:

$$\begin{aligned} \hat{\kappa}_{\boldsymbol{k}} &= \frac{\tilde{P}^{\mathrm{tot}}_{\boldsymbol{k}} - P_{\boldsymbol{k}}}{G_{\boldsymbol{k}}} \\ &= \kappa_0 + n_{\boldsymbol{k}}\,, \end{aligned} \tag{4.4}$$

where $n_{\boldsymbol{k}} = P^N_{\boldsymbol{k}}/G_{\boldsymbol{k}}$. Note $\hat{\kappa}_{\boldsymbol{k}}$ is not the Fourier transform of $\hat{\kappa}(\boldsymbol{x})$, which has a constant value, but the measurement of κ_0 at frequency $\boldsymbol{k}$. The optimal estimator of κ_0 should be the total contribution from all $\boldsymbol{k}$ appropriately weighted. The minimum variance estimator of κ_0 is

$$\hat{\kappa}_0 = \frac{\sum \hat{\kappa}_{\boldsymbol{k}}/\sigma^2_{\boldsymbol{k}}}{\sum 1/\sigma^2_{\boldsymbol{k}}}\,, \tag{4.5}$$

where $\sigma^2_{\boldsymbol{k}} = \langle n_{\boldsymbol{k}} n^{\star}_{\boldsymbol{k}} \rangle = P^2_{\boldsymbol{k}}/n_{\boldsymbol{k}_b} G^2_{\boldsymbol{k}}$. Therefore, the weight at each $\boldsymbol{k}$ is inverse proportional to the $\sigma^2_{\boldsymbol{k}}$. The reconstruction noise of the minimum variance estimator is now

$$\mathrm{Var}(\hat{\kappa}_0) = \frac{1}{\sum 1/\sigma^2_{\boldsymbol{k}}}\,. \tag{4.6}$$

More generally, in the non-Gaussian case, power spectrum at different $\boldsymbol{k}$ are correlated. We can write all variables in the form of matrices and vectors:

$$\mathbf{G}\kappa_0 = \tilde{\mathbf{P}}^{\mathrm{tot}} - \mathbf{P} - \mathbf{P}^N\,, \tag{4.7}$$

where $\mathbf{G}$, $\mathbf{P}$ and $\mathbf{P}^N$ are the $N_{\boldsymbol{k}} \times 1$ matrix composed by $G_{\boldsymbol{k}}$,$P_{\boldsymbol{k}}$ and $P^N_{\boldsymbol{k}}$ respectively, and $N_{\boldsymbol{k}}$ is the number of all frequencies. The minimum variance (least square) estimator can be written as

$$\mathbf{G}^t\mathbf{C}^{-\mathbf{1}}\mathbf{G}\hat{\kappa}_0 = \mathbf{G}^t\mathbf{C}^{-\mathbf{1}}(\tilde{\mathbf{P}}^{\mathrm{tot}} - \mathbf{P} - \mathbf{P}^N)\,, \tag{4.8}$$

where $\mathbf{C}_{\boldsymbol{k},\boldsymbol{k}'} = \langle P^N_{\boldsymbol{k}} P^N_{\boldsymbol{k}'} \rangle$ is the covariance matrix for the power spectrum. Since the terms introduced by lensing do not dominate in the noise, we have neglected their contribution in the covariance matrix $\mathbf{C}$. All noise related calculations in the following text will use $\hat{P}$ instead of $\tilde{P}^{\rm tot}$, while the estimator of lensing quantities will always use $\tilde{P}^{\rm tot}$. Therefore the minimum variance estimator for the non-Gaussian source is

$$\hat{\kappa}_0 = (\mathbf{G}^t\mathbf{C}^{-1}\mathbf{G})^{-1}\mathbf{G}^t\mathbf{C}^{-1}(\tilde{\mathbf{P}}^{\rm tot} - \mathbf{P} - \mathbf{P}^N)\,. \tag{4.9}$$

and the variance of the estimator is

$$\mathrm{Var}(\hat{\kappa}_0) = (\mathbf{G}^t\mathbf{C}^{-1}\mathbf{G})^{-1}\,. \tag{4.10}$$

Continuous case

We now consider the continuous case. By definition $\tilde{T}(\boldsymbol{k}) = \tilde{T}_{\boldsymbol{k}}L^3$, $P(\boldsymbol{k}) = P_{\boldsymbol{k}}L^3, G(\boldsymbol{k}) = G_{\boldsymbol{k}}L^3$ and $\mathbf{C}(\boldsymbol{k},\boldsymbol{k}') = \mathbf{C}_{\boldsymbol{k},\boldsymbol{k}'}L^6$. As such, taking the continuum limit leads to

$$\begin{aligned}\hat{\kappa}_0 &= \frac{1}{Q}\int \frac{\mathrm{d}^3k}{(2\pi)^3}\frac{\mathrm{d}^3k'}{(2\pi)^3}\mathbf{C}^{-1}(\boldsymbol{k},\boldsymbol{k}')G(\boldsymbol{k}')(\tilde{P}^{\rm tot}(\boldsymbol{k}) - P(\boldsymbol{k}) - P^N(\boldsymbol{k}))\\ &= \int \frac{\mathrm{d}^3k}{(2\pi)^3}\tilde{P}^{\rm tot}(\boldsymbol{k})\mathcal{F}(\boldsymbol{k}) - V\,, \end{aligned}\tag{4.11}$$

where

$$\mathcal{F}(\boldsymbol{k}) = \frac{1}{Q}\int \frac{\mathrm{d}^3k'}{(2\pi)^3}\mathbf{C}^{-1}(\boldsymbol{k},\boldsymbol{k}')G(\boldsymbol{k}')\,, \tag{4.12}$$

$$Q = \int \frac{\mathrm{d}^3k}{(2\pi)^3}\frac{\mathrm{d}^3k'}{(2\pi)^3}\mathbf{C}^{-1}(\boldsymbol{k},\boldsymbol{k}')G(\boldsymbol{k})G(\boldsymbol{k}')\,, \tag{4.13}$$

$$V = \int \frac{\mathrm{d}^3k}{(2\pi)^3}(P(\boldsymbol{k}) + P^N(\boldsymbol{k}))\mathcal{F}(\boldsymbol{k})\,. \tag{4.14}$$

V is the mean variance of the smoothed temperature field. Note that when the source is Gaussian, the optimal filter is $\mathcal{F}(\boldsymbol{k}) = G(\boldsymbol{k})/P(\boldsymbol{k})^2/Q$. This is the estimator presented in chapter 3. Using $\tilde{T}(\boldsymbol{k})\tilde{T}^\star(\boldsymbol{k}) = (2\pi)^3\delta^{\rm 3D}(0)\tilde{P}^{\rm tot}(\boldsymbol{k}) = L^3\tilde{P}^{\rm tot}(\boldsymbol{k})$ and Parseval's theorem, we can write the calculation in real space instead of Fourier space

$$\begin{aligned}
\hat{\kappa}_0 &= L^{-3}\int \frac{\mathrm{d}^3k}{(2\pi)^3}\tilde{T}(\boldsymbol{k})\tilde{T}^{\star}(\boldsymbol{k})\mathcal{F}(\boldsymbol{k}) - V \\
&= L^{-3}\int \mathrm{d}^3x \tilde{T}_{w1}(\boldsymbol{x})\tilde{T}_{w2}(\boldsymbol{x}) - V \\
&= \tilde{T}_{w1}(\boldsymbol{x})\tilde{T}_{w2}(\boldsymbol{x}) - V\,, \qquad (4.15)
\end{aligned}$$

where the two window functions to smooth the temperature field are the decomposition of $\mathcal{F}$, i.e., $W_1W_2 = \mathcal{F}$. The noise of $\hat{\kappa}_0$ is

$$\mathrm{Var}(\hat{\kappa}_0) = L^{-6}\int \frac{\mathrm{d}^3k}{(2\pi)^3}\int \frac{\mathrm{d}^3k'}{(2\pi)^3}\mathbf{C}(\boldsymbol{k},\boldsymbol{k}')\mathcal{F}(\boldsymbol{k})\mathcal{F}(\boldsymbol{k}')\,, \qquad (4.16)$$

where $\mathbf{C}(\boldsymbol{k},\boldsymbol{k}') = \langle(\hat{P}(\boldsymbol{k}) - P(\boldsymbol{k}))(\hat{P}(\boldsymbol{k}') - P(\boldsymbol{k}'))\rangle$ is the covariance matrix of the power spectrum. Similar to Eq. (4.9) in the discrete case, we have neglected the contribution from lensed terms in the estimate of the covariance matrix $\mathbf{C}$. It is clear that the noise for $\hat{\kappa}_0$ equals Q^{-1}. $\mathbf{C}(\boldsymbol{k},\boldsymbol{k}')$ will be measured from simulations.

Since

$$\begin{aligned}
\tilde{P}(\boldsymbol{k}_\perp, k_\parallel) &= |\mathbf{J}|^{-1}P(\mathbf{J}^{-1}\boldsymbol{k}_\perp, k_\parallel) \\
&\approx (1+2\kappa)[P(k) + \Delta P(\boldsymbol{k})(\kappa + \gamma_1\cos 2\theta_{\boldsymbol{k}_\perp} \\
&\quad +\gamma_2 \sin 2\theta_{\boldsymbol{k}_\perp})]\,, \qquad (4.17)
\end{aligned}$$

where $\theta_{\boldsymbol{k}_\perp}$ is the angle between $\boldsymbol{k}_\perp$ and the transverse coordinate (chapter 3), the first order Taylor expansion gives

$$\begin{aligned}
G^{\gamma_1}(\boldsymbol{k}) &= \Delta P(\boldsymbol{k})\cos 2\theta_{\boldsymbol{k}_\perp}\,, \qquad (4.18) \\
G^{\gamma_2}(\boldsymbol{k}) &= \Delta P(\boldsymbol{k})\sin 2\theta_{\boldsymbol{k}_\perp}\,. \qquad (4.19)
\end{aligned}$$

We will rewrite the equation in spherical coordinates using $\boldsymbol{k} = (k,\theta,\phi)$, $k_\perp^2/k^2 = \sin^2\theta$, and $\theta_{\boldsymbol{k}_\perp} = \phi$.

4.4 Numerical methods

We will calculate the reconstruction noise for sources at redshift: 1.25, 3 and 5. The power spectrum covariance matrix will be measured from N-body simulations. Unless mentioned explicitly, WMAP5 cosmological parameters are used throughout the chapter: $\Omega_m = 0.258, \Omega_\Lambda = 0.742, \Omega_b = 0.0441, \sigma_8 = 0.796, n_s = 0.963, h = 0.719, \tau = 0.087$ (Dunkley et al., 2008).

4.4.1 Simulations

We generate the N-body simulations with the CUBEPM code. The total number of nodes used in simulation is n^3. We used $n = 2$ and 3 in our simulations. The simulations are run on the SUNNYVALE cluster of CITA, which is a Beowulf cluster composed of 200 Dell PE1950 nodes. For each node there are 2 quad core Intel(R) Xeon(R) E5310 @ 1.60GHz processors, 4GB of RAM, and 2 gigE network interfaces.

HI gas is distributed in galaxies. If we ignore the bias between the galaxies and dark matter distributions, we could approximately use dark matter to represent 21-cm sources distributions at these redshifts (the simple toy model in chapter 3) since HI sources are expected to trace dark matter fairly well. We have generated the source distribution with CUBEPM and we output the 3-D particles distribution at redshifts 1.25, 3 and 5. To correctly measure the covariance matrix of power spectrum, we generate 400 independent simulations at $z = 0$ (RH05). At $z = 1.25$, we have run 186 simulations. Because the non-linearlity is less at higher redshift, we have run 100 simulations at $z = 3$ and 90 simulations at $z = 5$. The density field is produced by assigning the mass of particles to nearby grids with Cloud in Cell (CIC) method.

As we will illustrate later, the 3-D Fisher information (as defined in RH05) of the amplitude of source power spectrum has similarity with the reconstruction noise. We thus address the convergence as a function of the number of simulations with a Fisher

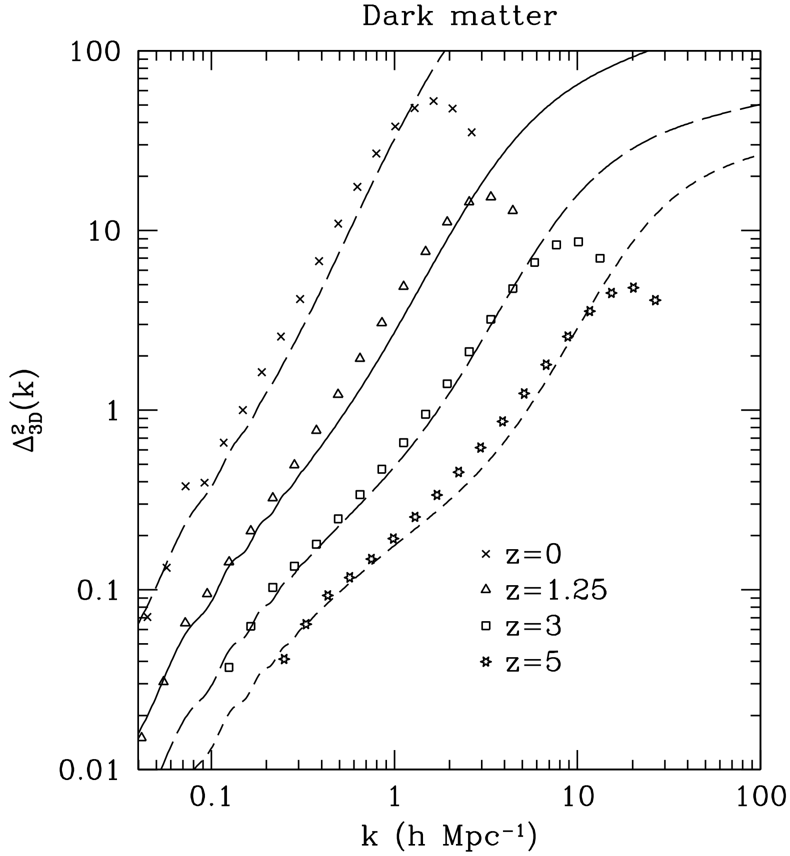

Figure 4.1 Power spectrum at z=0, 1.25, 3, and 5. The data points are from simulations, and the lines are the predicted non-linear power spectra generated by CAMB. The non-linearity increases with lower redshift. The power spectra from simulations at small scales drop due to the resolution limit. The parameters of the simulation are shown at table 4.1.

redshift	$L(h^{-1}Mpc)$	$\Delta^2(2\pi/L)$	n_{grid}	$\Delta^2(k_{Ny})$	k_s	$\Delta^2_{lin}(k_s)$
0	256	0.022	256	197	0.12	0.54
1.25	300	0.0036	512	39.3	0.25	0.31
3.0	100	0.015	1024	40.5	0.47	0.19
5.0	50	0.020	1728	26.5	0.99	0.16

Table 4.1 Source simulation parameters. We choose these simulations to optimize the computation load while maintaining the required resolution. The simulation parameter choice is validated by looking at the convergence of the Fisher information content of the power spectrum in both linear and non-linear scales. We found the upper limit for the non-linear power spectrum at the fundamental modes, which has the same size as the box, is around 0.02; and the lower limit for non-linear power spectrum at the Nyquist frequency k_{Ny}, is around 20. The values of non-linear power spectrum are estimated from CAMB. k_s are the saturation scales of the Fisher information, as shown in Fig. 4.2. We also list the values of the linear power spectra at these saturation scales for all four redshifts.

information plot. We check the values of $\sum_{k,k' \le k_{max}} \bar{\mathbf{C}}^{-1}_{k,k'}$, where $\bar{\mathbf{C}}$ is the covariance matrix of the normalized power spectrum $P(k)/\langle P(k) \rangle$, that is, the Fisher information is inversely proportional to the variance of the amplitude of normalized power spectrum. We have considered runs with different resolutions and box-sizes for each redshift. The box-size needs to be big enough so that not much of the linear modes are cut-off by the limited box-size. On the other hand, the non-linear structure needs to be resolved at small enough scales so that the saturation effect can be seen. We confirmed the convergence at linear scales by comparing the results with the Gaussian prediction, and at non-linear scales by comparing with higher resolution simulations. We found the upper limit for the non-linear power spectrum at the fundamental modes, which has the same size as the box, is around $\Delta^2 \lesssim 0.02$; and the lower limit for non-linear power spectrum at the Nyquist frequency k_{Ny}, is around $\Delta^2 \gtrsim 23$. The simulations parameters we finally chose are given in table 4.1, and the power spectra are shown at Fig. 4.1. The values of the non-linear power spectrum are estimated from CAMB, and thus can be obtained before running simulations. The actual power spectra from N-body simulations are always lower than that at k_{Ny} because of the finite resolution. The sources at $z = 1.25$ are produced with simulations with $300h^{-1}$Mpc box, and 256^3 particles on 512^3 grids. For $z = 3$, because the matter distribution is more linear, we need to increase the resolution to reach the non-linear structures at smaller scales. We did this by using a smaller box with $L = 100h^{-1}$Mpc and more refined grids with $n_{grid} = 1024$. At $z = 5$, the non-linearity is lower, and we choose a $50h^{-1}$Mpc box with 1728^3 grids. For the sources at $z = 3$ and $z = 5$, we have assigned the particles distributions to 512^3 grids though the N-body simulations were performed on finer meshes for the sake of computation efficiency. We also include the Fisher information at z=0, repeating the simulations at RH05. We ran 400 simulations with 256^3 grids, 128^3 particles. The box-size is $256h^{-1}$Mpc, and the cosmological parameters are: $\Omega_m = 0.29, \Omega_\Lambda = 0.71, \Omega_b = 0.046, \sigma_8 = 0.97$, and $h = 0.71$.

The inversion of a large covariance matrix with a few hundred of elements on each dimension can be numerically challenging. Following RH05, we divide the power spectra to $N_b = 20$ bins uniformly distributed in log scale, and calculate the associated band power spectra and their covariance matrix. Subsequently, we invert the $N_b \times N_b$ matrix instead. The Fisher information can be seen in Fig. 4.2. All results are normalized to the values for a $300h^{-1}$Mpc box, i.e., multiply the Fisher information measured by $(300/L)^3$ because larger volume has more independent modes. It is clear that the Fisher information curves grow as k^3 at linear scales, then turn flat at quasi-linear scales ($k \sim 0.5 - 2h\text{Mpc}^{-1}$ for $z = 1.25$), before going up again at non-linear scales with a quasi-Gaussian scaling. There is a gradual evolution of the plateau from $z = 0$ to $z = 5$.

We show the convergence of the Fisher information as a function of both the number of simulations $n_{\rm sim}$ and the resolution in Fig. 4.3. In the top panel, $n_{\rm sim}$ are reduced from 186 to 100, 50, 30 and 20. The relative difference from the 186 runs result is 2% for 100 runs, and 4% for 50 runs. When the number of simulations is less than 30, the deviation is comparable to the Fisher information itself. We emphasize here that the required number

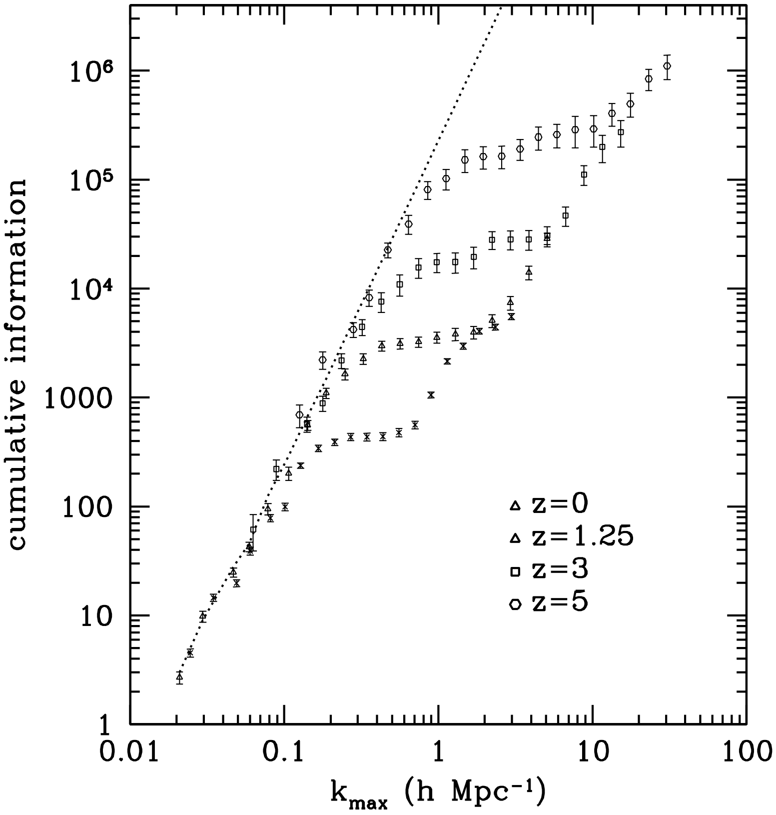

Figure 4.2 The 3-D Fisher information content of the amplitude of power spectrum at $z = 0, 1.25, 3.0, 5.0$ respectively. All results are normalized to the values for a $300h^{-1}$Mpc box. The dashed line is the Gaussian prediction. It is clear that the Fisher information curves grow as k^3 at linear scales, then turn flat at quasi-linear scales ($k \sim 0.5-2h\text{Mpc}^{-1}$ for $z = 1.25$), before going up again at non-linear scales. As expected, there is a gradual evolution of the plateau from $z = 0$ to $z = 5$.

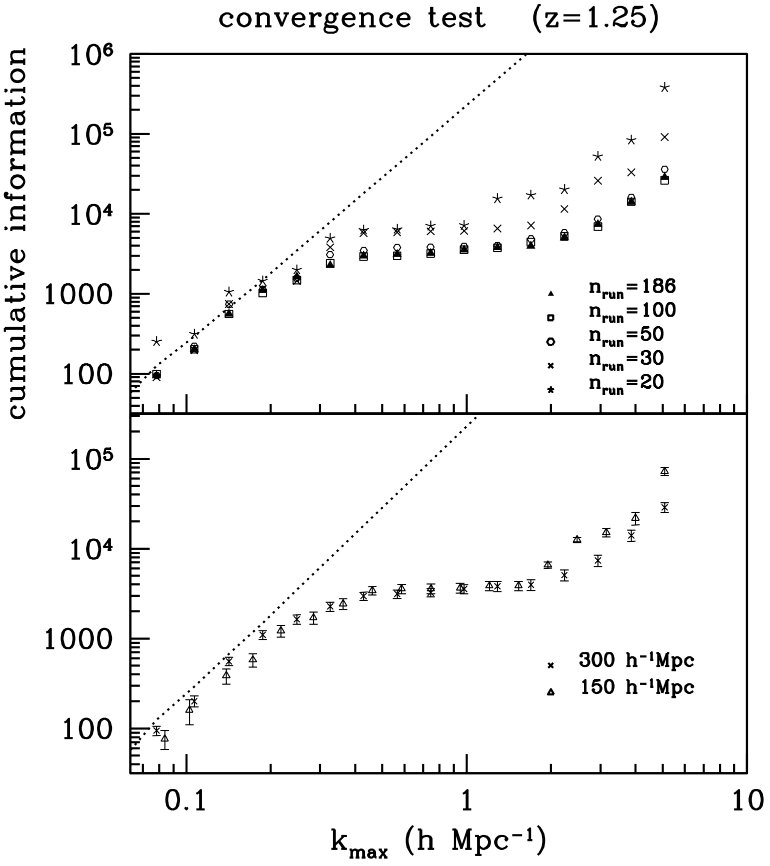

Figure 4.3 Convergence test of the Fisher information as a function of simulations. We increase the number of simulations in the top panel. The relative difference from the 186 runs result is 2% for 100 runs, and 4% for 50 runs. When the number of simulations is less than 30, the deviance is comparable to the Fisher information itself. The bottom panel is for different box-sizes, both types have 186 runs. The error bars are generated by the bootstrap method. The Fisher information from simulations with $L = 150h^{-1}$ Mpc agrees reasonably well with the one from $300h^{-1}$Mpc simulations. As such, the lower limit of power spectrum at fundamental mode should be equal to or higher than the value found in the $150h^{-1}$ Mpc simulations.

of simulations also depends on the number of bins N_b : N_b has to be much smaller than the number of simulations for the covariance matrix to be non-singular. In the bottom panel, we show the convergence of two set of simulations with different box-sizes. Both types have 186 runs. The Fisher information from simulations with $L = 150h^{-1}$ Mpc agrees reasonably well with the one from $300h^{-1}$Mpc simulations. We thus deduce that the lower limit for the power spectrum at fundamental mode should be equal to or higher than the value found in the $150h^{-1}$ Mpc simulations.

4.4.2 Lensing reconstruction

We now study the lensing reconstruction noise that will be measured directly from reconstructed lensing maps with the optimal quadratic estimator described in section 4.3.

Power spectrum covariance matrix

The power spectrum covariance matrix can be written as $\langle P(\boldsymbol{k})P(\boldsymbol{k}')\rangle - \langle P(\boldsymbol{k})\rangle\langle P(\boldsymbol{k}')\rangle$. Because the source is isotropic and stationary, $\langle P(\boldsymbol{k})P(\boldsymbol{k}')\rangle$ can be expressed in a symmetric way $\mathbf{C}(k, k', \cos\theta_{\boldsymbol{k},\boldsymbol{k}'})$, where $\cos\theta_{\boldsymbol{k},\boldsymbol{k}'}$ is the angle between the two vectors on the two shells denoted by k and k'. We can then expand $\mathbf{C}$ in spherical harmonic functions, and because it is independent of ϕ, the expansion can be written as the sum of the Legendre functions

$$\mathbf{C}(k, k', \cos\theta_{\boldsymbol{k},\boldsymbol{k}'}) = \sum_{l=0}^{\infty} \mathbf{C}_l(k, k') P_l(\cos\theta_{\boldsymbol{k},\boldsymbol{k}'}). \tag{4.20}$$

Here l is a even integer because $\mathbf{C}(k, k', \cos\theta_{\boldsymbol{k},\boldsymbol{k}'})$ is an even function of θ. The window function can be written as an integral of the form

$$W^{\kappa}(\boldsymbol{k}) \propto \int \frac{\mathrm{d}^3 k}{(2\pi)^3}' \mathbf{C}^{-1}(\boldsymbol{k}, \boldsymbol{k}')[G_0^{\kappa}(\boldsymbol{k}') + G_2^{\kappa}(\boldsymbol{k}')] \, , \tag{4.21}$$

where

$$G_0^{\kappa}(\boldsymbol{k}) = 2P + 2P'k/3 \, , \tag{4.22}$$

$$G_2^{\kappa}(\boldsymbol{k}) = P'k(-\cos^2\theta_{\boldsymbol{k}'} + 1/3) \, , \tag{4.23}$$

which are proportional to the zeroth order and second order Legendre function $P_0(\cos\theta_{\boldsymbol{k}'})$ and $P_2(\cos\theta_{\boldsymbol{k}'})$. For each $\boldsymbol{k}$, we can always choose the z-axis along it so that $\cos\theta_{\boldsymbol{k},\boldsymbol{k}'}$ is equal to $\cos\theta_{\boldsymbol{k}'}$ without loss of generality. Because the orthogonal property of the Legendre functions, only zeroth and second terms in $\mathbf{C}^{-1}$ remain, all the higher orders cancel out by the integral over $\boldsymbol{k}'$.

In other words, we decompose the optimal window function in two orthogonal components from the zeroth and second order of the covariance matrix expansion terms respectively. For Gaussian sources, the power spectrum modes are uncorrelated with other modes in different directions or on different shells, therefore all $\mathbf{C}_l$ are equal. For non-Gaussian sources, $\mathbf{C}_2$ is about an order of magnitude higher than $\mathbf{C}_0$, so the additional information of lensing obtained by using $\mathbf{C}_2$ is negligible (private communication with Joachim Harnois-Deraps, Harnois-Deraps et al. (2009)).

In Fig. 4.4, we show the optimal non-Gaussian window, and the optimal Gaussian window function in Fourier space. The latter window has an almost power-law slope, while the former one has both positive and negative values due to the complicated behavior of inverse covariance matrix $\mathbf{C}^{-1}$. We use crosses to represent the absolute values of the negative part of the optimal non-Gaussian window.

In this chapter, we will reconstruct the estimator using the zeroth order only. Modes on a shell can be binned as a group first. We could simplify the estimator by calculating its 1-D equivalent, e.g., replacing $P(\boldsymbol{k})$ with $P(k)$. Working in lower dimension facilitates the numerical calculation, especially the inversion of covariance matrix. Rather than working with a 6-D array $\mathbf{C}(\boldsymbol{k},\boldsymbol{k}')$, we now only need to calculate the covariance matrix

$$\mathbf{C}(k,k') = \langle P(k)P(k')\rangle - \langle P(k)\rangle\langle P(k')\rangle\,. \tag{4.24}$$

We treat the modes on each k shell as independent components and the window function corresponding to this shell is calculated from the overall contribution of these modes. We use another trick in the calculation of window function by replacing $\sum \mathbf{C}^{-1}(\boldsymbol{k},\boldsymbol{k}')G(\boldsymbol{k}')$ with $\sum \mathbf{C}_0^{-1}(k,k')G_0(k')$.

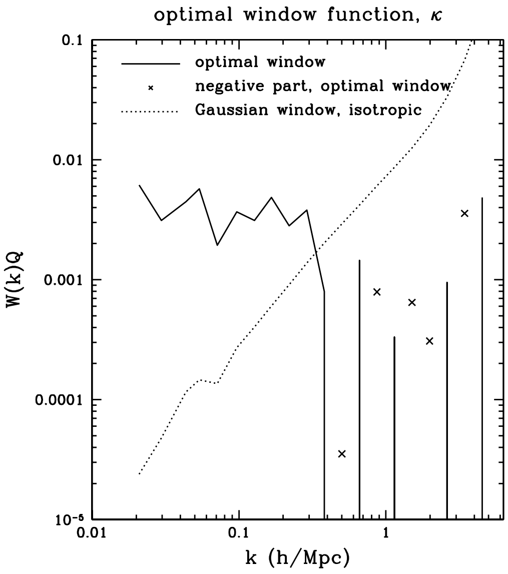

Figure 4.4 The optimal non-Gaussian window, and the optimal Gaussian window function in Fourier space. The latter window has an almost power-law slope, while the former one has both positive and negative values due to the complicated behavior of inverse covariance matrix $\mathbf{C}^{-1}$. We use crosses to represent the absolute values of the negative part of the optimal non-Gaussian window.

When considering the shear, the calculation is very similar. $G^{\gamma 1} = P'k\sin^2\theta\cos 2\phi$ can also be written as the sum of Legendre functions:

$$G_0^{\gamma 1} = 2P'k\cos 2\phi/3\,, \tag{4.25}$$

$$G_2^{\gamma 1} = P'k(1/3-\cos^2\theta)\cos 2\phi\,. \tag{4.26}$$

Therefore only the zeroth and second order mode of covariance remain, and we will also just use the zeroth order in the reconstruction of γ_1. For γ_2, one just need to replace $\cos 2\phi$ by $\sin 2\phi$. The convergence can be calculated from the reconstructed shear. We will call this convergence, γ_E. Similar to chapter 3, we can choose the axis of the coordinate to be parallel to the direction of the mode measured, so that $\gamma_E = \gamma_1$.

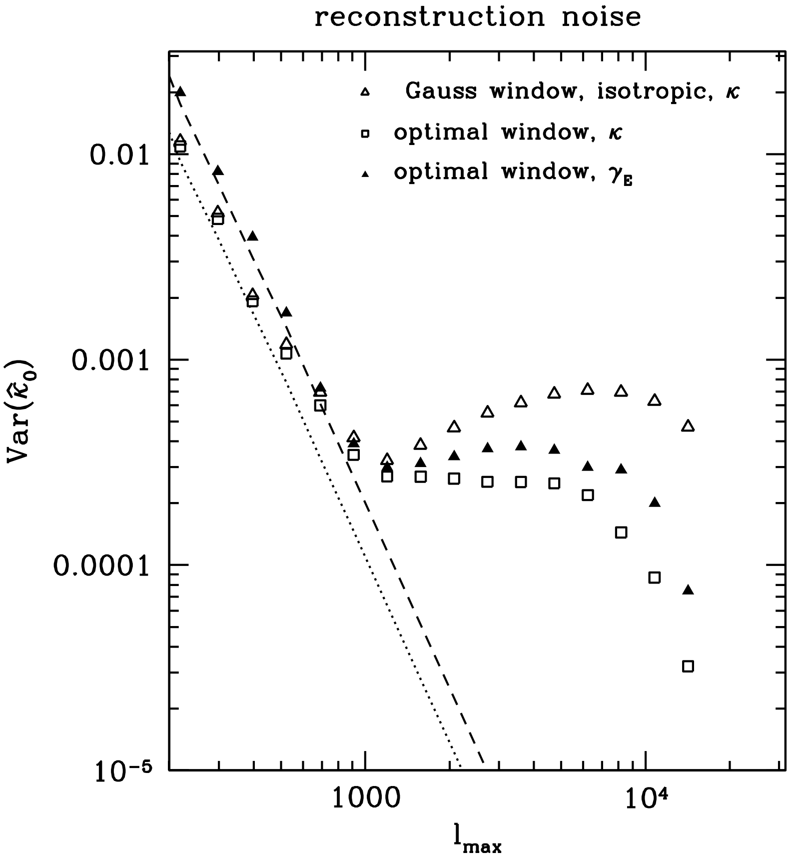

Figure 4.5 The contribution from one box at $z = 1.25$ to the variance of various estimators. The variances are averaged over 186 independent boxes. The reconstruction noises for the Gaussian window, the optimal window for κ, and the optimal window function for γ_E are presented. At linear scales, i.e., when $\ell < 1000$, the two estimators have the same reconstruction noise. The non-Gaussian window has a factor of a few less noise than the Gaussian estimator, and displays a plateau above the saturation scale. The plateau means that the non-Gaussianity of sources caused by the effective number of cells increases slowly below the saturation scales, i.e., the structures are intrinsically correlated below such scales and external factor such as experimental noise will not have much effect. The Gaussian estimator used in chapter 3 does not decrease at saturation scale, and even starts to increase at smaller scales. This is an artifact of the non-optimality of the Gaussian estimator for non-Gaussian sources. We also show the results from Gaussian sources as dashed (shear) and dotted lines (convergence).

Convolution

The optimal estimator contains the inverse covariance matrix $\mathbf{C}^{-1}$, which we measure using around $100 - 200$ sources. For 512^3 grids, there are about 256 modes covering the range from the box-size scale to the Nyquist scale. The inversion of the 256×256 covariance matrix will be unstable, because the number of eigenvalues is no more than the total number of independent samples. As such, we will keep using the 20 bins in our calculation. The binned window function are less optimal, but the difference is a minor disadvantage compared to the stability of the inversion. Another important term in the window function is the gradient of the power spectra. Since the power spectrum is close to a power-law, we calculate the gradient by finding the tangent of the power spectrum at saturation scale k_s measured from the Fisher information on the log-log plot. We find $P'k = n_{\rm NL}P$, and $n_{\rm NL} = -1.6, -1.9$ and -2.1 for $z = 1.25, 3$ and 5 respectively.

On the 512^3 grids, we calculate the window function with the band power, inverse matrix, and gradient. The Nearest Grid Point (NGP) method is used to map these band values to the 3-D grids in Fourier space. From Eq. (4.17), we know that the optimal window function does not only depend on the amplitude k, but is also a function of $k_\perp$. Note that when considering the shear, there is an extra angular dependence on $\theta_{\boldsymbol{k}_\perp}$, the direction on the transverse plane that we need to take into account when computing the window function.

Finally, the 3-D sources are convolved with the optimal window function. Since the window functions are generated in Fourier space, we can simply transform the temperature fields to Fourier space by Fast Fourier Transform (FFT), multiply by the window functions, and transform them back to real space. κ or γ_E maps are integrals of the covariance maps of the two convolved fields along the line of sight. Since we are interested in the zero mode, which is the average value of the covariance, we calculate the reconstruction noise of the combined estimator of κ and γ_E (as in chapter 3), which decrease the noise level. To validate the procedure, we compare the reconstruction noises with the analytical predictions in the Gaussian simulations with the same power spectra as the simulated sources.

4.5 Numerical results and discussion

4.5.1 Reconstruction noise

The optimal estimators were derived in the constant $\kappa(\gamma)$ limit (zero-mode). For the other scales, we need to scale the estimator by a normalization factor $b(\ell)$. The derivation of $b(\ell)$ can be found in appendix A. Not surprisingly, $b(\ell) \leq 1$, and decreases from 1 to 0 when ℓ increases, for any given estimator. For the scales we are interested in – above the characteristic scale ℓ_a– $b(\ell)$ is close to 1, and the noise properties are similar to the zero-mode one. To illustrate the comparison between different estimators at various ℓ with the of zero-mode reconstruction, we treat the experimental noise as a hard cut-off at k_c. $\ell_a \sim \ell_c/2 = \chi(z_s)k_c/2$. This is motivated by the fact that the reconstructed noise is proportional to k^{-3} on linear scales, and the contribution to the reconstruction at $k \leq k_c/2$ are small and can be treated like the zero-mode. The values of ℓ_a are shown in Fig. 4.8 and 4.9.

In Fig. 4.5, the zero-mode κ reconstruction noises from various estimators are presented as functions of various cut-off scales at $z = 1.25$. For this plot, the noises correspond to the contribution from a single simulation box of width $300h^{-1}$Mpc, and is measured over 186 independent boxes. As a reference, we present the reconstruction noises from the Gaussian window, non-Gaussian window for κ and the non-Gaussian window for $\gamma_{\rm E}$. On linear scales, i.e., when $\ell < 1000$, the two estimators of κ have the same reconstruction noise. At smaller scales, the non-Gaussian window reduces the noise by factor of a few compared to the Gaussian estimator, and has a plateau after the saturation scale. The experimental noise smears out the structure in the sources, and regions which are originally independent become correlated. In other words, lower noise level leads to a larger effective number of independent source cells. Similarly, the effective number of cells increases if the experiment is resolved at smaller scales, therefore the reconstructed noise is lower. The plateau means that the non-Gaussianity of sources cause the effective number of cells to increase very slowly below the saturation scale, i.e., the structures are intrinsically correlated below this scale and external factors such as experimental noise will have little effect. The Gaussian estimator we used in chapter 3 also saturates at saturation scales, and even increases at smaller scales. This is an artifact effect coming from the fact that the Gaussian estimator is non-optimal for non-Gaussian sources. Note also that shear has better S/N level than kappa. To compare with the simulated sources, we also show the results from Gaussian predictions calculated using Wick's theorem. In fact, the reconstruction noises in the Gaussian case can be approximated by the following power-law relationship:

$$\begin{aligned}\mathrm{Var}(\kappa_0^{Gauss}) &\approx 3\pi^2 L^{-3} k_c^{-3} (1+\frac{n_{\rm NL}}{3})^{-2}\,, \\ \mathrm{Var}(\gamma_{E0}^{Gauss}) &\approx 45\pi^2 L^{-3} k_c^{-3} n_{\rm NL}^{-2}\,. \end{aligned} \tag{4.27}$$

Note that $n_{\rm NL}$ can not be -3 or 0, otherwise the variance of the estimator is infinity and the lensing signal can not be reconstructed. An intuitive explanation is that, in these cases the variance of temperature field are conserved even after being lensed, therefore lensing maps can not be distinguished from unlensed ones and lensing can not be extracted. This stems from the fact that we only consider the zero order, $\mathbf{C}_0$, in the covariance matrix of matter power spectrum. Lensing can still be solved for $n_{\rm NL} = -3, 0$ cases if $\mathbf{C}_2$ is taken into account in the estimator.

4.5.2 Saturation effects

We define the saturation scale k_s as the wavenumber where the noise from the Gaussian prediction is equal to the average amplitude of reconstructed noise in the plateau. We have discussed the saturation effects of lensing reconstruction in chapter 3, where we

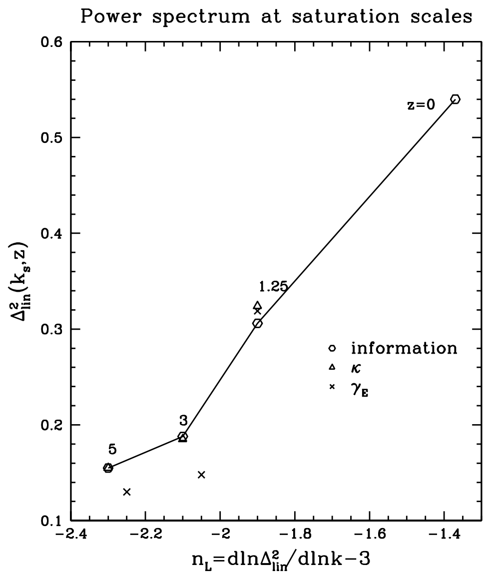

Figure 4.6 Power spectra at the saturation scales of the lensing reconstruction noises for different redshifts. It can be seen that $\Delta^2_{lin}(k_s)$ increases with steeper shape of power spectrum. The points measured from the κ saturation almost overlap with those measured from the Fisher information saturation because the shapes of the source power spectra are very close to power-laws.

found that at $z = 7$ the S/N saturates at the equivalent scale k_s where the non-linear power spectrum of source $\Delta^2(k_s) \sim 0.2$. The corresponding linear power spectrum at k_s is about 0.17, which is consistent with the results at other redshifts shown in table 4.2. It is not clear why the value of $\Delta^2_{lin}(k_s)$ is slightly higher at z=7. It could come from either the fact that there were less sources ($n_{sim} = 20$) or that the estimator was less optimal in the works of chapter 3. Note that since we used a Gaussian estimator in chapter 3 and no $\mathbf{C}^{-1}$ term was involved, the convergence of the reconstructed noise should be much better than that for the Fisher information shown in Fig. 4.3. Since the shape of the matter power spectrum persists except when the non-linear scales shift to larger scales, we expect to see similar saturation effects at lower redshifts. This is confirmed by Fig. 4.5 and later by Fig. 4.6.

As we mentioned earlier, the saturation effect resembles the Fisher information satu-

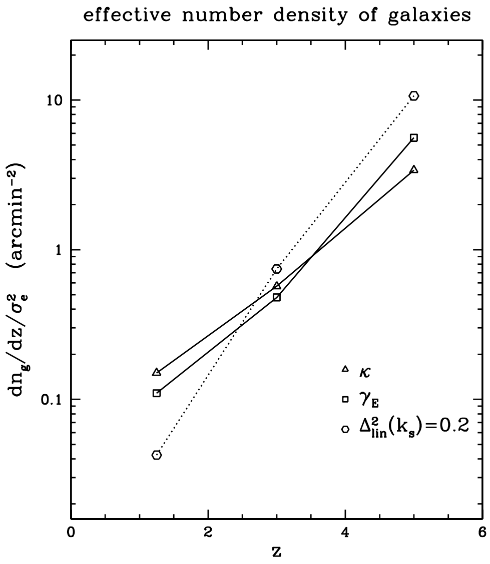

Figure 4.7 The effective number densities of sources at various redshifts. When the redshift is higher, the non-linearity of sources goes down. $dn_g/dz/\sigma_e^2$ increases quickly with redshift. We also plot the number density for a toy model where $\Delta^2_{lin}(k_s) = 0.2$ for all redshifts.

redshift	estimator	$dn_g/dz/\sigma_e^2$(arcmin^{-2})	$L_{cell}(h^{-1}$Mpc)	ℓ_s	k_s	$\Delta^2_{lin}(k_s)$
1.25	κ	0.15	19.	753	0.27	0.32
1.25	γ_E	0.11	21.	753	0.27	0.32
3.0	κ	0.57	13.	2136	0.46	0.19
3.0	γ_E	0.48	14.	1671	0.36	0.15
5.0	κ	3.4	7.0	5657	0.99	0.16
5.0	γ_E	5.6	5.9	4514	0.79	0.13

Table 4.2 Equivalent number densities of galaxies and saturation scales at different redshifts. The characteristic scale of the independent cell L_{cell} is approximately the saturation scale k_s. The number density n_g is calculated with Eq. (4.28). Because n_g itself depends on the box-size of the simulated source and the variance of intrinsic ellipticities σ_e^2, we compare $dn_g/dz/\sigma_e^2$ instead, which increases rapidly with redshift. Note that shear has a higher number density than convergence.

ration effect for the initial amplitude of the 3-D dark matter power spectrum in RH05. Their cumulative Fisher information increases at linear scales. On trans-linear scales, the Fisher information is degenerate with that from larger scales and the cumulative Fisher information does not increase. It is possible to view their Fisher information as a special case of our lensing reconstruction calculation if the gradient term G equals the derivative of the power spectrum with regard to an overall amplitude. Neyrinck et al. (2006) have used halo models to explain this saturation effect: on linear scales, the cumulative Fisher information increases with higher ℓ, as the volume and number of halos do. On trans-linear scales, the 2-halo term first dominates, which washes out the fluctuation in 1-halo term. On non-linear scales, the contribution from small mass halos dominates, and the Fisher information increases again – at less than 1% of the Gaussian information. Our numerical results confirmed the reports of RH05, and our results agree with theirs.

The saturation scales, as well as the non-linear scales, change with redshifts. To illustrate this we investigate the evolution of Δ^2_{lin} at the saturation scale. In Fig. 4.6, we plot $\Delta^2_{lin}(k_s)$ at three redshifts and we give the corresponding numbers in table 4.2. Because the optimal estimator is a function of the gradient of the power spectrum, we will try to see the evolution of $\Delta^2_{lin}(k_s)$ with the gradient of power spectrum at the saturation scale. It can be seen that $\Delta^2_{lin}(k_s)$ increases with steeper shape of power spectrum. The points from κ almost overlap with those from the Fisher information, because the shapes of source power spectra are very close to power-laws.

4.5.3 Effective number density

To describe the information content gained in the 21-cm lensing in a more intuitive way, we define n_g, the equivalent effective surface number density of galaxies which gives the same noise level at the scale where S/N equals to one:

$$\mathrm{Var}(\bar{\hat{\kappa}}) = \frac{\sigma_e^2}{n_g \cdot Area} \approx \frac{1}{N_{cell}}\,. \tag{4.28}$$

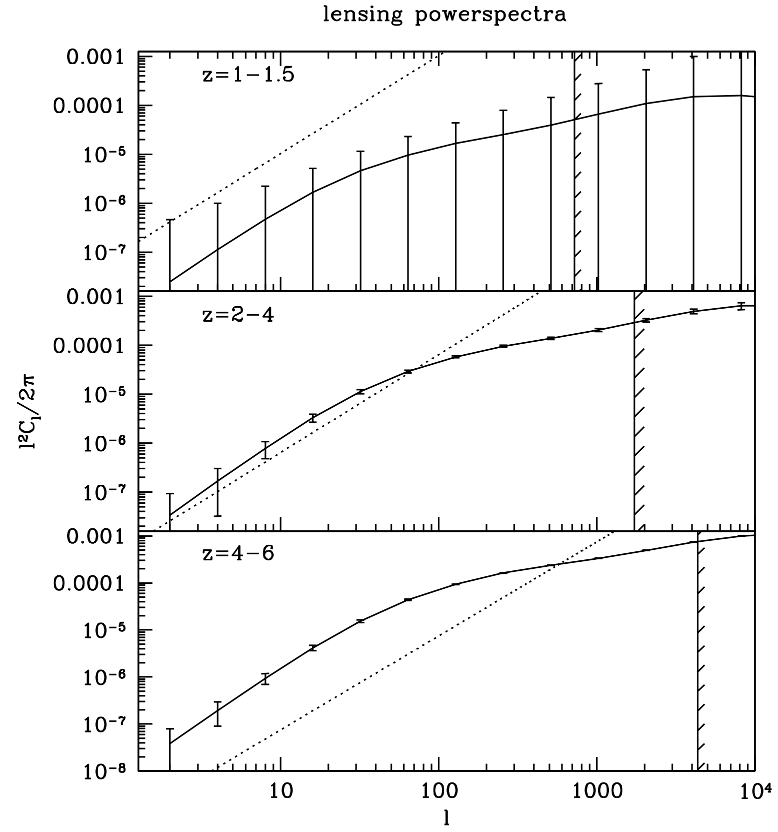

Figure 4.8 The noise power spectra from the effective number of galaxies, compared with the lensing signals. The solid lines are the lensing power spectrum. The error bars are the contribution from both the lensing reconstruction noises and the cosmic variance. The noise dominates over the signal for sources at $z \sim 1-1.5$, becomes comparable to but less than signal in $z \sim 2-4$, and further decreases at $z \sim 4-6$. This corresponds to a sensitivity to lens structures around $z = 0.5, 1.0$ and 1.5 respectively. Note that the noise we display here is only valid for $\ell < \ell_a$, since we calculate n_g in the regime where the noise has similar behavior to zero mode. ℓ_a are plotted as vertical lines here and in Fig. 4.9.

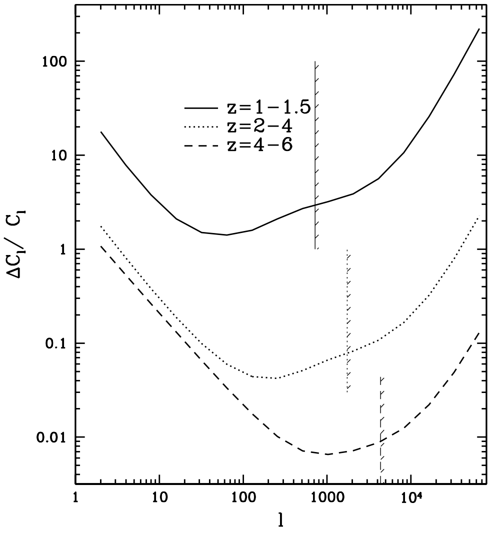

Figure 4.9 The ratio of the error bars and lensing signals. The vertical lines are ℓ_a, and the noise model is valid for $\ell < \ell_a$. For $\ell \gg \ell_a$, a scaling factor $b(\ell)$ need to be applied. The calculation for $b(\ell)$ is done in chapter 3, and we found it decreased from 1 gradually. Therefore the noises at these small scales will be higher than what are shown in this plot. However, these scales are not of interest to us in this chapter. The error bars for $2-4$ bin are at a few percents level for $\ell \sim 20-500$.

Similarly, we can define the number of effective independent cells, N_{cell}, and an effective number density for the noise of γ_E. The characteristic scale of the independent cell is approximately the saturation scale k_s. The cell size is

$$L_{cell} = 2\pi/k_s \times [27(n_{\rm NL}+3)^{-2}/8\pi]^{1/3}\,, \tag{4.29}$$

$$L_{cell} = 2\pi/k_s \times [45 n_{\rm NL}^{-2}/8\pi]^{1/3} \quad \text{for} \quad \gamma_E\,, \tag{4.30}$$

where $\ell_s = k_s\chi(z)$ is the saturation scale that increases with redshift. n_g at different redshifts are shown in table 4.2. Because n_g itself depends on the box-size of the simulated source and the variance of the intrinsic ellipticities σ_e^2, we compute $dn_g/dz/\sigma_e^2$ instead. From the table, we see that the effective number density of galaxies increases rapidly with redshift, and that the shear has higher effective number density than the convergence. We also plot $dn_g/dz/\sigma_e^2$ versus z in Fig. 4.7, where one can see that $dn_g/dz/\sigma_e^2$ increases quickly with redshift. We also show the number density for a toy model where $\Delta^2_{lin}(k_s) = 0.2$ for all redshifts.

In Fig. 4.8, we compare the shot noises $\ell^2 C^N_\ell/2\pi$ from the effective number densities with the lensing signals. The error bars of the lensing power spectra are also plotted, which are estimated by

$$\Delta C_\ell = \sqrt{\frac{2}{(2\ell+1)\Delta\ell f_{sky}}}(C^\kappa_\ell + C^N_\ell)\,, \tag{4.31}$$

where $\Delta\ell \approx \ell/2$ because we used $\ell = 2^n\,(n = 1, 2, ...)$ bins, and consider a half sky survey (Chang et al., 2008) with the fraction of the sky $f_{sky} = 1/2$. Here we assume the noise of κ is Gaussian, i.e., the eight-point function of temperature is Gaussian though the temperature distribution itself is non-Gaussian. Although not obvious, this point is supported by numerically tests. We also neglect the non-Gaussianity of the lensing observables themselves, which is not dominant at $\ell < 500$ but could become an issue at higher ℓ (Dore et al., 2009). It can be seen that the noise dominates in $z \sim 1 - 1.5$, becomes comparable to signal in $z \sim 2 - 4$, and further decreases at $z \sim 4 - 6$. For redshift bin $z \sim 4 - 6$, we assume that $\mathrm{d}n_g/dz$ varies slowly, and use $n_g = dn_g/dz|_{z=5}\Delta z$ with $\Delta z = 2$, which should provide the correct order of magnitude. Ideally, simulated sources should be generated at all redshifts between $z = 4$ and $z = 6$, and the total number density is then the integral of $\mathrm{d}n_g/dz$ over all z. The noise is calculated similarly for $1 - 1.25$ and $2 - 4$ source redshift bins.

The noise is only valid for $\ell < \ell_a \sim \ell_s$, because we calculate n_g in the regime where the noise has a similar behavior to the zero mode. We plot ℓ_a as vertical lines in Fig. 4.9. The full noise calculation can be done similarly to chapter 3, however the zero-mode κ calculation works well enough because we are only interested in large scales (small ℓ's). The noise at small scales will be due to the increased non-Gaussianity at lower redshifts. For $\ell > \ell_a$, $b(\ell)$ is smaller than 1, and the noise will be higher.

The relative error bars are shown in Fig. 4.9. The error bars for $z \sim 2 - 4$ bin are at a few percents level for $\ell \sim 20 - 500$. Schmidt (2008) pointed out that linear

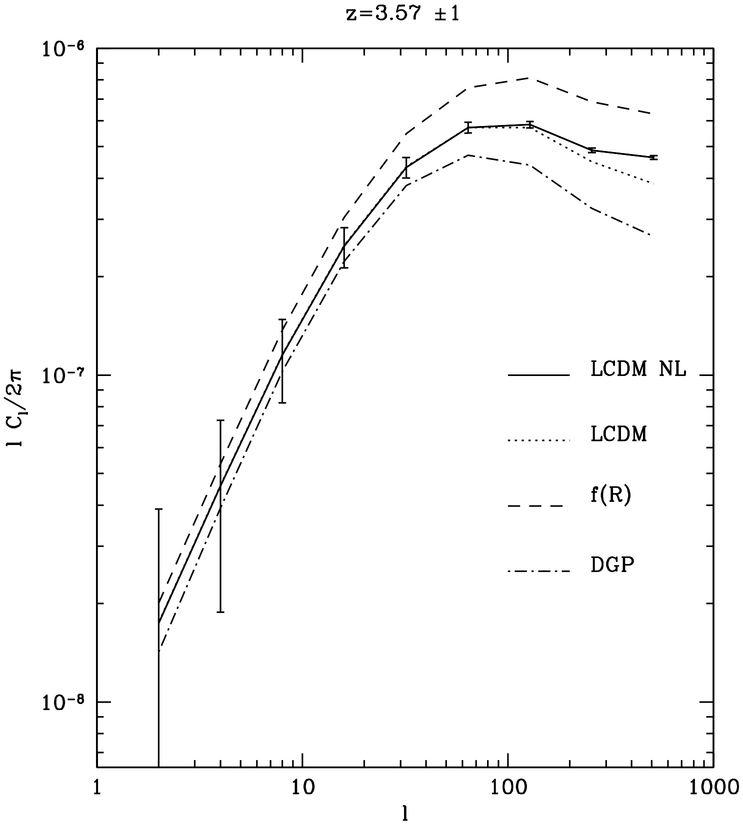

Figure 4.10 The lensing power spectra from different models of dark energy (private communication with Fabian Schmidt). The average redshift is 3.57, where the effective number density of sources is calculated by interpolation with the values from Fig. 4.7. The error bars are calculated in the same way as Fig. 4.8.

scales detection of lensing for source galaxies at $z \sim 1 - 3$ could be a good way to distinguish three modified gravity models from a smooth dark energy: $f(R)$ gravity, the DGP model, and the TeVeS theory. We show the lensing power spectrum from different models of dark energy (private communication with Fabian Schmidt) in Fig. 4.10. The average redshift is 3.57, where the effective number density of sources is calculated by interpolation with the values from Fig. 4.7. The error bars are calculated in the same way as Fig. 4.8. Because of the high precision of the lensing reconstruction, we can expect to see a promising use of 21-cm sources to constrain modified gravity.

Here we discussed how to use 21-cm lensing to constrain dark energy from measuring the lensing power spectrum. One can also investigate a parametrized dark energy model by constraining w_0, w_a using lensing tomography (Dore et al., 2009).

4.6 Summary

In this chapter, we discussed the possibility to constrain dark energy models with the lensing from 21-cm intensity emission at redshifts $1 - 6$. First we derived the optimal quadratic estimator for non-Gaussian sources, which can be constructed numerically from simulations. Then we investigated the reconstruction noise with a large number of simulations, and revealed that there is a saturation scale for the reconstruction noise at all redshifts. We calculated the effective number densities of 21-cm sources and compared the corresponding shot noises to the lensing signals. We conclude that 21-cm sources are promising lensing sources, because they can be observed at high redshift, and may soon be mapped over half the sky (Chang et al., 2008). The lensing reconstructed from 21-cm sources at $z \sim 3$, has a few percent error bars on linear scales at $z \sim 1$, and can be used to constrain modified gravity models.

Chapter 5

Conclusion

Weak gravitational lensing has proven to be an important component of modern precision cosmology. Since the first detections (Van Waerbeke et al., 2000; Bacon et al., 2000; Wittman et al., 2000; Kaiser et al., 2000), cosmic shear has been used to map the distribution of dark matter directly, without using tracers.

Weak lensing is powerful in understanding the accelerating expansion of our universe, which is the most challenging question of contemporary cosmology. The acceleration is widely believed to be caused by dark energy (or vacuum energy) which has negative pressure. Alternatively , modifications to General Relativity on large scales could also lead to cosmic acceleration. Almost all new cosmological surveys are designed to find the answer to the dark energy problem (Albrecht et al., 2006). The ability of constrain the dark energy is often quantified by the Figure of Merit (FoM), which is the area of the 95% contour ellipse on the $w_0 - w_a$ plane. Among the four major experimental methods of dark energy probe, BAO, supernovae, cluster counting and weak lensing, weak lensing is the one with potentially most powerful constraint ability, i.e., highest Figure of Merit.

There are many next generation cosmic shear experiments, e.g., PanSTARRS, LSST, JDEM, DUNE, etc. The optimal strategy for such surveys is still under debate. For example, would a wide and shallow survey be superior to a deep and narrow one? To answer this question, we need to understand the uncertainties of lensing statistics from the theoretical side. In this thesis, I have discussed the lensing S/N for the ideal case, where experimental noise is simplified as a hard cut-off at different scales, i.e., experiments can only detect structures at $k \leq k_c$. In reality, the experimental noise are more complicated, however, most of them are exponential functions of scales or power law (shot noise). Using hard cut-off as noise can help us to separate the effect caused by intrinsic non-Gaussianity of sources and lenses from those produced externally by such as the set up of individual experiments.

To make weak lensing predictions, numerical simulation is a necessary and powerful tool. Weak lensing probes the matter distribution on nonlinear scales, where perturbative methods become inadequate. We therefore have to rely on N-body cosmological simulations to generate realizations of the dark matter distribution. N-body simulations also allow the easy generation of mock lensing data using Monte-Carlo ray-tracing simulations. The lensing signal, e.g., the convergence power spectrum (or cross power spectrum between different source redshift bins), can be obtained from the simulations. Using these simulations, we can compute the statistics of lensing observables, including their average values, but equally importantly their covariance as well. Because the lensing field itself is non-Gaussian, the error bars are correlated at different scales, meaning that simulations must be used. Accurate determination of the error covariance requires a large number of lensing simulations. This is often a non-trivial task (Hamilton et al., 2006).

Inspired by previous works on the Fisher information content of the amplitude of the dark matter power spectrum (Rimes & Hamilton, 2005), in chapter 2 we discussed the Fisher information content of the lensing power spectrum, to investigate whether the Fisher information provided by lensing observables saturates on quasi-linear scales. In the standard ray tracing method to simulate lensing maps, the lensing light cone is constructed from a collection of outputs from N-body simulations. Because of the finite number of simulations, outputs from simulations boxes are usually stacked multiple times, which breaks the Independence of the matter distribution between different redshifts, though the degree of redundancy can be reduced by shifting and rotating these boxes in a random manner. We developed a new method to calculate the covariance of lensing convergence power spectrum, which does not have the spurious correlation problem. The basic idea is to calculate the covariance matrix of matter power spectrum at each redshift with all simulations, and project these covariance matrices to get the covariance matrix of lensing convergence power spectrum. We used this method to measure the covariance matrix from 300 high resolution (1024^3 grids, $200h^{-1}$Mpc box size) N-body simulations. We checked the influence of lensing non-Gaussianity on the Fisher information content of the amplitude of lensing convergence power spectrum. The effect is mild, compared with the 3-D dark matter case. The Fisher information on the amplitude is reduced by about factor of 10 at $\ell \sim 10000$. We can similarly compute the effect of non-Gaussianity on cosmological parameter constraints. In particular, we checked the change of the dark energy FoM caused by the non-Gaussianity. We found the FoM is lowered by a factor of 1.3-1.6 for four types of future lensing surveys. Although the problem has been addressed multiple times by previous works (Cooray & Hu, 2001; Hu & White, 2001; Semboloni et al., 2007; Takada & Jain, 2008; Eifler et al., 2008), this is the first robust examination on the non-Gaussianity issue using simulations.

In chapter 3 and 4, we discussed 21-cm line lensing. Most current research in weak lensing uses galaxies as sources. Another candidate of lensing source could be the 21-cm emission from the neutral hydrogen gas at the epoch of reionization ($z \sim 7-20$). Ongoing and upcoming surveys to explore the universe at the epoch of reionization such as 21CMA, GMRT, SKA, MWA, and LOFAR, makes this choice possible. Galaxies have resolvable shapes, so that the lensing effect can be measured by the statistical correlation of ellipticity of a group of galaxies. However, similar to the CMB, the 21-cm line emission is a diffuse background radiation. Structures in the 21-cm line emission is caused by minihalos (halos with small radius, $r \lesssim 10$Kpc), which are neutral regions and have irregular shapes since the reionization from denser regions might be anisotropic. Nevertheless, the resolution limit of first generation 21-cm experiments (21CMA, GMRT, LOFAR) will not be sufficient ($\ell \ll 10^4$) to resolve these object ($\ell \sim 10^6$) (Pen, 2004; Pen et al., 2008). We therefore treat the 21-cm emission is a diffuse field, rather than as compact objects.

In chapter 3, we study the limits of accuracy for measured weak lensing maps from the diffuse 21-cm radiation at the epoch of reionization using simulations. Lensing distributions can be reconstructed from the observed brightness temperature maps of emission from 21-cm line sources. Because the 21-cm sources trace the neutral hydrogen distribution, which itself traces total gas (i.e., ignore the non-uniform ionization fraction), which in turn traces the dark matter distribution (ignore the galaxy bias), we can simply use dark matter distribution from simulations to represent 21-cm line sources distribution (Metcalf & White, 2007, 2008). This is reasonable for a conceptual understanding of the problems introduced by the non-Gaussianity in the 21-cm sources. Previous optimal quadratic lensing estimators which minimize their variance were derived assuming sources distribution were Gaussian, and reconstructing deflection angles or gravitational potentials (Zahn & Zaldarriaga, 2006). We improve on these estimators by using shear and convergence instead of deflection angles (potentials). This is a generalization of the deflection (potential) estimator, and is more optimal for non-Gaussian sources. The cross power spectrum of shear and convergence is an unbiased estimator of lensing power spectrum which does not require knowledge of the source four-point function. We find that non-Gaussianity provides a limit to the accuracy of weak lensing reconstruction, even if instrumental noise is reduced to zero, i.e., the structures of the sources can be resolved at infinitesimal scales ($k_c \to \infty$). This can be intuitively explained by the fact that different pixels in the sources are not only correlated but also follows non-Gaussian distributions. Therefore the reconstruction noise will be limited by the number of regions which are large enough to be approximately independent to each other, and we will call these regions effective independent cells. We found the reconstruction noise using the optimal estimator for Gaussian sources, actually worsens as the instrumental resolution improved beyond trans-linear scales. The best reconstruction result is equivalent to Gaussian distributed sources with effective independent cell of side length $2.0h^{-1}$ Mpc. Using a source full map from z=10-20, this limiting sensitivity allows mapping of dark matter at a Signal-to-Noise ratio (S/N) greater than 1 out to $\ell \lesssim 6000$, which is better than any other proposed technique for large area weak lensing mapping.

Recently, because of the development of new instruments and the advancement of experimental techniques, much attentions were also attracted to map the 21-cm sources at the post-reionization universe ($z \sim 1-6$) (Peterson et al., 2006; Tegmark & Zaldarriaga, 2008; Wyithe et al., 2008; Chang et al., 2008). The aggregate emission from these 21-cm sources are to be mapped instead of detection of individual galaxies, and this method is called intensity mapping (Pen et al., 2008). The intensity mapping of 21-cm emission for upcoming experiments, such as CHIME, which is aiming at map the 21-cm intensity at $z \sim 1-3$ for half sky in a two year operation time, makes it possible to use them as lensing sources to study the lenses structures at $z \sim 0.5-1.0$, where the cosmic structures are most sensitive to different models of dark energy. Though the main goal of CHIME is to investigate the BAO from the 21-cm distribution, we found the 21-cm line lensing are much more than a by-product science of the survey.

In chapter 4, we studied the possibility of constraining dark energy models with lensing of the post-reionization 21-cm sources ($z \sim 1-6$). To do so, we first built the optimal quadratic estimators (for κ and γ) for non-Gaussian sources. These optimal estimators need to utilize the inverse covariance matrix $\mathbf{C^{-1}}$ of source power spectrum, which can only be obtained numerically from simulations. We applied the estimators on the sources from hundreds of simulations, and investigated the reconstruction noise. We continue to compare the lensing signal and reconstruction noise for ideal experiment where the instrumental noise is approximated as hard cut-off. We revealed that there is a saturation scale for the reconstruction noise at all redshifts. Unlike the Gaussian optimal estimators, where the reconstruction noise rises above trans-linear scales, the reconstruction for the optimal estimators decreases at linear scales, turns flat at trans-linear scales, and drops again at highly non-linear scales. We found the "V" shape of the lensing reconstruction noise with Gaussian optimal estimators are the artifacts of applying non-optimal estimators to sources. We calculated the effective number densities of independent cells, n_g, for the 21-cm sources and compared the shot noises produced by n_g to the lensing signals. We found that although n_g is not very high at $z \sim 1 - 1.5$, it increases quickly with redshifts. The S/N is further boosted dramatically by the large sky coverage. The lensing reconstructed from 21-cm sources at $z \sim 3$, has a few percent error bars at linear scales ($\ell \sim 100$) at $z \sim 1$, and can be used to distinguish the vacuum energy model from alternative gravity models such as $f(R)$ gravity, DGP model and TeVeS theory (Schmidt, 2008).

In summary, we have studied the non-Gaussianity problem in the lensing signal and noise in this thesis, using large numbers of high resolution simulations. We first investigated the effect of non-Gaussian covariance matrix of the lensing convergence power spectrum using a new robust simulation method. Next we discussed the lensing reconstruction problem from 21-cm sources at both the epoch of reionization ($z \sim 7 - 20$), and post-reionization age ($z \sim 1 - 6$). In the former case, we found the 21-cm lensing is competent to galaxy surveys; while in the latter case, we found the upcoming surveys at $z \sim 1 - 3$ with large sky coverage could help to constraint dark energy models at $z \sim 0.5$. Both results describe a promising future for 21-cm lensing.

References

Albrecht A., Bernstein G., Cahn R., Freedman W. L., Hewitt J., Hu W., Huth J., Kamionkowski M., Kolb E. W., Knox L., Mather J. C., Staggs S., Suntzeff N. B., 2006, ArXiv Astrophysics e-prints

Amara A., Réfrégier A., 2007, MNRAS, 381, 1018

Amblard A., Vale C., White M., 2004, New Astronomy, 9, 687

Astier P. et al., 2006, A&A, 447, 31

Bacon D. J., Refregier A. R., Ellis R. S., 2000, MNRAS, 318, 625

Barber A. J., Thomas P. A., Couchman H. M. P., 1999, MNRAS, 310, 453

Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15

Barkana R., Loeb A., 2005, ApJ, 624, L65

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

Baugh C. M., Efstathiou G., 1993, MNRAS, 265, 145

Benjamin J., Heymans C., Semboloni E., van Waerbeke L., Hoekstra H., Erben T., Gladders M. D., Hetterscheidt M., Mellier Y., Yee H. K. C., 2007, MNRAS, 381, 702

Bernardeau F., Colombi S., Gaztañaga E., Scoccimarro R., 2002, Physics Reports, 367, 1

Bernardeau F., van Waerbeke L., Mellier Y., 1997, A&A, 322, 1

Bertschinger E., 1992, in Martinez V. J., Portilla M., Saez D., eds, New Insights into the Universe Vol. 408 of Lecture Notes in Physics, Berlin Springer Verlag, Large-Scale Structures and Motions: Linear Theory and Statistics. pp 65–+

Bertschinger E., 1998, ARA&A, 36, 599

Carroll S. M., Press W. H., Turner E. L., 1992, ARA&A, 30, 499

Chang T.-C., Pen U.-L., Peterson J. B., McDonald P., 2007, ArXiv e-prints, 709

Chang T.-C., Pen U.-L., Peterson J. B., McDonald P., 2008, Physical Review Letters, 100, 091303

Ciardi B., Madau P., 2003, ApJ, 596, 1

Cole S. et al., 2005, MNRAS, 362, 505

Cooray A., 2004, New Astronomy, 9, 173

Cooray A., Hu W., 2001, ApJ, 554, 56

Crittenden R. G., Natarajan P., Pen U.-L., Theuns T., 2001, ApJ, 559, 552

Croft R. A. C., Metzler C. A., 2000, ApJ, 545, 561

Dodelson S., 2003, Modern cosmology. Modern cosmology / Scott Dodelson. Amsterdam (Netherlands): Academic Press. ISBN 0-12-219141-2, 2003, XIII + 440 p.

Dodelson S. et al., 2002, ApJ, 572, 140

Dodelson S., Gaztañaga E., 2000, MNRAS, 312, 774

Dodelson S., Zhang P., 2005, Phys. Rev. D, 72, 083001

Dore O., Lu T., Pen U.-L., 2009, in preparation

Dunkley J., Komatsu E., Nolta M. R., Spergel D. N., Larson D., Hinshaw G., Page L., Bennett C. L., Gold B., Jarosik N., Weiland J. L., Halpern M., Hill R. S., Kogut A., Limon M., Meyer S. S., Tucker G. S., Wollack E., Wright E. L., 2008, ArXiv e-prints

Efron B., Tibshirani R. J., 1993

Efstathiou G., Bond J. R., White S. D. M., 1992, MNRAS, 258, 1P

Efstathiou G., Davis M., White S. D. M., Frenk C. S., 1985, ApJS, 57, 241

Eifler T., Schneider P., Hartlap J., 2008, ArXiv e-prints

Eisenstein D. J. et al., 2005, ApJ, 633, 560

Eisenstein D. J., Hu W., 1998, ApJ, 496, 605

Eisenstein D. J., Hu W., 1999, ApJ, 511, 5

Eisenstein D. J., Zaldarriaga M., 2001, ApJ, 546, 2

Fisher R. A., 1936, Annals Eugen., 179, 7

Hamana T., Colombi S. T., Thion A., Devriendt J. E. G. T., Mellier Y., Bernardeau F., 2002, MNRAS, 330, 365

Hamana T., Mellier Y., 2001, MNRAS, 327, 169

Hamilton A. J. S., Rimes C. D., Scoccimarro R., 2006, MNRAS, 371, 1188

Harnois-Deraps J., Pen U.-L., Dore O., 2009, in preparation

Hartlap J., Simon P., Schneider P., 2007, A&A, 464, 399

Heavens A., Refregier A., Heymans C., 2000, MNRAS, 319, 649

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

Hilbert S., Metcalf R. B., White S. D. M., 2007, MNRAS, 382, 1494

Hirata C. M., Ho S., Padmanabhan N., Seljak U., Bahcall N. A., 2008, Phys. Rev. D, 78, 043520

Hirata C. M., Padmanabhan N., Seljak U., Schlegel D., Brinkmann J., 2004, Phys. Rev. D, 70, 103501

Hirata C. M., Seljak U., 2003, Phys. Rev. D, 68, 083002

Hirata C. M., Seljak U., 2004, Phys. Rev. D, 70, 063526

Hoekstra H., Mellier Y., van Waerbeke L., Semboloni E., Fu L., Hudson M. J., Parker L. C., Tereno I., Benabed K., 2006, ApJ, 647, 116

Hoekstra H., Yee H. K. C., Gladders M. D., Barrientos L. F., Hall P. B., Infante L., 2002, ApJ, 572, 55

Hu W., 2001, ApJ, 557, L79

Hu W., Okamoto T., 2002, ApJ, 574, 566

Hu W., White M., 2001, ApJ, 554, 67

Hui L., Gaztañaga E., Loverde M., 2007, Phys. Rev. D, 76, 103502

Hui L., Gaztañaga E., Loverde M., 2008, Phys. Rev. D, 77, 063526

Iliev I. T., Shapiro P. R., Ferrara A., Martel H., 2002, ApJ, 572, L123

Jain B., 2002, ApJ, 580, L3

Jain B., Seljak U., 1997, ApJ, 484, 560

Jain B., Seljak U., White S., 2000, ApJ, 530, 547

Kaiser N., 1996, ArXiv Astrophysics e-prints

Kaiser N., Wilson G., Luppino G. A., 2000, ArXiv Astrophysics e-prints

Kassim N. E., Lazio T. J. W., Erickson W. C., Crane P. C., Perley R. A., Hicks B., 2000, in Proc. SPIE Vol. 4015, p. 328-340, Radio Telescopes, Harvey R. Butcher; Ed. The Low-Frequency Array (LOFAR): opening a new window on the universe. pp 328–340

Kilbinger M., Benabed K., Guy J., Astier P., Tereno I., Fu L., Wraith D., Coupon J., Mellier Y., Balland C., Bouchet F. R., Hamana T., Hardin D., McCracken H. J., Pain R., Regnault N., Schultheiss M., Yahagi H., 2008, ArXiv e-prints

Komatsu E., Dunkley J., Nolta M. R., Bennett C. L., Gold B., Hinshaw G., Jarosik N., Larson D., Limon M., Page L., Spergel D. N., Halpern M., Hill R. S., Kogut A., Meyer S. S., Tucker G. S., Weiland J. L., Wollack E., Wright E. L., 2008, ArXiv e-prints

Lee J., Pen U.-L., 2001, ApJ, 555, 106

Lewis A., Challinor A., 2006, Physics Reports, 429, 1

Limber D. N., 1954, ApJ, 119, 655

Loverde M., Hui L., Gaztañaga E., 2008, Phys. Rev. D, 77, 023512

Lu T., Pen U.-L., 2008, MNRAS, 388, 1819

Lu T., Pen U.-L., Dore O., 2009, in preparation

Ma Z., Hu W., Huterer D., 2006, ApJ, 636, 21

Maddox S. J., Efstathiou G., Sutherland W. J., Loveday J., 1990, MNRAS, 242, 43P

Mandel K. S., Zaldarriaga M., 2006, ApJ, 647, 719

Massey R. et al., 2007a, ApJS, 172, 239

Massey R. et al., 2007b, Nature, 445, 286

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

Merz H., Pen U., Trac H., 2005, New Astronomy, 10, 393

Metcalf R. B., White S. D. M., 2007, MNRAS, 381, 447

Metcalf R. B., White S. D. M., 2008, ArXiv e-prints, 801

Neyrinck M. C., Szapudi I., 2007, MNRAS, 375, L51

Neyrinck M. C., Szapudi I., Rimes C. D., 2006, MNRAS, 370, L66

Padmanabhan N., Seljak U., Pen U. L., 2002, in eprint arXiv:astro-ph/0210478 Mining Weak Lensing Surveys. pp 10478–+

Peacock J. A., 1999, Cosmological Physics. Cosmological Physics, by John A. Peacock, pp. 704. ISBN 052141072X. Cambridge, UK: Cambridge University Press, January 1999.

Peacock J. A., Dodds S. J., 1996, MNRAS, 280, L19

Pen U., 2004, New Astronomy, 9, 417

Pen U.-L., Chang T.-C., Peterson J. B., Roy J., Gupta Y., Bandura K., 2008, ArXiv e-prints, 804

Pen U.-L., Chang T.-C., Peterson J. B., Roy J., Gupta Y., Hirata C. M., Odegova J., Sigurdson K., 2008, ArXiv e-prints

Pen U.-L., Lu T., van Waerbeke L., Mellier Y., 2003, MNRAS, 346, 994

Pen U.-L., Staveley-Smith L., Peterson J., Chang T.-C., 2008, ArXiv e-prints

Pen U.-L., Van Waerbeke L., Mellier Y., 2002, ApJ, 567, 31

Pen U.-L., Zhang T., van Waerbeke L., Mellier Y., Zhang P., Dubinski J., 2003, ApJ, 592, 664

Peterson J. B., Bandura K., Pen U. L., 2006, ArXiv Astrophysics e-prints

Riess A. G. et al., 2004, ApJ, 607, 665

Rimes C. D., Hamilton A. J. S., 2005, MNRAS, 360, L82

Rimes C. D., Hamilton A. J. S., 2006, MNRAS, 371, 1205

Rudd D. H., Zentner A. R., Kravtsov A. V., 2008, ApJ, 672, 19

Schmidt F., 2008, Phys. Rev. D, 78, 043002

Schneider P., van Waerbeke L., Jain B., Kruse G., 1998, MNRAS, 296, 873

Seljak U., 1998a, ApJ, 503, 492

Seljak U., 1998b, ApJ, 506, 64

Seljak U., Zaldarriaga M., 1996, ApJ, 469, 437+

Seljak U., Zaldarriaga M., 1999, Physical Review Letters, 82, 2636

Seljak U., Zaldarriaga M., 2000, ApJ, 538, 57

Semboloni E., Mellier Y., van Waerbeke L., Hoekstra H., Tereno I., Benabed K., Gwyn S. D. J., Fu L., Hudson M. J., Maoli R., Parker L. C., 2006, A&A, 452, 51

Semboloni E., van Waerbeke L., Heymans C., Hamana T., Colombi S., White M., Mellier Y., 2007, MNRAS, 375, L6

Smith K. M., Cooray A., Das S., Doré O., Hanson D., Hirata C., Kaplinghat M., Keating B., LoVerde M., Miller N., Rocha G., Shimon M., Zahn O., 2008, ArXiv e-prints

Smith K. M., Zahn O., Doré O., 2007, Phys. Rev. D, 76, 043510

Smith R. E., Peacock J. A., Jenkins A., White S. D. M., Frenk C. S., Pearce F. R., Thomas P. A., Efstathiou G., Couchman H. M. P., 2003, MNRAS, 341, 1311

Spergel D. N. et al., 2003, ApJS, 148, 175

Takada M., Jain B., 2008, ArXiv e-prints

Takahashi R., Yoshida N., Takada M., Matsubara T., Sugiyama N., Kayo I., Nishizawa A. J., Nishimichi T., Saito S., Taruya A., 2009, ArXiv e-prints

Tegmark M. et al., 2006, Phys. Rev. D, 74, 123507

Tegmark M., Hamilton A. J. S., Strauss M. A., Vogeley M. S., Szalay A. S., 1998, ApJ, 499, 555

Tegmark M., Taylor A. N., Heavens A. F., 1997, ApJ, 480, 22

Tegmark M., Zaldarriaga M., 2008, ArXiv e-prints

Trac H., Pen U., 2004, New Astronomy, 9, 443

Vafaei S., Lu T., van Waerbeke L., Heymans. C., Semboloni E., Pen U. L., 2008, in preparision

van Haarlem M. P., ed. 2000, Perspectives on Radio Astronomy: Science with Large Antenna Arrays

Van Waerbeke L., Mellier Y., 2003, ArXiv Astrophysics e-prints

Van Waerbeke L., Mellier Y., Erben T., Cuillandre J. C., Bernardeau F., Maoli R., Bertin E., Mc Cracken H. J., Le Fèvre O., Fort B., Dantel-Fort M., Jain B., Schneider P., 2000, A&A, 358, 30

Van Waerbeke L., Mellier Y., Pelló R., Pen U.-L., McCracken H. J., Jain B., 2002, A&A, 393, 369

Weinberg D. H., 1992, MNRAS, 254, 315

White M., 2004, Astroparticle Physics, 22, 211

White M., Hu W., 2000, ApJ, 537, 1

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

Wouthuysen S. A., 1952, AJ, 57, 31

Wyithe J. S. B., Loeb A., Geil P. M., 2008, MNRAS, 383, 1195

Zahn O., Zaldarriaga M., 2006, ApJ, 653, 922

Zaldarriaga M., Seljak U., 1999, Phys. Rev. D, 59, 123507

Zhan H., Knox L., 2004, ApJ, 616, L75

Zhang P., Pen U.-L., 2005, Physical Review Letters, 95, 241302

Zhang T.-J., Pen U.-L., Zhang P., Dubinski J., 2003, ApJ, 598, 818

Appendix A

Normalization and noise of quadratic estimator

In the end of section 3.4, the numerical results of the noise of the estimators are shown. Here we will develop the analytical expression for

$$E_\kappa(\boldsymbol{k}_\perp) = b_\kappa(\boldsymbol{k}_\perp)[\kappa(\boldsymbol{k}_\perp) + n(\boldsymbol{k}_\perp)]\,. \tag{A.1}$$

For shear, a similar relationship holds even though b and n are not isotropic.

When κ is spatially variable,

$$\tilde{T}_b(\boldsymbol{x}) = T_b(\boldsymbol{x}_\perp - \boldsymbol{D}_0 - \boldsymbol{\delta} D(\boldsymbol{x}_\perp), x_\parallel) = T_b(\boldsymbol{x}_\perp - \boldsymbol{D}_0, x_\parallel) - \boldsymbol{\nabla}_\perp T_b(\boldsymbol{x}_\perp - \boldsymbol{D}_0, x_\parallel) \cdot \boldsymbol{\delta} D(\boldsymbol{x}_\perp)\,, \tag{A.2}$$

where $\boldsymbol{D}(\boldsymbol{x}_\perp) = \boldsymbol{d}(\boldsymbol{x}_\perp)\chi(z_s)$, and $\boldsymbol{d}(\boldsymbol{x}_\perp)$ is the deflection angle. Therefore $\kappa = \boldsymbol{\nabla}_\perp \cdot \boldsymbol{\delta} D$. $\boldsymbol{\delta} D(\boldsymbol{x}_\perp) = \boldsymbol{D}(\boldsymbol{x}_\perp) - \boldsymbol{D}_0$, and $\boldsymbol{D}_0 = \boldsymbol{D}(0)$. $\delta D \ll \boldsymbol{x}_\perp$ and can be treated as small perturbations since $\kappa \ll 1$.

Fourier transforming Eq. (3.17),

$$E_\kappa(\boldsymbol{k}_\perp) = \int \mathrm{d}^2 x_\perp E_\kappa(\boldsymbol{x}_\perp) e^{-\mathrm{i}\boldsymbol{k}_\perp \cdot \boldsymbol{x}_\perp} = \frac{1}{L}\int \mathrm{d}^3 x \tilde{T}^\kappa_{\mathrm{w}_1}(\boldsymbol{x}) \tilde{T}^\kappa_{\mathrm{w}_2}(\boldsymbol{x}) e^{-\mathrm{i}\boldsymbol{k}_\perp \cdot \boldsymbol{x}_\perp} - (2\pi)^2 \delta^{\mathrm{2D}}(\boldsymbol{k}_\perp) V_\kappa\ . \tag{A.3}$$

$\tilde{T} = \tilde{T}_b + n$, and noise is uncorrelated with the signal. The product in real space can be represented as a convolution in Fourier space

$$\int \mathrm{d}^3 x e^{-\mathrm{i}\boldsymbol{k}_\perp\cdot\boldsymbol{x}_\perp} \tilde{T}^\kappa_{\mathrm{w}_1}(\boldsymbol{x}) \tilde{T}^\kappa_{\mathrm{w}_2}(\boldsymbol{x}) = \int \frac{\mathrm{d}^3 k'}{(2\pi)^3} \tilde{T}^\kappa_{\mathrm{w}_1}(\boldsymbol{k}'_\perp, k'_\parallel) \tilde{T}^\kappa_{\mathrm{w}_2}(\boldsymbol{k}_\perp - \boldsymbol{k}'_\perp, -k'_\parallel)\,. \tag{A.4}$$

$$\begin{aligned} \tilde{T}_b(\boldsymbol{k}) &= \int \mathrm{d}^3 x e^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} T_b(\boldsymbol{x}_\perp - \boldsymbol{D}(\boldsymbol{x}_\perp), x_\parallel) \\ &= e^{-\mathrm{i}\boldsymbol{k}_\perp\cdot\boldsymbol{D}_0}[T_b(\boldsymbol{k}) - \int \mathrm{d}^3 x e^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} \boldsymbol{\nabla}_\perp T_b(\boldsymbol{x}_\perp, x_\parallel) \cdot \boldsymbol{\delta} D(\boldsymbol{x}_\perp + \boldsymbol{D}_0)]\,, \end{aligned} \tag{A.5}$$

and the lensing introduced term can be further simplified as

$$\begin{aligned} &\int \mathrm{d}^3 x e^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} \boldsymbol{\nabla}_\perp T_b(\boldsymbol{x}_\perp, x_\parallel) \cdot \boldsymbol{\delta} D(\boldsymbol{x}_\perp + \boldsymbol{D}_0) \\ &= \int \mathrm{d}^3 x e^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} T_b(\boldsymbol{x}_\perp, x_\parallel)(\mathrm{i}\boldsymbol{k}_\perp - \boldsymbol{\nabla}_\perp) \cdot \boldsymbol{\delta} D(\boldsymbol{x}_\perp + \boldsymbol{D}_0) \\ &= \int \frac{\mathrm{d}^2 k'_\perp}{(2\pi)^2} T_b(\boldsymbol{k}_\perp - \boldsymbol{k}'_\perp, k_\parallel) \mathrm{i}(\boldsymbol{k}_\perp - \boldsymbol{k}'_\perp) \cdot \boldsymbol{\delta} D(\boldsymbol{k}'_\perp) e^{\mathrm{i}\boldsymbol{k}'_\perp\cdot\boldsymbol{D}_0}\,. \end{aligned} \tag{A.6}$$

The quadratic term in Eq. (A.3) can be written as

$$\begin{aligned} &\int \mathrm{d}^3 x e^{-\mathrm{i}\boldsymbol{k}_\perp\cdot\boldsymbol{x}_\perp} \tilde{T}^\kappa_{\mathrm{w}_1}(\boldsymbol{x}_\perp, x_\parallel) \tilde{T}^\kappa_{\mathrm{w}_2}(\boldsymbol{x}_\perp, x_\parallel) \\ = \; & e^{-\mathrm{i}\boldsymbol{k}_\perp\cdot\boldsymbol{D}_0} \int \frac{\mathrm{d}^3 k'}{(2\pi)^3} W^\kappa_1(\boldsymbol{k}'_\perp, k'_\parallel) W^\kappa_2(\boldsymbol{k}_\perp - \boldsymbol{k}'_\perp, -k'_\parallel) \\ & \times [T_b(\boldsymbol{k}'_\perp, k'_\parallel) T_b(\boldsymbol{k}_\perp - \boldsymbol{k}'_\perp, -k'_\parallel) \\ & \quad - T_b(\boldsymbol{k}_\perp - \boldsymbol{k}'_\perp, -k'_\parallel) \int \frac{\mathrm{d}^2 k''_\perp}{(2\pi)^2} T_b(\boldsymbol{k}'_\perp - \boldsymbol{k}''_\perp, k'_\parallel) \mathrm{i}(\boldsymbol{k}'_\perp - \boldsymbol{k}''_\perp) \cdot \boldsymbol{\delta} D(\boldsymbol{k}''_\perp) e^{\mathrm{i}\boldsymbol{k}''_\perp\cdot\boldsymbol{D}_0} \\ & \quad - T_b(\boldsymbol{k}'_\perp, k'_\parallel) \int \frac{\mathrm{d}^2 k'''_\perp}{(2\pi)^2} T_b(\boldsymbol{k}_\perp - \boldsymbol{k}'_\perp - \boldsymbol{k}'''_\perp, -k'_\parallel) \mathrm{i}(\boldsymbol{k}_\perp - \boldsymbol{k}'_\perp - \boldsymbol{k}'''_\perp) \cdot \boldsymbol{\delta} D(\boldsymbol{k}'''_\perp) e^{\mathrm{i}\boldsymbol{k}'''_\perp\cdot\boldsymbol{D}_0}] \\ & + \mathrm{Noise}\,. \end{aligned} \tag{A.7}$$

Using the relationship that

$$\langle T_b(\boldsymbol{k}'_\perp, k'_\parallel) T_b(\boldsymbol{k}_\perp - \boldsymbol{k}'_\perp, -k'_\parallel) \rangle = (2\pi)^3 \delta^{\mathrm{3D}}(\boldsymbol{k}_\perp, 0) P(\boldsymbol{k}'_\perp, k'_\parallel)\ , \tag{A.8}$$

we found that the expectation value of the first terms and the noise term in Eq. (A.7) can cancel the last term in Eq. (A.3). Note $\delta^{\mathrm{D}}(0) = \lim_{\Delta k\to 0}(\Delta k)^{-1} \sim (L/2\pi)$, and

$W_2^\kappa(\boldsymbol{k}_\perp-\boldsymbol{k}'_\perp,-k'_\parallel)\sim W_2^\kappa(\boldsymbol{k}'_\perp,k'_\parallel)$ since $\delta^{2\mathrm{D}}(\boldsymbol{k}_\perp)$ is nonzero only when $\boldsymbol{k}_\perp=0$. Similarly, the last two terms can be simplified. Both $e^{\mathrm{i}\boldsymbol{k}''_\perp\cdot\boldsymbol{D}_0}$ and $e^{\mathrm{i}\boldsymbol{k}'''_\perp\cdot\boldsymbol{D}_0}$ terms cancel $e^{-\mathrm{i}\boldsymbol{k}_\perp\cdot\boldsymbol{D}_0}$ since $\boldsymbol{k}''_\perp=\boldsymbol{k}_\perp$ and $\boldsymbol{k}'''_\perp=\boldsymbol{k}_\perp$ respectively. The normalization factor

$$\begin{aligned}b_\kappa(\boldsymbol{k}_\perp) &= \frac{2}{k_\perp^2}\int\frac{\mathrm{d}^3k'}{(2\pi)^3}W_1^\kappa(\boldsymbol{k}'_\perp,k'_\parallel)W_2^\kappa(\boldsymbol{k}_\perp-\boldsymbol{k}'_\perp,-k'_\parallel)\\ &\quad\times[(\boldsymbol{k}_\perp-\boldsymbol{k}'_\perp)\cdot\boldsymbol{k}_\perp P(\boldsymbol{k}_\perp-\boldsymbol{k}'_\perp,-k'_\parallel)+\boldsymbol{k}'_\perp\cdot\boldsymbol{k}_\perp P(\boldsymbol{k}'_\perp,k'_\parallel)]\,. \qquad (A.9)\end{aligned}$$

Similarly, replacing W_1^κ, W_2^κ by $W_1^{\gamma_1}, W_2^{\gamma_1}$ ($W_1^{\gamma_2}, W_2^{\gamma_2}$), and $k_\perp^2$ by $k_\perp^2\cos 2\theta_{\boldsymbol{k}_\perp}$ ($k_\perp^2\sin 2\theta_{\boldsymbol{k}_\perp}$), we find the normalization factor for γ_1 (γ_2).

The noise of the estimator can be calculated in the absence of lensing: $\langle|\hat\kappa(\boldsymbol{k}_\perp)|^2\rangle=\langle\hat\kappa(\boldsymbol{k}_\perp)\hat\kappa^\star(\boldsymbol{k}_\perp)\rangle$. Since $\langle|\hat\kappa(\boldsymbol{k}_\perp)|^2\rangle=(2\pi)^2\delta^{2\mathrm{D}}(0)N_\kappa(\boldsymbol{k}_\perp)$ and $\delta^{2\mathrm{D}}(0)=\lim_{\Delta k\to 0}(\Delta k)^{-2}\sim(L/2\pi)^2$, Wick's theorem gives

$$\begin{aligned}&N_\kappa(\boldsymbol{k}_\perp)\\ =\;&\frac{1}{b(\boldsymbol{k}_\perp)^2L}\int\frac{\mathrm{d}^2k'_\perp}{(2\pi)^2}\int\frac{\mathrm{d}k'_\parallel}{(2\pi)}\{P(\boldsymbol{k}_\perp-\boldsymbol{k}'_\perp,-k'_\parallel)P(\boldsymbol{k}'_\perp,k'_\parallel)[W_1^\kappa(\boldsymbol{k}_\perp-\boldsymbol{k}'_\perp,-k'_\parallel)W_2^\kappa(\boldsymbol{k}'_\perp,k'_\parallel)]^2\\ &+P(\boldsymbol{k}_\perp-\boldsymbol{k}'_\perp,-k'_\parallel)P(\boldsymbol{k}'_\perp,k'_\parallel)\mathcal{F}^\kappa(\boldsymbol{k}_\perp-\boldsymbol{k}'_\perp,-k'_\parallel)\mathcal{F}^\kappa(\boldsymbol{k}'_\perp,k'_\parallel)\}\,. \qquad (A.10)\end{aligned}$$

The first term is the convolution of $P(\boldsymbol{k})W_1^\kappa(\boldsymbol{k})^2$ and $P(\boldsymbol{k})W_2^\kappa(\boldsymbol{k})^2$, and the second term is the convolution of $P(\boldsymbol{k})\mathcal{F}^\kappa(\boldsymbol{k})$ with itself. The dimensionless quantity $k_\perp^2N_\kappa(\boldsymbol{k}_\perp)/(2\pi)$ is equivalent to $\ell^2C_\ell^N/(2\pi)$ in the literature.