THE UNIVERSITY *of* EDINBURGH

Cosmic Shear Analysis From Theory to Data

Institute for Astronomy
School of Physics and Astronomy

Marika Asgari

A thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy
to the
University of Edinburgh
June 2015

Lay Summary

The standard theory of gravity is general relativity, which expresses how matter and energy shape the space-time. The light of a far away galaxy travels through space-time which is affected by the presence of large scale structures and the general geometry of the Universe. According to general relativity light travels through the shortest path between two points in space-time. Therefore, the distortions caused by matter and energy to space-time, change the path of light. We can see this effect in the distorted images of galaxies. This is the fundamental idea behind gravitational lensing. A gravitational lens acts very much like an optical lens, and is sometimes called a cosmic telescope. The distortion of galaxy images can be strong, in which case more than one image of the same galaxy will be visible, or weak, where the images are only slightly sheared. These effects can be simulated using a wine glass and a small light source like a candle as in Fig. 1.

In weak gravitational lensing if a galaxy in the background is circular then its gravitationally lensed image is an ellipse. By studying the effects of gravitational lenses on background galaxy images we can deduce the properties of the mass of the lens: how much mass it has and how its mass is distributed.

Gravitational lensing happens in cosmological scales (cosmic shear) because of very large scale structures. These structures typical have a weak gravitational potential and hence have a weak gravitational lensing effect. Analysing gravitational lensing on cosmic scales provides us with a chance to answer fundamental questions about the Universe.

Our current knowledge of cosmology, tells us that the Universe is about 95% dark, which means that only a few percent of the Universe is made up of the “normal“ matter. The rest is divided into a dark matter and dark energy component[1]. About 70% of the Universe is made up of a unknown repulsive energy, known as the dark enerdy, which accelerates the expansion of the Universe. The dark matter on the other hand only reacts to ”normal“ matter via gravity. Therefore, with gravitational lensing we can detect these dark components, as they change the space-time and hence the path of light.

One of the main challenges of modern cosmology is analysing very large data sets. Data compression method have an important role to play in all disciplines of cosmology. In Chapter

[1]There is also a very small radiation component which was significant when the Universe was very young.

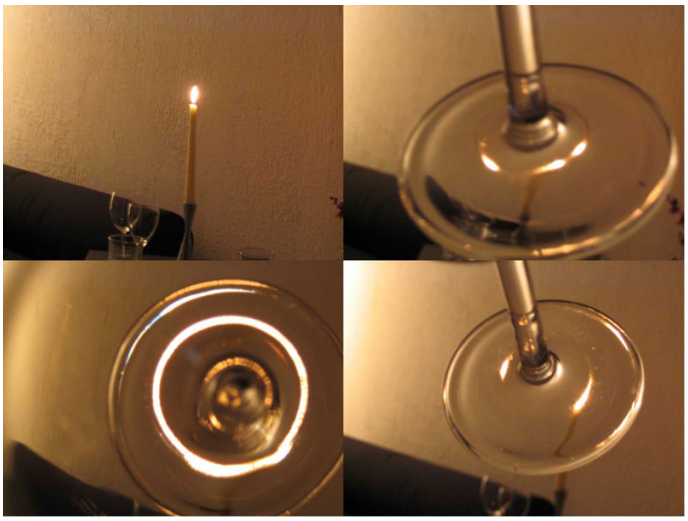

Figure 1: A wine glass and a light source is all you need to simulate gravitational lensing effects. The bottom left photo shows the "Einstein Ring", where the light of the candle is aligned with the center of the glass. The right panels show multiple images of the candle. These phenomena resemble strong lensing effects of massive gravitational lenses on background light sources such as galaxies further away from us.

3 I will show a data compression method and apply it to cosmic shear analysis numerically. I will show that it is robust and should be used for future analysis. My thesis is mainly focused on developing, testing and applying efficient methods of cosmic shear analysis. Cosmic shear signal is contaminated by errors in image analysis and intrinsic alignments of galaxies. The current favored methods are not the most efficient or accurate in separating this signal from the contaminants. I worked on two cosmic shear analysis methods which separate the signal from the contaminants efficiently. In Chapter 4 I will use one these methods together with the compression method from Chapter 3 to analyze cosmic shear data. I will find the likelihood of two cosmological parameters from the data. These are the parameters which cosmic shear is more sensitive to. One of the parameters is called σ_8 which is a measure of how many structures of different sizes exist in the Universe, while the other, $\Omega_{\rm m}$ measures the total amount of matter (both dark and normal matter). I compare my results with other analysis of this kind.

In Chapter 5 I will explore another cosmic shear analysis method using simulated data. This method is fundamentally different from the one used in Chapters 3 and 4, in that it relies on the frequency of the signal. These two methods can be complimentary in data analysis for the future. I test this method and explore its limitations and provide details of developing it in a robust manner which will work for any realistic scenario.

Abstract

One of the most important challenges in cosmology today is understanding the dark matter and dark energy which composite together 95% of the cosmic energy density of the Universe. Weak gravitational lensing by large scale structures is one of the most promising probes for understanding these components and therefore the Universe. The imaging surveys of the future will cover wider fields of view, more accurate redshift estimations and deeper galaxy images. This will leads to smaller statistical errors and tighter parameter constraints. This increased statistical precision will not be satisfactory, however, unless there are trustworthy and accurate methods to analyse the data in order to extract all the information they can offer.

In this thesis I will explore two cosmic shear analysis methods, COSEBIs (Complete Orthogonal Sets of E-/B-Integrals) and PCls (Pseudo Cls). Both of these methods are able to separate gravitational lensing effects (E-modes) from the contaminants (B-modes).

A prominent challenge for cosmological surveys is the estimation of accurate data covariances. N-body cosmological simulations are the most common method used for estimating the covariance, but a large number of simulations with high enough resolution have to be run to estimate accurate data covariances. This number grows with the number of data points used in the analysis. Running cosmological simulations is time consuming and expensive. Therefore, data compression is highly desirable for many disciplines. In Chapter 3 I introduce a method that optimally compresses the number of observables according to their sensitivity to the parameters to be estimated. I then apply this method to COSEBIs (Complete Orthogonal Sets of E-/B-Integrals), an analysis method for weak gravitational lensing, and show that the compressed observables are not sensitive to the choice of the input covariance matrix used to define them.

In Chapter 4 I set up a blind analysis of CFHTLenS[2], the state-of-the-art weak gravitational lensing survey, using COSEBIs and their compressed version. I present a likelihood analysis to estimate cosmological parameters from the data. This is the first time this form of optimised compression has been applied to data. I will also use tomographic redshift bins with COSEBIs and compressed COSEBIs for the first time. The tightest constraints I find for the best cosmological parameter combination is $\sigma_8(\Omega_\mathrm{m}/0.27)^{0.62} = 0.825^{+0.033}_{-0.044}$, which is consistent

[2]http://www.cfhtlens.org/

with previous analysis of CFHTLenS data.

In Chapter 5 I employ Gaussian and lognormal simulated shear fields to explore a flat sky Pseudo Cl analysis pipeline which I have developed. Although, shear two-point correlation functions are insensitive to the mask which are always present on galaxy images, their Fourier counterparts, shear power spectra, are biased by them. Therefore, the effects of masking should be considered in Fourier space analysis of weak gravitational lensing data. I use different masks and propagate errors to cosmological parameters using Fisher analysis to explore the limitations and strengths of Pseudo Cl method.

In the final Chapter I will conclude that the studies presented in this thesis strongly advocates and prefers the use of the presented methods in Chapters 3 and 4, for any future analysis of weak gravitational lensing data. In addition, the compression method in Chapter 3 can also be applied to other cosmological analysis. And finally to avoid biased results Pseudo Cl analysis for the future surveys have to be performed with the considerations detailed in Chapter 5.

Declaration

Except where otherwise stated, the research undertaken in this thesis was the unaided work of the author. Where the work was done in collaboration with others, a significant contribution was made by the author.
Chapter 3 has been published as Marika Asgari & Peter Schneider, A new data compression method and its application to cosmic shear analysis, in Astronomy & Astrophysics. Volume 578, id.A50, 10 pp.

M. Asgari
June 2015

Acknowledgements

I would like to thank my supervisors/collaborators here:
Peter Schneider my collaborator and ex-supervisor for Chapter 3. Catherine Heymans for her support and supervision of the CFHTLenS data analysis in Chapter 4 and for proof reading of this thesis. Andy Taylor, Tom Kitching, Benjamin Joachimi and Alan Heavens for the work on Pseudo Cls in Chapter 5. Andy Taylor for proof reading this thesis. Benjamin Joachimi for providing the shear field simulations used in Chapter 5. Bob Mann and Matthias Bartelmann, my viva examiners, for carefully reading my thesis and their constructive comments.

Contents

Lay Summary i

Abstract iii

Declaration v

Acknowledgements vii

Contents ix

List of figures xii

List of tables xx

1 A Brief Introduction to Cosmology 1
- 1.1 The Principles . . . 1
- 1.2 Friedmann–Lemaître–Robertson–Walker Models . . . 2
  - 1.2.1 Redshift . . . 5
  - 1.2.2 Angular Diameter Distance . . . 5
- 1.3 Structure Formation . . . 6
- 1.4 Cosmological Probes . . . 8
  - 1.4.1 Geometrical Probes . . . 9
  - 1.4.2 Probes of Structure Growth . . . 10
  - 1.4.3 Weak Gravitational Lensing: a Probe of Geometry and Structure Growth 10

2 Weak Gravitational Lensing and Statistics 13
- 2.1 Statistical Tools . . . 13
  - 2.1.1 Likelihood and Chi-square functions . . . 14
  - 2.1.2 Fisher Matrix . . . 16
  - 2.1.3 MCMC . . . 18
- 2.2 Basics of Gravitational Lensing . . . 20
  - 2.2.1 Lens Equation . . . 20
- 2.3 Very Weak Gravitational Lensing . . . 23
- 2.4 Two-Point Statistics . . . 24
  - 2.4.1 E-modes and B-modes . . . 26
- 2.5 General E/B-Mode Decomposition . . . 29
  - 2.5.1 Methods . . . 32

2.6 COSEBIs . . . 33
2.6.1 Filters . . . 33
2.6.2 Covariance . . . 37
2.6.3 COSEBIs: A Literature Review . . . 41

3 Data Compression for Precision Cosmology 45
3.1 Importance of Data Compression . . . 45
3.2 Formalism . . . 47
3.3 Application to COSEBIs . . . 49
3.4 Cosmological Model, Survey Parameters and Covariance . . . 50
3.5 Results . . . 51
3.5.1 Weight functions of Compact COSEBIs . . . 52
3.5.2 Fisher analysis . . . 54
3.6 Band power . . . 56
3.7 Summary and Discussion . . . 61

4 A Cosmic Shear Analysis of CFHTLenS 63
4.1 CFHTLenS: An Introduction . . . 63
4.2 A Blind Analysis: Tests on COSEBIs . . . 65
4.2.1 From Smooth Integration to Noisy Trapezoidal . . . 66
4.3 Clone Simulations . . . 68
4.3.1 Measuring COSEBIs and their Covariance . . . 70
4.3.2 Measuring Compressed COSEBIs . . . 83
4.3.3 MCMC . . . 87
4.4 CFHTLenS Analysis . . . 89
4.4.1 Measured Quantities . . . 90
4.4.2 MCMC Results . . . 102
4.5 Conclusion . . . 104

5 Flat Sky Pseudo-Cl Analysis 107
5.1 Formalism . . . 108
5.1.1 Applying Masks to the Shear Fields . . . 110
5.1.2 Mask Modelling: The Mixing Matrix . . . 112
5.1.3 Mask Smoothing: Apodisation . . . 115
5.2 Theory vs Simulations . . . 116
5.3 Error Propagation . . . 120
5.3.1 Lognormal Moments . . . 120
5.3.2 Fisher Analysis: Results . . . 128
5.4 Conclusion . . . 132

6 Conclusions 137
6.1 Summary . . . 138
6.2 Prospects . . . 141

A An Unsuccessful Compression Method 143

Publications **150**

List of Figures

1 A wine glass and a light source is all you need to simulate gravitational lensing effects. The bottom left photo shows the "Einstein Ring", where the light of the candle is aligned with the center of the glass. The right panels show multiple images of the candle. These phenomena resemble strong lensing effects of massive gravitational lenses on background light sources such as galaxies further away from us. ii

1.1 Classification of different cosmological scenarios in the Ω_m and Ω_Λ plane. In making this plot, radiation is neglected. The solid straight line indicates spatially flat models; any point above or below that line specifies a closed or an open Universe respectively. Since Ω_m is related to an attractive force, it tends to slow down the expansion, on the other hand a positive Ω_Λ is associated with a repulsive force which tries to accelerate the expansion. Furthermore, from the First Friedmann equation we know that dark energy takes over at late epochs, i.e. given a not too large Ω_m, a positive Ω_Λ always means an everlasting expansion. If Ω_m is too large compared to Ω_Λ, its attractiveness wins before the dark energy gets the chance to accelerate the expansion. In the upper left corner of the plot another type of universes exist, which never had a big bang singularity. The universes in the boundary are called loitering universes, because they asymptotically reach a fixed scale factor when their expansion is reversed (adopted from Peacock 1999). 3

1.2 The matter power spectrum for the fiducial set of cosmological parameters values given in Table. 3.1. The solid curve assumes the halo fit formula of Smith et al. (2003), while for the dashed curve the initial power spectrum is linearly extrapolated to $z = 0.1$. At large scales the curves agree, although at small scales (large k) due to non-linear boost, the halo fit formula predicts a larger power. The turn over in the curves depends on the shape parameter, Γ; for a larger Γ the peak moves to smaller scales (larger k). Physically the position of the peak depends on the horizon size at the epoch of matter radiation equality. 9

1.3 Constraints on cosmological parameters adopted from Heymans et al. (2013). σ_8 and Ω_m are the two parameters being measured. Combining CFHTLenS data with BAO, CMB and Hubble parameter priors, results in tighter constraints on these parameters. 11

2.1 A typical lensing configuration. The distances between the observer, the lens and the source planes are calculated in terms of angular diameter distance. The lens changes the position angle of the source from $\boldsymbol{\beta}$ to $\boldsymbol{\theta}$, as seen by the observer (adopted from Bartelmann & Schneider 2001). 20

2.2 A galaxy image is estimated by an ellipse for weak gravitational lensing. The ellipse is characterized by its semi-major and minor axes, a and b, and its orientation, φ with respect to the reference frame. 22
2.3 A spherical over-density results in tangential distribution of shear around it, while an under-density gives rise to a radial distribution as pictured in the above row. The B-modes do not correspond to the Born approximation, they form the divergence free part of the field, as shown in the lower row (adopted from Van Waerbeke & Mellier 2003). 28
2.4 The filter functions of Ring statitics. Z_+ (thick curves) and Z_- (thin curves) are shown for three different ratio values, $\eta = \theta_{\rm max}/\theta_{\rm min}$. $\theta_{\rm max}$ and $\theta_{\rm min}$ are the maximum and minimum angular separation of two galaxies in two rings. These filters are zero outside of their range and oscillate inside (adopted from Schneider & Kilbinger 2007). 33
2.5 The COSEBIs filter function in real space. The $T_\pm^{\rm Log}$ shown in the left panels have fairly evenly spaced roots in logarithmic scales, while the Lin-COSEBIs filters shown in the right panels have their root distributed relatively evenly in their base of support. The angular range considered for these plots is $[1', 400']$. 36
2.6 The weight functions $W_n(\ell)$ shown for three angular ranges, $[1', 400']$, $[20', 400']$ and $[1', 20']$. $W_n(\ell)$ functions are the Hankel transformations of $T_\pm(\vartheta)$ as in Eq. (2.81). The position of the first peak mostly depends on $\vartheta_{\rm max}$ and is rather insensitive to $\vartheta_{\rm min}$. There are two modes of oscillations shown in the zoomed in panels, which depend on $\vartheta_{\rm min}$ and $\vartheta_{\rm max}$. 38
2.7 A 3D representation of non-tomographic covariance of 15 E-mode COSEBIs for an angular range of $[1', 400']$, for a CFHT-like survey. The x and y axes correspond to the elements of the covariance matrix, and the value of the vertical axis shows the value of the covariance of the corresponding element. In the making of these plots the source ellipticity dispersion, σ_ϵ is assumed to be zero. A contour representation of the covariance is shown at its base. 39
2.8 A representation of how a tomographic covariance is built. In this diagram 3 redshift-bins (1,2,3) and 5 COSEBIs modes are assumed to be present. The zoomed in panel shows an instance of one of the covariance building blocks; the numbers 1-5 show the COSEBIs mode considered, for example 15 means the covariance of E_1 and E_5 is shown in that particular block. The numbers on the sides of the matrix show which redshift bins are considered, for instance 12 means the covariance of redshift-bins 1 and 2 is relevant. Due to symmetry, only a part of the covariance elements have to be calculated which are colored pink. 40
2.9 Correlation coefficients of non-tomographic COSEBIs for different angular ranges $[\vartheta_{\rm min}, \vartheta_{\rm max}]$ at $m = 9$, for a CFHT-like survey. Here M, the capital subscripts, are equal to the COSEBIs mode, m. 40
2.10 The correlation coefficients of COSEBIs for an angular range of $[1', 400']$ and 4 redshift bins. In total 15 COSEBIs modes were considered for each graph. 41
2.11 Dependence of three cosmological parameters, σ_8, $\Omega_{\rm m}$, $n_{\rm s}$, on the first five E-mode Log-COSEBIs for a single galaxy redshift distribution. Both the parameters and the E_n values are normalized to their fiducial values. The error-bars show the normalized noise, $\sqrt{C_{nn}}/E_n^f$. The parameters of a future large scale survey such as Euclid or LSST are assumed for this figure with an angular range of $[1', 400']$ (Reproduced from Asgari et al. 2012). 42

2.12 The minimum standard deviation of estimated parameters, σ versus $n_{\rm max}$, the number of COSEBIs modes used in the analysis. Here there are two free parameters, σ_8 and Ω_m, and the others are fixed. For each set of points one of the two parameters is marginalized over. A single redshift distribution and four redshift bins are considered between 0.58 and 1.3, in accordance to the data analysis part of this thesis in Chapter 4 (Reproduced from Asgari et al. 2012). 43

3.1 The filter functions $T^F_{+\mu ij}(\vartheta)$ (top panel) and $W^F_{\mu ij}(\ell)$ (bottom panel) are related to their progenitors via Eq. (3.20). For clarity I show $\vartheta T^F_{+\mu ij}(\vartheta)$ on a logarithmic ϑ-scale. 3 redshift bins and 20 COSEBIs filters defined between $\theta_{\rm min} = 1'$ and $\theta_{\rm max} = 400'$ are considered. Each parameter has a different filter function for each redshift pair bin (z-bin). A comparison between these functions and their progenitors (see Schneider et al. 2010 and Asgari et al. 2012) show that they have considerably fewer structures and oscillations. The filters for each parameter are normalized according to Eq. (3.23). Hence one can compare how effective each z-bin pair is for constraining the parameters. 53

3.2 The figure-of-merit, f, as a function of the number of COSEBIs, $n_{\rm max}$, used. 7 free parameters listed in Table. 3.1, and 8 tomographic redshift bins are considered here. The solid line shows the result for using Log-COSEBIs with the true underlying cosmology. It also represents the maximum information level for a given $n_{\rm max}$. The circles, stars and the Y-shaped symbols represent the f-values for First order, Second order, and their combination $\boldsymbol{E}^{\rm c}$, where $n_{\rm max}$ COSEBIs modes with the fiducial cosmological parameters are utilized in making them. 55

3.3 The figure-of-merit, f, as a function of σ_ϵ. f is normalized by its minimum value which corresponds to using COSEBIs with the correct covariance (the solid line). The intrinsic ellipticity dispersion of galaxies, σ_ϵ, is varied with respect to its true value, 0.3, to show the effects of using a wrong covariance. The markers show the value of f for first order, $\boldsymbol{F}$, second order, $\boldsymbol{S}$ and the combination of both $\boldsymbol{E}^{\rm c}$ CCOSEBIs. . . 57

3.4 The estimated top-hat filter function with $\ell_{\rm min} = 200$ and $\ell_{\rm max} = 400$, from $n_{\rm max}$ COSEBIs filters defined on $1' < \vartheta < 400'$. The changes between using 40 and 80 COSEBIs filters are small, so that no better representation is obtained by using an even higher value of $n_{\rm max}$. 59

3.5 The relative difference, $\delta_{\rm min}$, between the estimated top hat and the input as a function of the number of COSEBIs filters, $n_{\rm max}$, utilized for a few ℓ-ranges. In all cases the saturation level is reached before $n_{\rm max} = 77$. The minimum value of $\delta_{\rm min}$ is shown in Table. 3.5. In general, a higher number of modes are needed for a narrower top-hat filter, which is due to the spillage beyond the observed angular range (see Eq. 3.34). . 60

4.1 A comparison between two methods of finding E-COSEBIs. E^P_n is calculated from Eq. (4.3), while E^ξ_n estimated from Eq. (4.4). E_n with $n = 1 - 7$ are shown here for the fiducial cosmology in Table. 3.1, the redshift parameters of a CFHTLenS-like survey in Table. 4.3 and an angular range of $[1', 400']$. 66

4.2 The convergence of trapezoidal integration to the true value. As n increases due to more oscillations in $T_{\pm n}$ a higher number of bins is required to reach the same level of accuracy. The angular bins are linearly spaced. The true value of E_n is calculated using Eq. (4.3), while the trapezoidal (Trap) value is calculated from Eq. (4.4). . . . 68

4.3 The effect of noise on the estimated E_n for the first 7 modes. Here we only consider uncorrelated noise between the angular bins. n_θ is the number of angular bins for the 2PCFs. The black solid lines are the curves in Fig. 4.2, over plotted here for comparison. The error bars are calculated from the variance between the noise realizations. 69

4.4 Redshift distribution for the galaxies in redshift range $z_B \in [0.58, 1.3]$. The curves show the spectroscopic redshift of the galaxies, with a z_B that belongs to the redshift range given in the plot legends. The y-axis shows the average weighted number of galaxies in a simulated field. The galaxies in the Clone simulations have the same distribution as in the CFHTLenS data. The step like shape of this plot is due to the discrete redshift slices in the simulations. 70

4.5 Covariance matrices of COSEBIs for a single redshift distribution. Three angular ranges, $[1', 40']$, $[1', 20']$ and $[20', 40']$ are considered here. The x/y-axis show the COSEBIs mode considered. The COSEBIs are less correlated for $[20', 40']$ compared to the other cases. 73

4.6 Covariance matrices of COSEBIs for 4 redshift bins. Three angular ranges, $[1', 40']$, $[1', 20']$ and $[20', 40']$ are considered here. The x/y-axis show the redshift bin combination, for example '13' means redshift bins 1 and 3 are relevant. There are 7 COSEBIs modes for each combination. 74

4.7 E_n versus the number of angular bins, N_θ, for Clone simulations. The angular range here is $[1', 40']$. The error bars come from the covariance measured from the simulations. We can see here that using more than about 10^5 angular bins does not contribute to a better convergence to the true value of E_n, since the fluctuations in the measurements are within the statistical errors. 75

4.8 Measured B_n from the Clone simulations for a single redshift bin. Three angular ranges are considered here. For each plot the reduced χ^2 for the B-modes and the angular range is stated, showing that the simulated B-mode is consistent with zero for all the cases tested here. 77

4.9 Measured B_n from Clone simulations for 4 redshift bins. An angular range of $[1', 40']$ is considered here. The reduced χ^2 for the B-modes is stated in the legends, using the full data set with 70 degrees of freedom. Each plot shows the B_n for the combination of redshift bins written on its corner. The reduced χ^2 is also stated for the autocorrelation cases, where only the covariance of the B_n in the respective redshift bin is considered. 78

4.10 E-mode COSEBIs measured from Clone simulations. A single redshift distribution is assumed here. The error bars correspond to the field to field variation of the simulations. The solid lines show the theory values of the E_n for the underlying parameters of the simulations, given in Table. 4.4. Note that the COSEBIs modes are discrete and the theory lines are drawn only for convenience. The reduced χ^2 correspond to the difference between the theory and the simulated values. 79

4.11 E-mode COSEBIs measured from Clone simulations for 4 redshift bins and $[1', 40']$. The error bars correspond to the field to field variation of the simulations. The solid lines show the theory values of the E_n for the underlying parameters of the simulations, given in Table. 4.4. Note that the COSEBIs modes are discrete and the theory lines are drawn only for convenience. The reduced χ^2 shows the correspondence between the theory and the simulations. The value of the reduced χ^2 is also shown for the autocorrelations between the tomographic bins. 80

4.12 E-mode COSEBIs measured from Clone simulations for 4 redshift bins and $[1', 20']$. See the caption for Fig. 4.11 for more details. 81

4.13 E-mode COSEBIs measured from Clone simulations for 4 redshift bins and $[20', 40']$. See the caption for Fig. 4.11 for more details. 82

4.14 The covariance matrix of CCOSEBIs. The elements are measured from the field to field variance of the Clone simulations. Three angular ranges and two redshift binning schemes are considered here. The left panels show the covariance for a single redshift distribution, while the right panels belong to the 4 redshift bin case. The angular range for each plot is shown above it. Comparing the right column of this figure with Fig. 4.6 which shows the COSEBIs covariance for 4 redshift bins, shows how much the compression has reduced the dimensions of the data. 84

4.15 CCOSEBIs measured from the Clone simulations. The CCOSEBIs are linear combinations of the COSEBIs. A single redshift distribution is assumed here. The errors bars correspond to the field to field variation of the simulations, scaled to CFHTLenS area. The solid curves show the theory values of the E_μ^c for the true parameters of the simulations, given in Table. 4.4. Note that the CCOSEBIs modes are discrete and the theory lines are drawn only for convenience. The reduced χ^2 values with five degrees of freedom, are shown for the theory values for each angular range. The B_μ^c and their goodness-of-fit to zero are shown as well. The x-axis shows which derivatives are used for making the CCOSEBIs mode. The B_μ^c for $[20', 40']$ are offset from zero by -1 for visual purposes. 85

4.16 CCOSEBIs measured from the Clone simulations for 4 redshift bins. The CCOSEBIs are linear combinations of the COSEBIs. The errors bars correspond to the field to field variation of the simulations, scaled to CFHTLenS area. The curves show the theory values of the E_μ^c for the true parameters of the simulations, given in Table. 4.4. Note that the CCOSEBIs modes are discrete and the theory lines are drawn only for convenience. The reduced χ^2 values are shown for the goodness-of-fit between the theory and the measured E_μ^c for each angular range. The B_μ^c and their goodness-of-fit to zero are shown as well. The x-axis shows which derivatives are used for making the CCOSEBIs mode. The B_μ^c for $[20', 40']$ are offset from zero by -5 for visual purposes. 86

4.17 Likelihood contours for σ_8 and Ω_m measured from the Clone simulations. A single redshift distribution and 7 E-mode COSEBIs are assumed here. An MCMC algorithm is used for sample the likelihood. The 68% and 95% confidence regions are shown here. The true underlying parameters are shown with the star. 88

4.18 Likelihood contours for σ_8 and Ω_m measured from the Clone simulations for $\theta \in [1', 40']$. The left panel shows the likelihood when COSEBIs is the measured quantity, while the right one shows the same for CCOSEBIs. An MCMC algorithm is used to sample the likelihood. The 68% and 95% confidence regions are shown here. The true underlying parameters are shown with the star. 89

4.19 Likelihood contours for σ_8 and Ω_m measured from the Clone simulations with 4 redshift bins. An MCMC algorithm is used to sample the likelihood distribution. The 68% and 95% confidence regions are shown here. The true underlying parameters are shown with the star. The red contours show the results for COSEBIs versus the black contours for CCOSEBIs. 90

4.20 Measured B_n from the CFHTLenS data for a single redshift bin. Three angular ranges are considered here. For each plot the reduced χ^2 for the estimated B_n and the angular range are stated. Note that the B_n modes are correlated (see Fig. 4.5). 91

4.21 Measured B_n from the CFHTLenS data for 4 redshift bins. An angular range of $[1', 40']$ is considered here. The reduced χ^2 for the B-modes with respect to zero is declared in the legends. Each plot shows the B_n for the combination of redshift bins written on its corner. The reduced χ^2 is also declared for the autocorrelation cases. The data points shown here are correlated as seen in Fig. 4.6. 92

4.22 Measured B_n from the CFHTLenS data for 4 redshift bins. An angular range of $[1', 20']$ is considered here. See the caption of Fig. 4.21 for more details. 93

4.23 Measured B_n from the CFHTLenS data for 4 redshift bins. An angular range of $[20', 40']$ is considered here. See the caption of Fig. 4.21 for more details. 94

4.24 The E_n measured from the CFHTLenS data for a single redshift bin. The solid curves show the theory values of E_n with the best fit cosmological parameters of Kilbinger et al. (2013). The angular range of each panel is noted along with the value of the reduced χ^2 for the fit between the theory and the data. The errors are estimated from the Clone simulations (see Fig. 4.5). 96

4.25 Estimated E_n from the CFHTLenS data for 4 redshift bins. An angular range of $[1', 40']$ is considered here. The solid curves show the value of the theory E_n with parameters deduced from previous CFHTLenS data analysis. The reduced χ^2 values show the agreement between the theory and the estimated values of E_n. Each plot shows the E_n for the combination of redshift bins written on its corner. The reduced χ^2 is also written for the autocorrelation cases. 97

4.26 Estimated E_n from the CFHTLenS data for 4 redshift bins. An angular range of $[1', 20']$ is considered here. See the caption of Fig. 4.25 for more details. 98

4.27 Estimated E_n from the CFHTLenS data for 4 redshift bins. An angular range of $[20', 40']$ is considered here. See the caption of Fig. 4.25 for more details. 99

4.28 Estimated CCOSEBIs from the CFHTLenS data. The CCOSEBIs are linear combinations of the COSEBIs. A single redshift distribution is assumed here. The errors on the measurements come from the Clone simulations. The solid curves show the theory values of the E_μ^c for the cosmological parameters in Kilbinger et al. (2013) (see the start of this section for more details). Note that the CCOSEBIs modes are discrete and the theory values are connected for visual purposes. The reduced χ^2 values are shown for the theory values for each angular range. The B_μ^c and their goodness-of-fit to zero are shown as well. The x-axis shows which derivatives are used for making the CCOSEBIs mode. 100

4.29 Measured CCOSEBIs from CFHTLenS using 4 tomographic redshift bins. See the caption of Fig. 4.28 for more information. 101

4.30 Likelihood contours for σ_8 and Ω_m measured from the data. A single redshift distribution and 7 E-mode COSEBIs are assumed here. An MCMC algorithm is used to sample the likelihood. The 68% and 95% confidence regions are shown here. The best fit values in Kilbinger et al. (2013) are shown with the star. 102

4.31 Likelihood contours for σ_8 and Ω_m measured from the data for $\theta \in [1', 40']$. The left panel shows the likelihood when COSEBIs is the measured quantity, while the right one shows the same for CCOSEBIs. An MCMC algorithm is used to sample the likelihood. The 68% and 95% confidence regions are shown here. The best fit values in Kilbinger et al. (2013) are shown with the star. 103

4.32 Likelihood contours for σ_8 and Ω_m measured from the CFHTLenS data with 4 redshift bins. The black contours show 68% and 95% confidence regions for CCOSEBIs versus the red contours for COSEBIs. An MCMC algorithm is used to sample the likelihood. The best fit values in Kilbinger et al. (2013) are shown with the star. . . 104

5.1 Star and Checkerboard masks. The star mask contains randomly positioned circles with random areas picked from three ranges, 2% from $[0.1', 0.5']$, 5% from $[1', 25']$ and 3% from $[15', 100']$ square arcminutes. The checkerboard mask mimics a CCD gap pattern. Three pixels are masked in the gaps. These two masks are also combined to simulate a more realistic senario. 111

5.2 The general form of the smoothing kernel used for aposiding the masked shear fields. The total number of pixels used, N, determines the support size of the smoothing Kernel. The kernel is symmetric, hence N is an odd number. The functional form of the smoothing kernel is a Gaussian with zero mean and $\sigma = (N-1)/1.5$. 115

5.3 The logarithm of the absolute value of the mixing matrices for star and checkerboard mask. The left panel shows the mixing matrix for the original mask with no apodisation, whereas the right panels shows the same for an apodised mask with $N = 11$ pixels smoothing scale. 25 linear ℓ bins in $[245, 8830]$ are considered here, which corresponds to the widest binning that will be used in this thesis. 116

5.4 The estimated power spectra from the lognormal simulations versus their expected values from theory for the composite mask. Ten of the smallest ℓ bins are merged to make these power spectra ($n_\ell = 10$, $\Delta\ell \approx 180$). The solid curves show the expected theory values and the symbols show the estimated values from the simulations. The blue circles show the recovered C_ℓ after subtracting the noise contribution. The green squares show the remaining B-mode contribution in the recovered C_ℓ. The theory value of the noise power spectrum is shown by the dashed magenta line. The black and red curves show the theory E-/B-mode pseudo power spectra, respectively. The $\times$ and + symbols show their estimated values. The error bars correspond to field-to-field variations between the realisations of the shear fields. The top panel shows the results for the original mask, while for the bottom panel the mask is apodised with a kernel of size 11 pixels. B_{CL} and B_{PCL} show the value of the constant multiplicative bias of the mixing matrix, for the C_ℓ and $\widetilde{C}_\ell$, respectively. This bias is corrected for using the simulations. The effective relative uncovered area of the simulated images is shown via f_{image}. The apodisation covers a larger area of the field. 119

5.5 The ratio of estimated to theory PCls. The top plot shows this ratio for the original star and checkerboard mask which has sharp edges, while the bottom plots shows the same for smoothed masks, with different kernel sizes. The black point show the ratio for the E-mode $\widetilde{C}_\ell$, whereas the red ones correspond to the B-modes. The blue solid line shows where a perfect correspondence between the theory and the estimated quantities from the simulations would lie. The gray area shows the expected cosmic variance. . 121

5.6 The ratio of recovered to theory E-mode C_ℓ. The top plot shows this ratio for the original composite mask, while the bottom plots shows the same for smoothed masks, with different kernel sizes. The black point show the ratio for the C_ℓ before noise correction, and the red points show this ratio after noise correction. The blue solid line shows where a perfect correspondence between the theory and the estimated quantities from the simulations would lie. The light gray area with red edges show the expected cosmic variance for the noise corrected C_ℓ ratio, while the dark gray area shows the cosmic variance before noise correction. 122

5.7 The ratio of bias to error on the estimated parameters with respect to the binning, using C_ℓ as the observable. n_ℓ is the number of original ℓ bins merged to make the new binning. A larger n_ℓ corresponds to a wider bin. The estimated parameters are σ_8 (red solid line) and Ω_m (black dashed line). All the cosmological parameters are fixed to their fiducial values in Table. 5.1 expect for these two parameters. A Fisher analysis is used for estimating the bias and error on the estimated parameter. To find the σ for each parameter the other one is marginalized over. The top panels show this ratio for original masks with no apodisation. The lower panels display $|B|/\sigma$ for the apodised versions of the masks, with Ap1 corresponding to the smallest kernel, Ap2 to the medium size and Ap3 to the largest (see Eq. 5.33). The columns correspond to the three masks used in this thesis. 130

5.8 The ratio of bias to error on the estimated parameters with respect to the binning, using $\widetilde{C}_\ell$ as the observable. See the caption of Fig. 5.7 for more details. 131

5.9 Constraints on σ_8 and Ω_m. Here $n_\ell = 10$ which means $\Delta\ell \approx 180$. The one and two σ contours are drawn around the best fit value for the top panel which uses the Cls. The bottom panel corresponds to the forward modelling which uses PCls as observables. The contours for this panel show the one σ constraints. The contours are shown for Gaussian and lognormal fields. The Gaussian contours are shifted in σ_8 by 0.1 for an easier comparison. The fiducial values of the parameters is shown by + and the shifted value by ×. The black and red contours belong to the original composite mask, while the green and blue ones show the estimated errors for its apodised version. The medium sized kernel with 11 pixels is used here. 133

5.10 Constraints on σ_8 and Ω_m with $n_\ell = 20$ ($\Delta\ell \approx 360$). See the caption of Fig. 5.9 for more details. 134

List of Tables

3.1 The fiducial cosmological parameters consistent with the WMAP 7-years results (Komatsu et al. 2011), and the underlying true parameters consistent with Planck (Planck Collaboration et al. 2014b). The normalization of the power spectrum, σ_8, is the standard deviation of perturbations in a sphere of radius $8h^{-1}$Mpc today. Ω_m, Ω_Λ, and Ω_b are the matter, the dark energy and the baryonic matter density parameters, respectively. w_0 is the dark energy equation of state parameter, which is equal to the ratio of dark energy pressure to its density. The spectral index, n_s, is the power of the initial power spectrum. The dimensionless Hubble parameter, h, characterizes the rate of expansion today. 51

3.2 The parameters of a fiducial large future survey. α, β, and z_0 characterize the total redshift distribution of sources, while z_min and z_max indicate the minimum and the maximum redshifts of the sources use for the cosmic shear analysis. Here A is the survey area in units of deg^2, σ_ϵ is the galaxy intrinsic ellipticity dispersion, and $\bar{n}$ is the mean number density of sources per square arcminute in the field. 51

3.3 The elements of the normalized compression matrix in percentage, $100 \times \mathbf{B}_{\mu n}$, for 1 redshift bin and three parameters. The first column shows the subscript of E^c_μ, where one parameter subscripts belong to the first order statistics, $\boldsymbol{F}$, and double subscripts belong to the second order statistics $\boldsymbol{S}$. The first row shows the value of n. 54

3.4 The normalized compression matrix elements in percentage, $100 \times \mathbf{B}$, for 2 redshift bins and three parameters. The first column shows the subscript of E^c_μ, where one parameter subscripts belong to the first order statistics, $\boldsymbol{F}$, and double subscripts belong to the second order statistics $\boldsymbol{S}$. The absolute values of the columns of $\mathbf{B}$ for each redshift pair decreases rapidly after n=3. 55

3.5 Examples of band-power estimation from 80 COSEBIs for $\xi_\pm(\vartheta)$ in the interval $\vartheta \in [1', 400']$. The first column shows the ℓ-range of the top-hat function $\hat{W}(\ell)$. The rest of the columns show percentage values for minimum relative difference between the estimated and the top hat, δ_min, the absolute minimum (Lower Bound) value for δ_min, δ_LB (see Eq. 3.34), and the relative difference between the estimated band power and its true value, $\delta_\mathrm{band} = |(\hat{E} - E)/\hat{E}|$, for different ℓ weightings, respectively. δ_min and δ_LB values correspond to the $\mathrm{d}\ell\,\ell$ weighting. The values of these quantities for other cases are similar. The cosmological model used here is the fiducial model from Table. 3.1, with one redshift bin. 61

4.1 The redshift bins for tomography. Although CFHTLenS contains information about galaxies with lower redshifts, only galaxies between redshifts 0.58 and 1.3 are considered. The reason behind this choice is that the galaxy intrinsic ellipticity contamination between these redshift bins is insignificant in the analysis (see Heymans et al. 2013). 64

4.2 The prior on the free cosmological parameters and the value of the fixed cosmological parameters. Flat ΛCDM are the only models considered. The free parameters are σ_8 and Ω_m with flat priors given above. Here $\Omega_{tot} = \Omega_m + \Omega_\Lambda$. 64

4.3 The parameters of CFHTLenS-like survey. The true parameters of the CFHTLenS survey will be discussed later in this Chapter. α, β, and z_0 characterize the total redshift distribution of sources, while z_{min} and z_{max} indicate the minimum and the maximum redshifts of the sources used for the cosmic shear analysis. Here A is the survey area in units of deg^2, σ_ϵ is the galaxy intrinsic ellipticity dispersion, and $\bar{n}$ is the effective mean number density of sources per square arcminute in the field. 67

4.4 The value of the cosmological parameters for the Clone simulations (Harnois-Déraps et al. 2012). 76

4.5 The constraints on Σ_8 from CFHTLenS data. Column one shows the data that is used to find the constraints. α is calculated using by fitting to $\Sigma_8 = \sigma_8(\Omega_m/0.27)^\alpha$. The value of the best fit Σ_8 and its associated errors are estimated from the distribution of Σ_8 by fixing α, and finding the values for Σ_8 around its maximum that correspond to 68% of the full area of the distribution. E_n and E^c_μ represent the COSEBIs and CCOSEBIs, respectively. The results are shown for one and four redshift bins with three angular ranges. 103

5.1 The fiducial cosmological parameters consistent with Planck results. The normalization of the power spectrum, σ_8, is the standard deviation of perturbations in a sphere of radius $8h^{-1}$Mpc today. Ω_m, Ω_Λ, and Ω_b are the matter, the dark energy and the baryonic matter density parameters, respectively. w_0 is the dark energy equation of state parameter, which is equal to the ratio of dark energy pressure to its density. The spectral index, n_s, is the power of the initial power spectrum. The dimensionless Hubble parameter, h, characterizes the rate of expansion today. 117

Chapter 1

A Brief Introduction to Cosmology

Understanding the Universe from its earliest time through its evolution is the grand purpose of cosmology. The uniqueness of the Universe makes this task rather challenging which differentiates cosmology from any other science in their methodology. The use of special methods is required to tackle cosmological questions; these are generally statistical tools such as correlation functions and power spectra.

The structure of this chapter is as follows: First the cosmological principles of isotropy and homogeneity are introduced, followed by the Friedmann–Lemaître–Robertson–Walker models, where the cosmological parameters are presented; these parameters will be used in later chapters.

1.1 The Principles

The simplest model one can assume for the Universe is a homogeneous and isotropic model, where matter and energy are distributed homogeneously in space and there is no preferred direction. Hopefully that is, to some extent, the case for our Universe. The observations of the galaxy distribution (assuming that galaxies trace the matter distribution) show no very large scale inhomogeneities or structures, despite the obvious smaller scale clustering. Furthermore, analysis of the Cosmic Microwave Background (CMB, around redshift of $z \sim 1100$) radiation indicates an isotropic Universe. Adding the reasonable Copernican principle, which states that we are not situated in a central or a special position in the Universe, to these findings results in homogeneity.

Modern cosmology started its life with the work of Edwin Hubble, who found the relation between the distance and velocity of galaxies, indicating an isotropic universal expansion called the Hubble expansion law. Mathematically it can be written as:

$$v = H_0 \, D \, , \tag{1.1}$$

where v is the velocity of the galaxies, D is the distance to them and H_0 is called the Hubble parameter and is usually written with respect to a dimensionless parameter, h, in this form:

$$H_0 = 100\,h\ \mathrm{km\,s^{-1}Mpc^{-1}}\ . \tag{1.2}$$

The Hubble parameter is one of the cosmological parameters that is sought after in many of the cosmological probes.

A homogeneous expansion is more conveniently characterized in terms of comoving coordinates, $\mathbf{x}$. They are defined as the physical coordinates of points in space at a fixed cosmic time. Consequently, for a radially expanding universe the physical coordinates at a different time are

$$\mathbf{r}(t) = a(t)\mathbf{x}\ , \tag{1.3}$$

where $a(t)$, the scale factor, depends on time and is scaled to unity for the present time. Taking the derivative of $\mathbf{r}$ with respect to time gives

$$\mathbf{v}(\mathbf{r},t) = \dot{\mathbf{r}} = \dot{a}\mathbf{x} = \frac{\dot{a}}{a}\mathbf{r} =: H(t)\mathbf{r}(t)\ , \tag{1.4}$$

which is a generalization of Hubble law to all times.

1.2 Friedmann–Lemaître–Robertson–Walker Models

Friedmann–Lemaître–Robertson–Walker (FLRW) models correspond to the cosmological models with an isotropic and homogeneous metric of the form

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - a^2(t)[\mathrm{d}\chi^2 + f_K^2(\chi)(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)]\ , \tag{1.5}$$

with t corresponding to the cosmic time; χ is the comoving radial coordinate, θ and ϕ are the angular coordinates and $f_K(\chi)$ is the comoving angular diameter distance which depends on the curvature parameter, K (Schneider 2009a),

$$f_K(\chi) = \begin{cases} K^{-1/2}\sin(K^{1/2}\chi) & (K>0), \\ \chi & (K=0), \\ (-K)^{-1/2}\sinh((-K)^{1/2}\chi) & (K<0). \end{cases} \tag{1.6}$$

Inserting this metric into Einstein field equations results in

$$\left(\frac{\dot{a}(t)}{a(t)}\right)^2 = H^2(t) = \frac{8\pi G}{3}\rho(t) - \frac{Kc^2}{a^2(t)}\ , \tag{1.7}$$

and

$$\frac{\ddot{a}(t)}{a(t)} = -\frac{4\pi G}{3}\left(\rho(t) + 3\frac{P(t)}{c^2}\right) , \tag{1.8}$$

two equations describing the dynamics of the scale factor, a, with respect to the density, ρ, and the pressure, P, of the matter and energy in the Universe. These equations are generally called the Friedmann equations. A more useful form of the first equation is obtained through inserting the following definitions and relations.

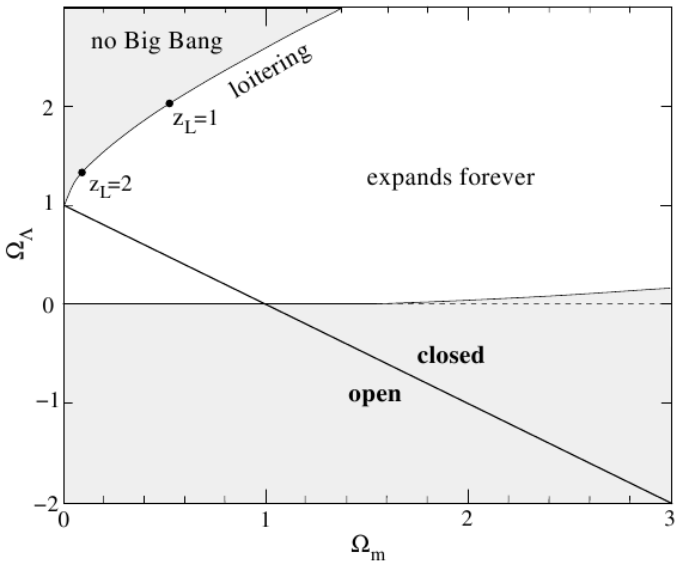

Figure 1.1: Classification of different cosmological scenarios in the Ω_m and Ω_Λ plane. In making this plot, radiation is neglected. The solid straight line indicates spatially flat models; any point above or below that line specifies a closed or an open Universe respectively. Since Ω_m is related to an attractive force, it tends to slow down the expansion, on the other hand a positive Ω_Λ is associated with a repulsive force which tries to accelerate the expansion. Furthermore, from the First Friedmann equation we know that dark energy takes over at late epochs, i.e. given a not too large Ω_m, a positive Ω_Λ always means an everlasting expansion. If Ω_m is too large compared to Ω_Λ, its attractiveness wins before the dark energy gets the chance to accelerate the expansion. In the upper left corner of the plot another type of universes exist, which never had a big bang singularity. The universes in the boundary are called loitering universes, because they asymptotically reach a fixed scale factor when their expansion is reversed (adopted from Peacock 1999).

There are three different matter and energy components in the Universe, that play a role in the dynamics of the scale factor, each with their own equation-of-state (EoS) governing their evolution. They are the pressure-free (collision-less) matter (the baryonic matter is neglected for the moment, because of its very low fraction of the total density), the ultra-relativistic particles, usually called radiation, and the dark energy, an unknown energy component. Dark energy is interpreted in different ways, e.g. the most common interpretation is to assume it is

the constant vacuum energy. In general the EoS is in this form: $P = wc^2\rho$, where w is

$$w = \begin{cases} 0 & \text{for matter,} \\ 1/3 & \text{for radiation,} \\ -1 & \text{for vacuum.} \end{cases} \tag{1.9}$$

The value of w for the dark energy can in general depend on cosmic time and evolve; nevertheless, in the present work we assume that it is constant and to avoid confusion we use w_0 as our EoS parameter.

Inserting $P = wc^2\rho$ into the first law of thermodynamics for an adiabatic system (in cosmology it is generally assumed that the Universe is adiabatic), results in $\rho(t) = \rho(t_0)a(t)^{-3(w+1)}$, a relation between the density and the scale factor.

We define the present critical density,

$$\rho_{\mathrm{cr0}} := \frac{3H_0^2}{8\pi G} , \tag{1.10}$$

which corresponds to the total mean density of the Universe assuming it is spatially flat, $K = 0$. Inserting the ρ_{cr0} definition into Eq. 1.7 yields

$$H^2 = H_0^2 \frac{\rho}{\rho_{\mathrm{cr0}}} - Kc^2 . \tag{1.11}$$

Specializing the above equation to today reveals this relation

$$K = \frac{H_0^2}{c^2}\left(\frac{\rho_0}{\rho_{\mathrm{cr0}}} - 1\right) . \tag{1.12}$$

Defining the density parameters by dividing the density of each matter and energy component at the present time with ρ_{cr0} results in

$$\begin{aligned} &\Omega_{\mathrm{m}} := \frac{\rho_{\mathrm{m0}}}{\rho_{\mathrm{cr0}}} \quad \text{for matter,} \qquad \Omega_{\mathrm{r}} := \frac{\rho_{\mathrm{r0}}}{\rho_{\mathrm{cr0}}} \quad \text{for radiation,} \qquad \Omega_{\Lambda} := \frac{\rho_{\Lambda 0}}{\rho_{\mathrm{cr0}}} \quad \text{for dark energy,} \\ &\Omega_{\mathrm{tot}} := \frac{\rho_{\mathrm{tot0}}}{\rho_{\mathrm{cr0}}} \quad \text{for the total matter and energy content.} \end{aligned} \tag{1.13}$$

A similar procedure is carried out to define the baryonic matter density parameter, Ω_b.

Inserting the evolution relations of each density component and making use of the definitions of the density parameters leads us to

$$H^2 = H_0^2\left[\Omega_{\mathrm{r}}a^{-4} + \Omega_{\mathrm{m}}a^{-3} + (1 - \Omega_{\mathrm{tot}})a^{-2} + \Omega_{\Lambda}a^{-3(w_0+1)}\right] , \tag{1.14}$$

the more practical First Friedmann equation. The radiation density parameter is irrelevant in cosmic shear analysis, since the structures that lay in the path of light coming from source galaxies, live in late epochs where the scale factor is large enough for the radiation to not play a substantial part in the evolution of the Universe.

Finding the values of the parameters is one of the main aims of cosmologists, because these values govern the past and the future of the Universe. Fig. 1.1 illustrates what type of universe is expected from different values of Ω_m and Ω_Λ neglecting radiation, and assuming $w_0 = -1$.

1.2.1 Redshift

Our expanding Universe results in a shift in the spectrum of the objects such as galaxies to larger wavelengths. This effect is called the cosmological redshift (redshift for short from here on). From the Hubble expansion law (Eq. 1.1), we see that the velocity of more distant objects is higher which makes their redshift larger. The following relation between the scale factor and redshift comes from time dilation considerations for a FLRW metric (see Eq. 1.5)

$$1 + z := \frac{\nu_\mathrm{e}}{\nu_\mathrm{obs}} = \frac{\lambda_\mathrm{obs}}{\lambda_\mathrm{e}} = \frac{1}{a(t)} \,, \tag{1.15}$$

where ν_e and ν_obs are the emitted and observed frequencies, and the redshift, z, is defined such that it is zero for the light emitted in the present. The redshift can be interpreted as a Doppler effect for small distances where the metric in Eq. (1.5) can be approximated by a Minkowski metric. The relativistic Doppler effect is given by,

$$1 + z = \sqrt{\frac{v + c}{v - c}} \,, \tag{1.16}$$

where c is the speed of light. For this equation to hold $v < c$ however from Eq. (1.1) we can see that the expansion velocity of an object at $D > c/H_0$ is larger than c, hence for distant objects this approximation fails. The redshift of objects is calculated using photometry and spectroscopy.

1.2.2 Angular Diameter Distance

There are several ways to define distances of objects in a curved space-time. However, in gravitational lensing the most widely-used distance measure is the angular diameter distance, which is a measure of the distance to an object with physical diameter of $2R$, that spans an angle ω on the sky: $D_\mathrm{ang} = 2R/\omega$. Recalling the metric (1.5), and equating $\mathrm{d}s^2 = 4R^2$ and $\mathrm{d}\theta^2 = \omega$, yields

$$D_\mathrm{ang} = a\, f_K(\chi) \,. \tag{1.17}$$

1.3 Structure Formation

The Universe cannot be truly homogeneous on all scales, otherwise we wouldn't be here. To see how we came about in the Universe, the boring, completely homogeneous world model must be changed. The starting point is a perturbation to the models described in the previous section, i.e. assuming that each structure seen today in the Universe has its roots in a very tiny perturbations seeded in the early Universe. The justification for this assumption comes from the Cosmic Microwave Background (CMB) observations from COBE to Planck. The CMB temperature maps these satellites and balloon missions have made, show that the Universe at the epoch of recombination was homogeneous with very small anisotropies of the order of 10^{-5}. The origin of the perturbations is believed to be inflated quantum fluctuations in the very early Universe.

To describe these perturbations, the density contrast, δ is defined as

$$\delta(\mathbf{x}, t) = \frac{\rho(\mathbf{x}, t) - \bar{\rho}}{\bar{\rho}} \ , \tag{1.18}$$

where $\bar{\rho}$ is the mean density.

On average the density contrast is zero in a homogeneous Universe, since the structures are homogeneously distributed. Consequently, cosmologists are compelled to use higher order statistics, e.g. the two-point statistics. In Fourier space two-point statistics are characterized by power spectra, the counterparts of two-point correlation functions in real space. In the case of structure formation the relevant statistic is the matter power spectrum, $P_\delta(k, t)$, written in terms of ensemble averages in Fourier space,

$$\langle \hat{\delta}(\mathbf{k}, t)\, \hat{\delta}^*(\mathbf{k}', t) \rangle = (2\pi)^3 \delta_{\mathrm{D}}(\mathbf{k} - \mathbf{k}')\, P_\delta(k, t) \ , \tag{1.19}$$

where $\hat{\delta}(\mathbf{k}, t)$ is the Fourier transform of the matter density contrast, δ_{D} is the Dirac delta function and k is the magnitude of $\mathbf{k}$, the Fourier mode. The ensemble average is defined as an average over several realizations of the random field. However, only one realization (the Universe) is observable. To solve this problem ergodicity is usually assumed which enables one to substitute the ensemble average by an average over approximately independent patches of one realization. Since there are no characteristic scales for the initial perturbations, the initial power spectrum is in the form of a power law,

$$P_\delta(k) \propto k^{n_{\mathrm{s}}} \ , \tag{1.20}$$

where n_{s} is called the spectral index and has different values regarding different inflationary scenarios. The fluctuations in the Universe are not causally connected prior to horizon crossing, hence Harrison and Zel'dovich proposed that the primordial power spectrum should be scale

invariant, which results in $n_s = 1$. Assuming a larger n_s results in the formation of many black holes at high redshifts (early times) since there is too much power on small scales. On the other hand a small n_s generates an infinite gravitational potential on large scales.

Nowadays, we have good evidence that n_s is slightly smaller than one which is in agreement with single-field inflation theories. The power spectrum changes as the perturbations evolve to form structures of different sizes. To account for this evolution one must consider the interactions between different matter and energy components in the Universe. In general such considerations are divided into two different regimes, the linear for large scales and early times and the non-linear for small scales and late times. In the former regime, the perturbations growth can be described by linear approximations in δ. For example considering scales much smaller than the horizon in the matter dominated era, we can express the behavior of collision-less matter by the collision-less Boltzmann equations. These equations can be linearized in terms of the density contrast and together form a second order, homogeneous differential equation,

$$\frac{\partial^2 \delta}{\partial t^2} + \frac{2\dot{a}}{a}\frac{\partial \delta}{\partial t} - \frac{3H_0^2 \Omega_m}{2a^3}\delta = 0 \, . \tag{1.21}$$

The general solution to this equation is

$$\delta(\mathbf{x}, t) = D_+(t)\Delta_+(\mathbf{x}) + D_-(t)\Delta_-(\mathbf{x}) \, , \tag{1.22}$$

with $D_+(t)$ standing for the growing mode in contrast to $D_-(t)$, the decaying mode. The decaying mode is irrelevant to structure formation, since it dies out rapidly. A similar procedure can be performed for other matter and energy components that cluster (here dark energy is assumed to remain homogeneous), with the inclusion of relativistic considerations for super-horizon perturbations. A more detailed study of structure evolution is found in Schneider (2009a), Peacock (1999), and Dodelson (2003).

From inflation theories we know that at very early times all of the perturbations are larger than the horizon scale. All kinds of super-horizon perturbations grow $\propto a^2$ in the radiation dominated era, and $\propto a$ in the matter dominated era, because physical effects like pressure do not apply to them. As they enter the horizon their physical differences become important. The evolution of a dark matter perturbation that enters the horizon before the matter-radiation equality, freezes until matter becomes dominant, when it grows $\propto a$. Baryons and photons couple as they enter the horizon before recombination epoch. In their case, the pressure plays an important role by opposing gravitational forces: it suppresses the perturbations' growth, and makes them oscillate. The oscillations continue for radiation even after they decouple from baryonic matter, in contrast to the baryons which fall into the more dominant dark matter potential wells and follow their evolution. To account for the scale dependence in Fourier space

the transfer function, $T(k)$, is defined via

$$\frac{\hat{\delta}(\mathbf{k}, a_0)}{\hat{\delta}(\mathbf{k}_\mathrm{L}, a_0)} \equiv T(k) \frac{\hat{\delta}(\mathbf{k}, a_\mathrm{i})}{\hat{\delta}(\mathbf{k}_\mathrm{L}, a_\mathrm{i})} \ , \tag{1.23}$$

where we consider a large scale perturbation, $\hat{\delta}(\mathbf{k}_\mathrm{L})$, which enters the horizon when the Universe is completely matter dominated, and compare its behavior with a perturbation of scale k, at an initial time corresponding to a_i, when both of them are well outside the horizon and at the present, characterized by a_0.

To summarize, the linear power spectrum is

$$P_\delta(k, t) = A \, k^{n_\mathrm{s}} \, T^2(k) \, D_+^2(t) \ , \tag{1.24}$$

where A is the normalization of power spectrum, and is usually expressed in terms of a constant at a given scale. For weak gravitational lensing σ_8, the standard deviation of linear perturbations in a sphere of radius 8 Mpc h^{-1} today, is the most commonly used constant. This scale is roughly the boundary between a linear and an non-linear perturbation, since σ_8 is of order unity. Due to the suppression effects for small scales discussed above the power spectrum has a turn-over scale that depends on the shape parameter, Γ, via the transfer function; assuming cold dark matter (CDM) is the dominant matter form results in

$$\Gamma = \Omega_\mathrm{m} \, h \exp[-\Omega_\mathrm{b}(1 + \sqrt{2h}/\Omega_\mathrm{m})] \ . \tag{1.25}$$

Recent theoretical attempts to solve the non-linear structure evolution have been carried out, for example in Bartelmann et al. (2014). Nevertheless, numerical methods and N-body simulations are the main tools utilized to approximate the behavior of these scales. This especially important when baryons or non-standard forces and particles are included, since the theoretical attempts so far have only been concerned with cold dark matter particles which follow relatively simple physics. In this thesis I use Smith et al. (2003) fitting formula which is based on the "halo model" (see Seljak 2000; Peacock & Smith 2000, for example). Fig. 1.2 shows the linear and non-linear power spectra for the fiducial parameters in Table. 3.1, in Sect. 3.4.

1.4 Cosmological Probes

There are several probes of cosmology which attempt to differentiate between cosmological models using observations. These observations probe the properties of the Universe in different ways. They can be categorized in three groups: geometrical probes, structure growth probes and probes of both.

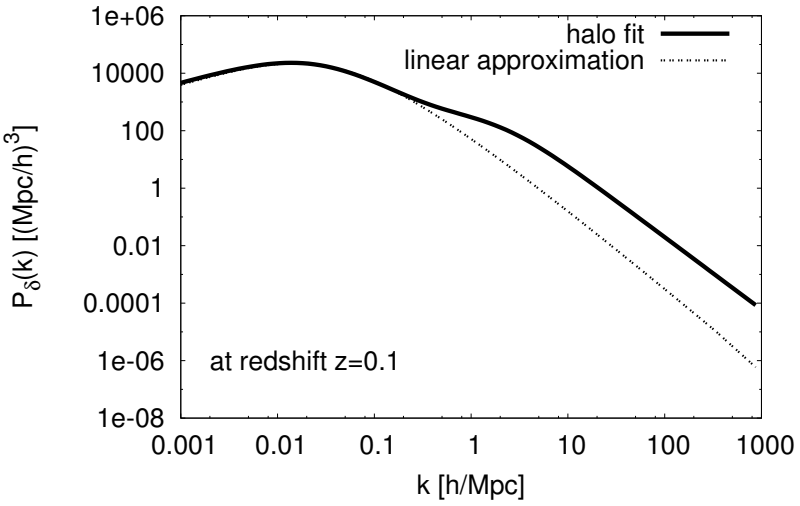

Figure 1.2: The matter power spectrum for the fiducial set of cosmological parameters values given in Table. 3.1. The solid curve assumes the halo fit formula of Smith et al. (2003), while for the dashed curve the initial power spectrum is linearly extrapolated to $z = 0.1$. At large scales the curves agree, although at small scales (large k) due to non-linear boost, the halo fit formula predicts a larger power. The turn over in the curves depends on the shape parameter, Γ; for a larger Γ the peak moves to smaller scales (larger k). Physically the position of the peak depends on the horizon size at the epoch of matter radiation equality.

1.4.1 Geometrical Probes

The geometrical probes of cosmology, are mainly sensitive to distances and the overall geometry of the Universe. These probes rely on the measurement of standard candles or rulers. A standard candle or ruler has a constant[1] luminosity or physical size, which is used to find distances in the Universe. As we saw in Sect. 1.2 the measured cosmological distances depend on the specific metric of the Universe and hence its geometry. The main probes in this category are, cosmic microwave background (CMB), baryonic acoustic oscillations (BAO), and supernova type Ia (SNe) measurements.

CMB and BAO measure the same standard candle in different eras. Before the recombination of electrons and photons, the Universe is in a plasma state, where photons and baryons are coupled together via Thompson scattering, and oscillate. After recombination, the photons decouple from the baryons as the Universe cools down and turns neutral. As a result, the decoupled photons will freeze and their structures will remain almost untouched throughout the history of the Universe. There is an oscillatory mode which has the largest amplitude at the time of recombination, which leaves its signature on the photons and baryons alike. CMB probes this mode in photons, by measuring the small temperature differences in an otherwise blackbody radiation, while BAO does the same with baryons, by measuring clustering of large

[1]This is not always a constant, however, it can be standardized.

scale structures (see Planck Collaboration et al. 2014a; Delubac et al. 2015, for CMB and BAO results, respectively).

Supernovae type Ia are standardized candles. There is a specific relation between their luminosity and the width of their light curve. By observing these supernovae one can measure a luminosity distance which is also dependent on the metric of the Universe. The first evidence for an accelerated expansion of the Universe was detected using SNe (see for example Riess et al. 2007, and references therein).

1.4.2 Probes of Structure Growth

The probes of structure growth, are sensitive to the large scale structures distribution and their abundance. They use luminous matter as tracers of the total matter distribution, although, dark matter is main component of the total matter distribution by far. Luminous matter qualitatively follows the same pattern as the dark matter, nevertheless, the details of their distributions are different. Aside from biases between the luminous and dark matter abundance distributions, there are stochastic processes which result in differences between these two. Redshift surveys probe the three dimensional distribution of structure in the sky. This enables them to study the evolution of structure through redshift (cosmic time). These probes can measure the power spectrum of the matter distribution, which depends on the cosmological model (see for example Peacock et al. 2001; Tegmark et al. 2004, and references therein).

1.4.3 Weak Gravitational Lensing: a Probe of Geometry and Structure Growth

In this thesis I will develop and test methods for weak gravitational lensing (cosmic shear) and apply them on simulated and real data. Weak gravitational lensing is a probe of both the geometry and structure growth of the Universe. It is sensitive to the gravitational effects of matter and the overall metric of the Universe. In a weak gravitational lensing analysis the light from a background source such as a galaxy, is measured. The gravitational potential of these structures distort light, which results in distorted images of background galaxies. The light travel path, also depends on the curvature of the Universe and the evolution of structures. Light propagates along null geodesics hence a change in the geometry of space-time leads to a change in the path of light. The overall metric of the Universe and the local inhomogeneities affect the geodesics (For a more detailed discussion on the metrics, geodesics and other related concepts see Plebanski & Krasinski 2006 and Schutz 2003).

The first detection of cosmic shear signal was at the start of this millennium (see Bacon et al. 2000; Kaiser et al. 2000; Van Waerbeke et al. 2000; Wittman et al. 2000). Since the cosmic shear signal is very weak, it is very challenging to measure. The state-of-the-art cosmic shear survey is CFHTLenS. I will use the data from this survey in Chapter 4, and estimate cosmological parameters. The CFHTLenS team applied a few cosmic shear methods

to constrain cosmological parameters. Fig. 1.3 shows the results from Heymans et al. (2013) which uses redshift binning to constrain σ_8 and Ω_m. They also combine their results with CMB (WMAP7, see Komatsu et al. 2011), BAO (BOSS, see Anderson et al. 2012) and constraints on the Hubble parameter (R11, see Riess et al. 2011).

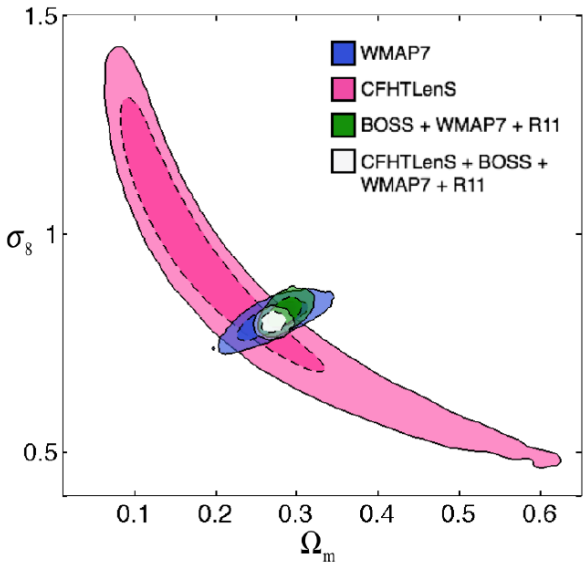

Figure 1.3: Constraints on cosmological parameters adopted from Heymans et al. (2013). σ_8 and Ω_m are the two parameters being measured. Combining CFHTLenS data with BAO, CMB and Hubble parameter priors, results in tighter constraints on these parameters.

Chapter 2

Weak Gravitational Lensing and Statistics

Weak gravitational lensing by large scale structure also known as "cosmic shear" refers to the study of the effects of gravitational potential inhomogeneities on light bundles which travel through the Universe. This effect can be estimated by measuring the distortions in the images of distant objects, for example galaxies. The cosmic shear signal is very weak and therefore, it can only be measured statistically, over an ensemble sample of galaxies.

This Chapter will start with a section for the statistical tools utilized throughout this thesis. Next we will go through the basics of gravitational lensing in Sect. 2.2 and the theoretical background of cosmic shear analysis in Sect. 2.3, which was briefly introduced in Chapter 1 as a promising probe of cosmology. In addition we will see some of the challenges of such analysis and focus on E-/B-mode decomposition in Sect. 2.4 and Sect. 2.5. COSEBIs, a method for separating the gravitational lensing modes from the others, which is going to be used in Chapters 3 and 4, will be introduced in Sect. 2.6. I will end this Chapter by showing some of the theoretical results using COSEBIs which will become relevant in Chapter 4.

2.1 Statistical Tools

Given a data set and a model a basic question comes to mind, what is the likelihood of the model given this data set? The answer lies in the likelihood function, $L(\Phi|\mathbf{x})$, of a set of model parameters, Φ, given a set of observables, $\mathbf{x}$, which is defined as the probability, $P(\mathbf{x}|\Phi)$, of the occurrence of the data assuming the model is correct. Note that it is sensible to write the conditional probability above in the form presented and not vice versa, since there is one true model, and an infinite number of data sets that can belong to it which statistically converge to the best fit values derived from the model.

Finding the likelihood distribution of the underlying parameters of a model based on a set

of observables that are functions of the said parameters, gets more challenging as the number of model parameters increases. For a model with a large parameter space one needs to use clever methods to capture the functional form of the likelihood in a relatively short computational time. The Monte Carlo Metropolis Hastings (MCMC) algorithm is one of the methods that provides a fast solution to this problem, which I will use in the data analysis in Chapter 4.

Provided with the likelihood function, three quantities can be extracted: the best-fit parameters, the errors on their fitted value, and a statistical measure for the goodness-of-fit in the form of a figure-of-merit quantity. The figure-of-merit analysis serves the purposes for the theory chapters, because it demonstrates the capability of the methods in extracting information from cosmological data, in the form of cosmological parameters.

In the first subsection I introduce the likelihood and Chi-square, χ^2, functions. The second subsection is dedicated to Fisher matrix, its definition, meaning and treatment. And finally in the last subsection I introduce the MCMC algorithm and explain the way it is used for the analysis.

2.1.1 Likelihood and Chi-square functions

By definition, the likelihood of a set of parameter values given a set of data, $L(\Phi|\mathbf{x})$, is equivalent to a constant factor of the probability of that data set given the parameter values, $P(\mathbf{x}|\Phi)$. Let us for simplicity assume that this constant factor is one, so that $L(\Phi|\mathbf{x}) = P(\mathbf{x}|\Phi)$. The likelihood is a function of the parameters, Φ, and it represents how likely it is for these set of parameters to be true, for a fixed data set. An important function related to the likelihood is the log-likelihood function,

$$\mathfrak{L} = -\ln(L) \,, \tag{2.1}$$

since it is closely related to the χ^2 function in the case of a Gaussian distributed data set. The χ^2 function for a general case is

$$\chi^2 = \sum_{i,j}^{N} [x_i - \mu_i(\Phi)](C^{-1})_{ij}[x_j - \mu_j(\Phi)] \,, \tag{2.2}$$

where μ_i are the expected values of the observables, x_i, and C^{-1} is the inverse covariance matrix, containing the errors on the estimates on x_i. The χ^2 fitting is a generalized (weighted) form of least-square fitting. To see why notice that if the covariance is a diagonal matrix with equal diagonal components, σ^2, the simple least-square formula is attained,

$$\chi^2_{\text{simple}} = \frac{1}{\sigma^2} \sum_{i}^{N} [x_i - \mu_i(\Phi)]^2 \,, \tag{2.3}$$

and minimizing this simplified χ^2 results in the fitting.

The first quest for a likelihood analysis is to find the best fit values. This is done by finding the global maximum of the likelihood or equivalently the global minimum of the log-likelihood function. This means that the gradient of the log-likelihood should be zero at the best fit parameter point,

$$\left.\frac{\partial \mathfrak{L}(\Phi|\mathbf{x})}{\partial \phi_i}\right|_{\Phi=\bar{\Phi}} = 0 \,, \tag{2.4}$$

where $\bar{\Phi}$ are the best-fit parameter values, and ϕ_i are the components of Φ. To find $\bar{\Phi}$ we assume a trial point $\Phi^{(0)}$ and Taylor expand the derivative of $\mathfrak{L}$ at $\bar{\Phi}$ with respect to their values at the trial point,

$$\mathfrak{L}_{,\phi_i}(\bar{\Phi}|\mathbf{x}) = \mathfrak{L}_{,\phi_i}(\Phi^{(0)}|\mathbf{x}) + \sum_j \mathfrak{L}_{,\phi_i\phi_j}(\Phi^{(0)}|\mathbf{x})(\bar{\phi}_j - \phi_j^{(0)}) + \dots \,, \tag{2.5}$$

where partial derivatives are replaced with commas followed by subscripts. Equating $\mathfrak{L}_{,\phi_i}(\bar{\Phi}|\mathbf{x})$ to zero for all of the parameters, and disregarding the derivatives higher than second order leads to

$$0 \simeq \mathrm{D}\mathfrak{L}(\Phi^{(0)}|\mathbf{x}) + \sum_j \mathrm{D}\mathfrak{L}_{,\phi_j}(\Phi^{(0)}|\mathbf{x})(\bar{\phi}_j - \phi_j^{(0)}) = \mathrm{D}\mathfrak{L}(\Phi^{(0)}|\mathbf{x}) + \mathrm{D}^2\mathfrak{L}(\Phi^{(0)}|\mathbf{x})(\bar{\Phi} - \Phi^{(0)}) \,, \tag{2.6}$$

where $\mathrm{D}\mathfrak{L}(\Phi^{(0)}|\mathbf{x})$ and $\mathrm{D}^2\mathfrak{L}(\Phi^{(0)}|\mathbf{x})$ are the gradient and the Hessian matrix of $\mathfrak{L}$ at $\Phi^{(0)}$. Rearranging the above relation an estimation for the true parameter values are found,

$$\bar{\Phi} \simeq \Phi^{(0)} - [\mathrm{D}^2\mathfrak{L}(\Phi^{(0)}|\mathbf{x})]^{-1}\mathrm{D}\mathfrak{L}(\Phi^{(0)}|\mathbf{x}) \,. \tag{2.7}$$

In this approximation I have assumed that the log-likelihood function is (nearly) quadratic and hence the likelihood function is (nearly) a multivariate Gaussian distribution with respect to the parameters. This is a reasonable approximation because of the central limit theorem which states that, for a large data set with randomly distributed deviations around its mean, the overall shape of the probability distribution converges to a normal (Gaussian) distribution. There are also other compelling reasons to assume Gaussianity which will become clear later on.

A multivariate Gaussian distribution for $\mathbf{x}$ has the following form:

$$L(\boldsymbol{\mu}(\Phi)|\mathbf{x}) = \frac{\mathrm{e}^{-\chi^2/2}}{(2\pi)^{N/2}\sqrt{\det C}} \,; \tag{2.8}$$

with the χ^2 function defined in Eq. (2.2), the likelihood in general will not be Gaussian in parameter space. Nevertheless, it is conventional and benign to assume Gaussianity near the maximum of the likelihood in most realistic cases.

Furthermore, the log-likelihood function related to the likelihood of Eq. (2.8) is

$$2\mathfrak{L}(\boldsymbol{\mu}(\Phi)|\mathbf{x}) = \ln(\det C) + N\ln(2\pi) + (\mathbf{x} - \boldsymbol{\mu}(\Phi))^{\mathrm{T}} C^{-1} (\mathbf{x} - \boldsymbol{\mu}(\Phi)) \ . \tag{2.9}$$

2.1.2 Fisher Matrix

A Fisher analysis may be the easiest and the most convenient method to analyze a set of data or observables. Consider a model with a number of parameters, ϕ_j, and a data set with a number of data points, x_i, then Fisher analysis determines the amount of information contained in the data set, x_i, about the parameters, ϕ_j, assuming the probability distribution of the data given the model is Gaussian, in other words the likelihood function is Gaussian. In general this assumption is satisfied near the maximum of the likelihood function. Another way of looking at the Fisher matrix is to acknowledge its close relation to the covariance matrix of the model parameters; in this sense Fisher matrix analysis is a compact way of propagating errors from the observable, to the parameter space.

Formally, the Fisher matrix is defined as the ensemble average of second derivatives of log-likelihood function at the maximum likelihood point,

$$F_{ij} \equiv \left\langle \frac{\partial^2 \mathfrak{L}}{\partial\phi_i \, \partial\phi_j} \right\rangle \tag{2.10}$$

To make the connection between this definition and the meaning of Fisher matrix explained above let us consider the Taylor expansion of $\mathfrak{L}_{,\phi_i}$, Eq. (2.5). The second term in this expansion contains the curvature matrix of the log-likelihood function, $\mathrm{D}^2\mathfrak{L}$, which is a measure of how fast the likelihood function drops near its maximum and the Fisher matrix is the expectation value of the curvature.

To calculate the Fisher matrix one does not need to find the likelihood function, instead by considering the derivatives of Eq. (2.9) and their expectation value the following analytical formula is found,

$$F_{ij} = \langle \mathfrak{L}_{,ij} \rangle = \frac{1}{2}\mathrm{Tr}[C^{-1} \, C_{,i} \, C^{-1} \, C_{,j} + C^{-1} \, M_{ij}] \ , \tag{2.11}$$

where $M_{ij} = \boldsymbol{\mu}_{,i} \, \boldsymbol{\mu}_{,j}^{\mathrm{T}} + \boldsymbol{\mu}_{,j} \, \boldsymbol{\mu}_{,i}^{\mathrm{T}}$ (see Tegmark et al. 1997 for example).

Fisher matrices have several advantages which come in the form of several theorems. The first and foremost is the Cramer-Rao theorem, which states that for any unbiased estimator the minimum one sigma error of parameter ϕ_i is $(F_{ii})^{-1/2}$ if all other parameters are held fixed. If all parameters are allowed to vary, the minimum one sigma error is $(F^{-1})_{ii}^{1/2}$. (see for example Kenney & Keeping 1951 and Kendall & Stuart 1960 for details).

The Fisher matrix defines the confidence contours, with constant probability in the

parameter space. Due to Gaussianity these contours are ellipsoids characterized by

$$\Delta\Phi^{\mathrm{T}}\, F\, \Delta\Phi = \beta\,, \tag{2.12}$$

where $\Delta\Phi = \Phi - \bar{\Phi}$ is the offset from the best-fit value, and β is a constant that depends on the confidence level desired and the number of free parameters (see Bassett et al. 2009 and Press et al. 2002). The volume of the ellipsoids are given by

$$V_n = V_{\mathrm{S}^n} \times \sqrt{\frac{\beta}{\det F}}\,, \tag{2.13}$$

where

$$V_{\mathrm{S}^n} = \frac{\sqrt{\pi^n}}{\Gamma(n/2+1)}\,, \tag{2.14}$$

is the volume of an n-dimensional sphere and Γ is the Gamma function.

In a cosmological analysis one usually has a number of free parameters, but wants to find the constraints on a smaller number of them. This task can be done in two different ways. One way is to assume that from the N parameters, m are fixed to their fiducial value (or their maximum likelihood value) and the remaining $N - m = n$ parameters are left free. The fixing of parameters is simply done by subtracting the rows and columns corresponding to the fixed parameters in the Fisher matrix, ergo reducing the size of the Fisher matrix from $N\times N$ to $n\times n$. The other way is to marginalize over the unwanted parameters. Assuming the Fisher matrix is composed of $\mathscr{N}_{n\times n}$, $\mathscr{M}_{m\times m}$ and $\mathscr{O}_{m\times n}$, the sub-matrices corresponding to the wanted, unwanted and mixed parameters, respectively, the Fisher matrix can be constructed as,

$$F = \begin{pmatrix} \mathscr{N} & \mathscr{O} \\ \mathscr{O}^{\mathrm{T}} & \mathscr{M} \end{pmatrix}. \tag{2.15}$$

The first step to marginalize the Fisher matrix is done by inverting it, to find the covariance of the parameters,

$$F^{-1} = C_\Phi = \begin{pmatrix} (\mathscr{N} - \mathscr{O}\mathscr{M}^{-1}\mathscr{O}^{\mathrm{T}})^{-1} & -(\mathscr{N} - \mathscr{O}\mathscr{M}^{-1}\mathscr{O}^{\mathrm{T}})^{-1}\mathscr{O}\mathscr{M}^{-1} \\ -(\mathscr{M} - \mathscr{O}^{\mathrm{T}}\mathscr{N}^{-1}\mathscr{O})^{-1}\mathscr{O}^{\mathrm{T}}\mathscr{N}^{-1} & (\mathscr{M} - \mathscr{O}^{\mathrm{T}}\mathscr{N}^{-1}\mathscr{O})^{-1} \end{pmatrix}. \tag{2.16}$$

The second step is to then reduce the above matrix to a smaller covariance matrix with the desired n parameters,

$$\tilde{C}_\Phi = (\mathscr{N} - \mathscr{O}\mathscr{M}^{-1}\mathscr{O}^{\mathrm{T}})^{-1}\,. \tag{2.17}$$

Finally by inverting the reduced covariance matrix, $\tilde{C}_\Phi$, the marginalized Fisher matrix is found,

$$\tilde{F} = \mathscr{N} - \mathscr{O}\mathscr{M}^{-1}\mathscr{O}^{\mathrm{T}}\,. \tag{2.18}$$

In the theory of matrices, $\tilde{F}$ is called the Schur complement of $\mathscr{M}$; it is frequently used in solving systems of equations (Zhang 2005).

If the two sets of parameters are independent, i.e. $\mathscr{O} = 0$, then $\tilde{F} = \mathscr{N}$, and the unwanted parameters have no effect on the desired ones. This case is equivalent to fixing the unwanted parameters.

Furthermore, one can impose independent priors to the analysis, by multiplying their probability to the likelihood function, which in the case of Gaussian priors translates to adding their Fisher matrix to the original Fisher matrix. For a set of independent priors on parameters, the prior Fisher matrix is diagonal, while using the results of another survey as a prior involves full Fisher matrices with mixed terms.

In this thesis, I will use this figure-of-merit,

$$f = \left(\frac{1}{\sqrt{\det \mathbf{F}}} \right)^{1/P} , \tag{2.19}$$

which gives a measure of the mean error on parameters, where P is the number of free parameters considered. Recall that $\sqrt{\det \mathbf{F}}$ is inversely proportional to the volume of the confidence regions, for a Gaussian distribution.

The Fisher matrix can also be used to propagate the bias in the measured observables to the estimated parameters. Taylor et al. (2007) show that for a Gaussian distributed likelihood the linear bias for a parameter, ϕ_μ, given the bias in the observables, $\mathbf{x}$, is

$$B_\mu = \left(F^{-1}\right)_{\mu\nu} \frac{\mathrm{d}\mu_i}{\mathrm{d}\phi_\nu} \left(C^{-1}\right)_{ij} \left(\mu_j - x_j\right) , \tag{2.20}$$

where μ_j is the expected value of x_j and Einstein summation rules apply (also see Knox et al. 1998; Kim et al. 2004). In Chapter 5 I will use this formula to estimated the bias introduced in cosmological parameters from a Pseudo Cl analysis of simulated galaxy shear fields.

2.1.3 MCMC

Although Fisher analysis is useful for testing the information content from a theoretical point of view, it is not used for data analysis. To constrain model parameters from data a full likelihood analysis is mandatory. For a likelihood analysis to take place, typically a statistic[1] is estimated from the data, and is compared to its value from theory assuming different sets of model parameters.

The main question that a likelihood analysis is trying to answer is: "given the likelihood of a set of statistics, what is the corresponding likelihood of the model parameters?" Note that we need to either assume a likelihood for our statistics or use other methods such as studying simulations to find the likelihood of the statistics. In the majority of cases the likelihood of

[1] A quantity that can be measured from the data alone, e.g. mean, variance.

the statistics is not known. In this case a Gaussian likelihood is the most common assumption, shown in Eq. (2.8).

The simplest way to estimate the parameter likelihood is to measure it on a grid in the parameter space. This method is reliable and straightforward, however, for a multidimensional parameter space it quickly becomes computationally expensive[2]. As a result, other methods have been developed to tackle this issue. The most widely used method is the Monte Carlo Markov Chain (MCMC) method (see Heavens 2009, for example). Here I explain the steps taken in a typical MCMC algorithm to estimate the likelihood of the model parameters.

A Markov chain is a random process where each state (point) in the chain depends *only* on the previous one. For an MCMC chain we need to first pick a start point, $\phi_{\rm start}$. Then the next point, $\phi_{\rm prop}$, is chosen from a proposal distribution, $P(\phi_{\rm prop}|\phi_{\rm start})$ which depends on the start point. The proposal distribution can have any functional form[3] and generally can be asymmetric. If the proposal distribution is wildly different from the likelihood distribution mapping of the likelihood becomes too slow and the chain would need a high number of points to converge to the true likelihood distribution. When the general form of the likelihood distribution is not known beforehand, a practical choice for the proposal distribution is a multivariate Gaussian distribution with the Fisher matrix as the inverse covariance matrix. The functional form of such a distribution is given in Eq. (2.8).

After picking a point randomly from the proposal distribution its likelihood is compared to the likelihood of the start point. If it is higher then the proposed point is added to the chain, however, a lower likelihood does not automatically exclude the new point. Instead it is accepted by a probability of

$$P(\text{acceptance}) = \frac{\mathrm{L}(\phi_{\rm prop})\mathrm{P}(\phi_{\rm prop}|\phi_{\rm start})}{\mathrm{L}(\phi_{\rm start})\mathrm{P}(\phi_{\rm start}|\phi_{\rm prop})}\,, \tag{2.21}$$

where $L(\phi)$ is the likelihood of the data belonging to the model with parameters ϕ. For simplicity most proposal distributions take a symmetric form, i.e. $P(\phi_{\rm prop}|\phi_{\rm start}) = P(\phi_{\rm start}|\phi_{\rm prop})$. In this case the acceptance probability reduces to the ratio of the likelihood of the proposed point to the start point. If the proposed point is not accepted the start point is repeated in the chain. This process is repeated with each point, that is added to the chain, as the new start point.

If the start point of a chain is outside the high likelihood volume a number of steps are needed for the chain to burn into this volume. These burn in points need to be excluded from the analysis, since they depend on the choice of the start point rather than the likelihood itself. Typically, a visual inspection is sufficient for finding the burn in chain.

There are several tests of convergence for MCMC chains. The aim of these tests is to find out if the chain mimics the likelihood well enough. These methods usually compare different

[2]For p parameters to make a grid with n points on each side n^p points need to be calculated.

[3]For example a Gaussian centered on the start point.

statistical properties of several chains with different starting point to measure the convergence. One of the more standard tests is the Gelmann & Rubin (1992).

2.2 Basics of Gravitational Lensing

The most basic equation relevant for gravitational lensing (GL) is the relation between the deflection angle, $\hat{\alpha}$, of a light ray passing a spherically symmetric body of mass M with an impact parameter, ξ, where ξ is much larger than the Schwarzschild radius, $R_s \equiv 2GM/c^2$,

$$\hat{\alpha} = \frac{4GM}{c^2\xi} \, . \tag{2.22}$$

The deflection angle obtained here is a factor of two larger than that expected from Newtonian physics. For a mass distribution the aforementioned equation can be generalized, since in the weak-field limit the sum of the deflection angles from mass elements is equal to the total deflection angle from the total mass (for a more detailed discussion about the basics of lensing see Schneider 2009b; Schneider et al. 1999, 2006).

2.2.1 Lens Equation

The lens equation is the fundamental equation governing the gravitational lensing effects. Considering a typical lensing system as in Fig. 2.1, the lens equation reads:

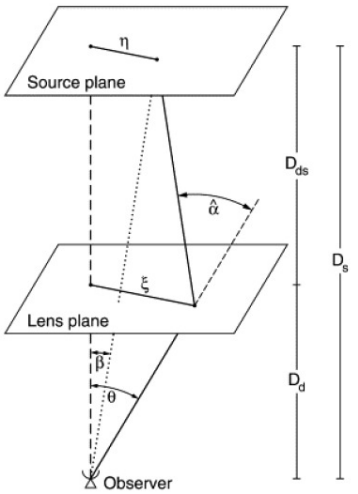

Figure 2.1: A typical lensing configuration. The distances between the observer, the lens and the source planes are calculated in terms of angular diameter distance. The lens changes the position angle of the source from $\boldsymbol{\beta}$ to $\boldsymbol{\theta}$, as seen by the observer (adopted from Bartelmann & Schneider 2001).

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \frac{D_{ds}}{D_s}\hat{\boldsymbol{\alpha}}(D_d\boldsymbol{\theta}) \equiv \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta}) \, , \tag{2.23}$$

where $\boldsymbol{\beta}$ is the source position on the sky (not an observable), $\boldsymbol{\theta}$ is the image position (observable), and $\boldsymbol{\alpha}$ is the scaled deflection angle. The lens equation is valid for the weak field regime, and thin lenses, which corresponds to mass distributions which have a small line-of-sight span compared to the distances between the observer, the lens and the source. For a thin lens we can assume that the light ray is a straight line in the vicinity of the lens and the deflection happens on the lens plane. In this case $\hat{\alpha}$ can be written in terms of the sum of the deflection angles from small mass elements at position $\boldsymbol{r'} = (\xi'_1, \xi'_2, r'_3)$,

$$\mathrm{d}m = \mathrm{d}^2\xi' \mathrm{d}r'_3 \, \rho(\boldsymbol{r'}) \, , \tag{2.24}$$

where $\boldsymbol{\xi} = D_\mathrm{d}\boldsymbol{\theta}$ is the position vector on the lens plane, while r_3 shows the position in the line-of-sight direction, perpendicular to the lens plane. If the light ray has an impact vector of $\boldsymbol{\xi} = (\xi_1, \xi_2)$ then Eq. (2.22) reads,

$$\hat{\boldsymbol{\alpha}} = \frac{4G}{c^2} \int \mathrm{d}^2\xi' \int \mathrm{d}r'_3 \, \rho(\boldsymbol{r'}) \frac{\boldsymbol{\xi} - \boldsymbol{\xi'}}{|\boldsymbol{\xi} - \boldsymbol{\xi'}|^2} \, , \tag{2.25}$$

where $\hat{\boldsymbol{\alpha}}$ is a two dimensional vector. Since $\rho(\boldsymbol{r'})$ is the only quantity in the above equation that depends on r'_3, the line-of-sight integral can be taken independent of the source position. As a result we can define a surface mass density for the lens as follows,

$$\Sigma(\boldsymbol{\xi}) \equiv \int \mathrm{d}r'_3 \, \rho(\xi'_1, \xi'_2, r'_3) \, . \tag{2.26}$$

We can rewrite Eq. (2.25) for the scaled deflection angle,

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{1}{\pi} \int_{R^2} \mathrm{d}\theta'^2 \, \kappa(\boldsymbol{\theta'}) \frac{\boldsymbol{\theta} - \boldsymbol{\theta'}}{|\boldsymbol{\theta} - \boldsymbol{\theta'}|^2} \, , \tag{2.27}$$

where the dimensionless surface mass density (convergence) was defined as,

$$\kappa(\boldsymbol{\theta}) \equiv \frac{\Sigma(D_\mathrm{d}\boldsymbol{\theta})}{\Sigma_\mathrm{cr}} \qquad \text{with} \qquad \Sigma_\mathrm{cr} = \frac{c^2}{4\pi G} \frac{D_\mathrm{s}}{D_\mathrm{d} D_\mathrm{ds}} \, , \tag{2.28}$$

where Σ_cr is the critical surface mass density which depends on the distances between the observer, the lens and the source (see Fig. 2.1).

Using the identity

$$\nabla_\theta \ln(\boldsymbol{\theta} - \boldsymbol{\theta'}) = \frac{\boldsymbol{\theta} - \boldsymbol{\theta'}}{|\boldsymbol{\theta} - \boldsymbol{\theta'}|^2} \tag{2.29}$$

we can rewrite Eq. (2.27) as the gradient of a deflection potential, $\psi(\boldsymbol{\theta})$,

$$\begin{aligned} \boldsymbol{\alpha}(\boldsymbol{\theta}) &= \frac{1}{\pi} \int_{R^2} \mathrm{d}\theta'^2 \, \kappa(\boldsymbol{\theta'}) \nabla_\theta \ln(\boldsymbol{\theta} - \boldsymbol{\theta'}) \\ &= \nabla\psi(\boldsymbol{\theta}) \, . \end{aligned} \tag{2.30}$$

In addition, via $\nabla^2 \ln|\boldsymbol{\theta}| = 2\pi\delta_D(\boldsymbol{\theta})$, a Poisson equation for the deflection potential and the convergence is obtained,

$$\nabla^2\psi = \nabla.\boldsymbol{\alpha} = 2\kappa \; . \tag{2.31}$$

The Jacobian matrix, $\mathscr{A}$, is responsible for a mapping from lens to source plane,

$$\mathscr{A}(\boldsymbol{\theta}) = \frac{\partial\boldsymbol{\beta}}{\partial\boldsymbol{\theta}} = \delta_{ij} - \frac{\partial^2\psi(\boldsymbol{\theta})}{\partial\theta_i\,\partial\theta_j} = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix}, \tag{2.32}$$

where the shear, $\gamma \equiv \gamma_1 + \mathrm{i}\gamma_2$, was defined, as a two component quantity.

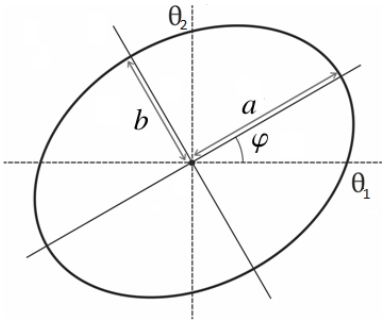

Figure 2.2: A galaxy image is estimated by an ellipse for weak gravitational lensing. The ellipse is characterized by its semi-major and minor axes, a and b, and its orientation, φ with respect to the reference frame.

In weak lensing observations the galaxies are approximated by ellipses, and characterized by their ellipticity (see Schneider 2009b). For the configuration in Fig. 2.2 the two components of the ellipticity are defined as,

$$\begin{pmatrix} \epsilon_1 \\ \epsilon_2 \end{pmatrix} = \frac{1-\beta}{1+\beta}\begin{pmatrix} \cos(2\varphi) \\ \sin(2\varphi) \end{pmatrix}, \tag{2.33}$$

where $\beta = a/b$. Like shear, ellipticity, ϵ, can be written as a complex quantity,

$$\epsilon = \epsilon_1 + i\epsilon_2 \; . \tag{2.34}$$

The observed ellipticity of a galaxy depends on the reduced shear,

$$g = \frac{\gamma}{1-\kappa} \; , \tag{2.35}$$

and its intrinsic ellipticity, ϵ^{int}, via

$$\epsilon^{\mathrm{obs}} = \frac{\epsilon^{\mathrm{int}} + g}{1 + g^*\epsilon^{\mathrm{int}}} \; , \tag{2.36}$$

given $|g| \leq 1$, which is the case for weak lensing (see Schramm & Kayser 1995; Seitz & Schneider 1997). The expectation value of the observed ellipticity in the absence of any

systematical errors is then,

$$\langle \epsilon^{\rm obs} \rangle = \langle \epsilon^{\rm int} \rangle + \langle g \rangle \ , \tag{2.37}$$

where $\langle \epsilon^{\rm int} \rangle$ will give rise to a shot noise term for the covariance of the observables as we will see in the following sections.

The shear defined above is in Cartesian coordinates, however, for practical reasons shear is usually defined with respect to a rotated reference frame. If ϕ is the orientation of the rotated frame, the tangential and the cross components of the shear are defined as follows,

$$\gamma_{\rm t} = -{\rm Re}(\gamma\, {\rm e}^{-2{\rm i}\phi}) \ , \qquad \gamma_{\times} = -{\rm Im}(\gamma\, {\rm e}^{-2{\rm i}\phi}) \ . \tag{2.38}$$

In the following sections we will expand the general formalism of gravitational lensing explained here to cosmological scales. Gravitational lensing has many other applications from microlensing which is used for finding exoplanets, galaxy-galaxy lensing which provides information about the environment in which galaxies reside, to lensing by clusters of galaxies which yield a robust measure of cluster mass distribution, and many more (see Hoekstra & Jain 2008, for cosmological application of weak gravitational lensing). Nevertheless, in this thesis I will focus on cosmic shear analysis explained below.

2.3 Very Weak Gravitational Lensing

In the very weak distortion regime relevant for comic shear, both the convergence and shear are small, and the expectation value of the ellipticity of the images is approximated by shear instead of reduced shear.

At first glance the 3D large scale structures in the Universe seem to make the lens equation, which uses the thin lens model, unusable. However, as light travels through a part of the Universe, we can assume, that part is separate from the others and apply the thin lens model to it. This is analogous to the Born approximation in quantum mechanics. This assumption is based on the following facts: The Universe on large scales is homogeneous and the largest structures are significantly smaller than the horizon size. The potential inhomogeneities are weak and change slowly with spatial position.

Over the course of light's journey there can be many infinitesimally thin lenses and the sum of their convergence contributions make up that of the entire intervening large scale structure. As a result, the effective convergence is written in terms of the comoving angular diameter distance, $f_{\rm K}(\chi)$, the matter density contrast δ, the scale factor, a, the matter density parameter, $\Omega_{\rm m}$, the Hubble constant, H_0, and a weight function, $g(\chi)$, which is the source–redshift distribution weighted with the lens efficiency factor,

$$\kappa(\boldsymbol{\theta}) = \int_0^{\chi_\mathrm{h}} \mathrm{d}\chi\; p_\chi(\chi)\, \kappa(\boldsymbol{\theta},\chi)$$
$$= \frac{3H_0^2\Omega_\mathrm{m}}{2c^2} \int_0^{\chi_\mathrm{h}} \mathrm{d}\chi\; g(\chi)\, f_\mathrm{K}(\chi) \frac{\delta(f_\mathrm{K}(\chi)\,\boldsymbol{\theta},\,\chi)}{a(\chi)}\,, \tag{2.39}$$

with

$$g(\chi) = \int_\chi^{\chi_\mathrm{h}} \mathrm{d}\chi'\; p_\chi(\chi')\, \frac{f_\mathrm{K}(\chi'-\chi)}{f_\mathrm{K}(\chi')}\,, \tag{2.40}$$

where χ_h is the comoving horizon scale, and $p_\chi(\chi)\mathrm{d}\chi = p_\mathrm{z}(z)\mathrm{d}z$ is the source redshift distribution (see Bartelmann & Schneider 2001, for the derivation of the above equations).

2.4 Two-Point Statistics

As in most fields of cosmology, cosmic shear studies make use of statistical methods, since the original shape of a galaxy is not known, and consequently no single ellipticity measurement can provide any cosmic shear information to the observer. Nevertheless, for an isotropic Universe, the ensemble average of the main cosmic shear observable (the ellipticity of the galaxies) averages to zero. As a result, to take advantage of the distortions one must use at least a second order statistical measure.

In this Chapter I will focus on real space tools, which are the standard statistical measurements that one commonly uses for weak lensing analysis. These are generally insensitive to masks, which are always present in the observations. In the final Chapter I will cover Fourier space analysis which is more challenging since the masking effects become important in such an analysis.

The two-point statistical properties of galaxy ellipticities are completely quantified by the two-point correlation functions (2PCFs),

$$\xi_\pm(\theta) = \langle\gamma_\mathrm{t}\gamma_\mathrm{t}\rangle(\theta) \pm \langle\gamma_\times\gamma_\times\rangle(\theta), \tag{2.41}$$
$$\xi_\times(\theta) = \langle\gamma_\mathrm{t}\gamma_\times\rangle(\theta)\,, \tag{2.42}$$

which correlate the tangential and cross components of the shear, $\gamma_{\mathrm{t/x}}$ (Eq. 2.38), of two galaxies separated by an angle θ in the sky. The tangential and cross components of the shear are measured with respect to the line connecting the two galaxies. Furthermore, since cosmologists believe that the Universe is parity invariant, the $\xi_\times$ vanishes, and we are left with $\xi_\pm$.

To make the connection between the 2PCFs and the underlying matter distribution let us first consider the relation between shear and convergence. To find this relation it is more

practical to use the following notation which is equivalent to what we had in Sect. 2.2.1. From Eq. (2.32) we can see the following relation between shear, convergence and the lensing potential,

$$\gamma = \frac{1}{2}\partial\partial\psi \,, \qquad \kappa = \frac{1}{2}\partial^2\psi \,, \tag{2.43}$$

with

$$\partial = \partial_1 + i\partial_2 \,, \qquad \partial^* = \partial_1 - i\partial_2 \qquad \text{and} \qquad \partial^2 = \partial\partial^* = \nabla^2 \,, \tag{2.44}$$

where $\partial_{1,2}$ is the partial derivative with respect to θ_1, θ_2. Eliminating ψ in Eq. (2.43) results in a relation between the shear and convergence,

$$\kappa = \partial^*\partial^*\partial^{-2}\gamma, \tag{2.45}$$

where the inverse Laplacian operator is

$$\partial^{-2} \equiv \int \frac{\mathrm{d}^2\boldsymbol{\vartheta'}}{2\pi} \ln|\boldsymbol{\vartheta} - \boldsymbol{\vartheta'}| \,. \tag{2.46}$$

In Fourier space the relation between κ and γ is more straightforward. Using the relation between the partial derivatives in real space with their Fourier counterparts

$$\mathscr{F}(\partial) = i\hat{\ell} \,, \qquad \mathscr{F}(\partial^*) = i\hat{\ell}^* \,, \tag{2.47}$$

where

$$\hat{\ell} = \ell_x + i\ell_y \,, \qquad \hat{\ell}^* = \ell_x - i\ell_y \,, \tag{2.48}$$

we find

$$\hat{\gamma}(\boldsymbol{\ell}) = \mathrm{e}^{2\mathrm{i}\phi_\ell}\hat{\kappa}(\boldsymbol{\ell}) \,, \tag{2.49}$$

where ϕ_ℓ is the polar angle of $\boldsymbol{\ell}$, $\hat{\kappa}$ and $\hat{\gamma}$ are the Fourier transforms of κ and γ. Consequently assuming a flat sky,

$$\langle\hat{\kappa}(\boldsymbol{\ell})\,\hat{\kappa}^*(\boldsymbol{\ell'})\rangle = \langle\hat{\gamma}(\boldsymbol{\ell})\,\hat{\gamma}^*(\boldsymbol{\ell'})\rangle = (2\pi)^2\delta_\mathrm{D}(\boldsymbol{\ell} - \boldsymbol{\ell'})\,P_\kappa(\ell) \,, \tag{2.50}$$

where $P_\kappa(\ell)$ is the convergence power spectrum and is related to the matter power spectrum, $P_\delta(k)$, via the Limber equation with a certain choice of weight functions (Kaiser 1998),

$$P_\kappa(\ell) = \frac{9H_0^4\Omega_\mathrm{m}^2}{4c^4}\int_0^{\chi_\mathrm{h}} \mathrm{d}\chi\, \frac{g^2(\chi)}{a^2}\, P_\delta\left(\frac{\ell}{f_\mathrm{K}(\chi)}, \chi\right) \,. \tag{2.51}$$

The above equation can be derived from Fourier transforming Eq. (2.39), for the convergence.

Therefore, the 2PCFs are related to P_δ, but also to P_κ, given by

$$\xi_+(\theta) = \int_0^\infty \frac{\mathrm{d}\ell\ \ell}{2\pi}\, \mathrm{J}_0(\ell\theta)\, P_\kappa(\ell) \qquad \xi_-(\theta) = \int_0^\infty \frac{\mathrm{d}\ell\ \ell}{2\pi}\, \mathrm{J}_4(\ell\theta)\, P_\kappa(\ell)\ , \tag{2.52}$$

where J_n is the n-th order Bessel function of the first kind.

In practice the 2PCFs are measured by correlating ellipticities of galaxies,

$$\xi_\pm(\theta) = \langle \epsilon_\mathrm{t}\epsilon_\mathrm{t}\rangle(\theta) \pm \langle \epsilon_\times\epsilon_\times\rangle(\theta), \tag{2.53}$$

which on average can be written in terms of their intrinsic ellipticity and shear, which in the case of mean shear reduces from Eq. (2.36) to

$$\epsilon = \epsilon_\mathrm{int} + \gamma\ . \tag{2.54}$$

By doing so a term containing the shear-shear correlation is produced along with other terms containing ϵ_int. The ellipticity correlation then can be written as,

$$\langle\epsilon\epsilon\rangle = \langle\gamma\gamma\rangle + \langle\epsilon_\mathrm{int}\epsilon_\mathrm{int}\rangle + \langle\epsilon_\mathrm{int}\gamma\rangle + \langle\gamma\epsilon_\mathrm{int}\rangle\ . \tag{2.55}$$

The first term is the quantity that we need, while the other terms contaminate the estimated 2PCFs. The second term in the above equation is the intrinsic-intrinsic correlation (II), which is only important for pairs of galaxies which are physically close in space and their intrinsic shape and orientation is affected by the same gravitational potential. Depending on which of the galaxies is closer to the observer one of the third or fourth terms vanish. Since the intrinsic shape of a background galaxy has no correlation with the shear of a foreground galaxy, only one of these terms is important. A background galaxy shape is sheared by the same gravitational field that a foreground galaxy resides in. This correlation (GI) can have an amplitude comparable to the first term (see Joachimi & Schneider 2010 for a comparison between the different terms). Nulling and boosting techniques explained in Joachimi & Schneider (2010), are a way to correct for these terms. In addition various attempt have been focused on modeling the intrinsic alignments signal (see for example Troxel & Ishak 2014, for a review on the subject and the references therein.) The significance of these intrinsic alignments is noted in King & Schneider (2003) and Heymans et al. (2004).

2.4.1 E-modes and B-modes

Converting a shear field to a convergence field does not necessarily result in a real field (see Eq. 2.45 for example). The reason is that aside from first order lensing effects there are other influential factors. These other factors fall into two categories according to whether their origin is physical or non-physical. The former may arise from higher order lensing effects

(contributions beyond the Born approximation used to derive Eq. 2.39, see also Schneider et al. 1998), source redshift clustering (Schneider et al. 2002c), or intrinsic galaxy alignments; while the latter case involves noise contributions and remaining systematic effects, for example, in galaxy shape measurements. First order weak gravitational lensing can only produce modes which are commonly referred to as E-modes. Whereas the modes which arise from the imaginary part of the estimated κ, are called B-modes. These modes are so named because of the similar mathematical properties of the shear field and the polarization of an electromagnetic field (both of them are polars). The largest contribution to B-modes by far, for the majority of observations, comes from errors in estimating the shapes of galaxies. This is due to incomplete or inaccurate seeing corrections. The seeing which is characterized by the point spread function (PSF), is the effect of the atmosphere and the properties and defects of the telescope on the images. It is very challenging to model the PSF effects for very small galaxy images. Since the physical contributions to the B-modes are very small, measuring a vanishing B-mode, suggests (but not guarantees) a satisfactory PSF correction.

In the presence of B-modes the gravitational field can be written in terms of a complex potential, ψ,

$$\psi(\boldsymbol{\theta}) = \psi_{\rm E}(\boldsymbol{\theta}) + {\rm i}\,\psi_{\rm B}(\boldsymbol{\theta})\ , \tag{2.56}$$

and convergence, κ,

$$\kappa(\boldsymbol{\theta}) = \kappa_{\rm E}(\boldsymbol{\theta}) + {\rm i}\,\kappa_{\rm B}(\boldsymbol{\theta})\ , \tag{2.57}$$

where the subscripts stand for E-modes and B-modes respectively. Likewise, the shear is also written in terms of E/B-modes,

$$\gamma_1 = \frac{1}{2}(\psi_{\rm E,11} - \psi_{\rm E,22}) - \psi_{\rm B,12}\ , \qquad \gamma_2 = \frac{1}{2}(\psi_{\rm B,11} - \psi_{\rm B,22}) - \psi_{\rm E,12}\ , \tag{2.58}$$

where commas followed by numbers indicate partial derivatives with respect to coordinates in the image plane (see Bartelmann & Schneider 2001, for example).

In real space the relation between κ and γ in terms of their derivatives can be found from Eq. (2.43),

$$\partial\kappa = \partial^*\gamma = \partial_1\gamma_1 + \partial_2\gamma_2 + {\rm i}(\partial_2\gamma_1 - \partial_1\gamma_2)\ . \tag{2.59}$$

We can write this relation in vectorial form,

$$\mathbf{u}(\boldsymbol{\theta}) \equiv \nabla\kappa = \begin{pmatrix} \gamma_{1,1} + \gamma_{2,2} \\ \gamma_{2,1} - \gamma_{1,2} \end{pmatrix}\ . \tag{2.60}$$

Inserting derivatives of Eq. (2.58), into the above equation, and making use of the Poisson

equation, $\nabla^2\psi_{\rm E,B} = 2\kappa_{\rm E,B}$, gives

$$\mathbf{u}(\boldsymbol{\theta}) = \begin{pmatrix} \kappa_{\rm E,1} - \kappa_{\rm B,2} \\ \kappa_{\rm E,2} + \kappa_{\rm B,1} \end{pmatrix} . \tag{2.61}$$

The divergence and curl of $\mathbf{u}$ depend only on E-/B-modes respectively,

$$\nabla^2\kappa_{\rm E} = \nabla.\mathbf{u} \equiv \partial_1 u_1 + \partial_2 u_2 \ , \qquad \nabla^2\kappa_{\rm B} = \nabla \times \mathbf{u} \equiv -\partial_2 u_1 + \partial_1 u_2 \ , \tag{2.62}$$

which is why these modes are called curl free and divergence free modes (see Fig. 2.3).

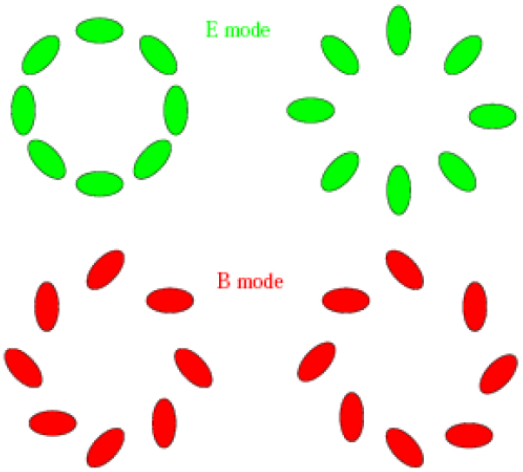

Figure 2.3: A spherical over-density results in tangential distribution of shear around it, while an under-density gives rise to a radial distribution as pictured in the above row. The B-modes do not correspond to the Born approximation, they form the divergence free part of the field, as shown in the lower row (adopted from Van Waerbeke & Mellier 2003).

The real part of κ corresponding to E-modes may arise from both first order lensing and the other effects. In contrast, the imaginary part, B-modes, only come from other effects. If E-modes are the only present contribution to the convergence, then the two components of shear would be interrelated (see Eq. 2.49), hence the 2PCFs would also be related via

$$\int_0^\infty {\rm d}\theta\, \theta\, \xi_+(\theta)\, {\rm J}_0(\theta\ell) = \int_0^\infty {\rm d}\theta\, \theta\, \xi_+(\theta)\, {\rm J}_4(\theta\ell) \ . \tag{2.63}$$

Although, in the presence of B-modes Eq. (2.50), should be rewritten as

$$\begin{aligned} \langle \hat{\kappa}_{\rm E}(\boldsymbol{\ell})\, \hat{\kappa}^*_{\rm E}(\boldsymbol{\ell}')\rangle &= (2\pi)^2 \delta_{\rm D}(\boldsymbol{\ell} - \boldsymbol{\ell}')\, P_{\rm E}(\ell) \ , \\ \langle \hat{\kappa}_{\rm B}(\boldsymbol{\ell})\, \hat{\kappa}^*_{\rm B}(\boldsymbol{\ell}')\rangle &= (2\pi)^2 \delta_{\rm D}(\boldsymbol{\ell} - \boldsymbol{\ell}')\, P_{\rm B}(\ell) \ , \\ \langle \hat{\kappa}_{\rm E}(\boldsymbol{\ell})\, \hat{\kappa}^*_{\rm B}(\boldsymbol{\ell}')\rangle &= (2\pi)^2 \delta_{\rm D}(\boldsymbol{\ell} - \boldsymbol{\ell}')\, P_{\rm EB}(\ell) \ , \end{aligned} \tag{2.64}$$

where the mixed power spectrum, $P_{\rm EB}$, vanishes due to parity symmetry. The 2PCFs should

also be altered to account for the E- and B-modes (see Schneider et al. 2002a),

$$\xi_+(\vartheta) = \int_0^\infty \frac{\mathrm{d}\ell\,\ell}{2\pi} \mathrm{J}_0(\ell\vartheta)[P_\mathrm{E}(\ell) + P_\mathrm{B}(\ell)] \,, \tag{2.65}$$

$$\xi_-(\vartheta) = \int_0^\infty \frac{\mathrm{d}\ell\,\ell}{2\pi} \mathrm{J}_4(\ell\vartheta)[P_\mathrm{E}(\ell) - P_\mathrm{B}(\ell)] \,. \tag{2.66}$$

From the above relations we see that the correlation functions depend on both E- and B-mode power spectra. However, to constrain cosmological models we wish to isolate the E-mode contributions.

2.5 General E/B-Mode Decomposition

In this section I will derive a general second-order decomposition of the form:

$$E = \frac{1}{2}\int_0^\infty \mathrm{d}\vartheta\,\vartheta\,[T_+(\vartheta)\xi_+(\vartheta) + T_-(\vartheta)\xi_-(\vartheta)] \,, \tag{2.67}$$

$$B = \frac{1}{2}\int_0^\infty \mathrm{d}\vartheta\,\vartheta\,[T_+(\vartheta)\xi_+(\vartheta) - T_-(\theta)\xi_-(\theta)] \,,$$

where $\xi_\pm(\vartheta)$ are the two-point correlation functions (2PCFs) of the shear field and $T_\pm(\vartheta)$ are filter functions, that have to be chosen such as to make the statistics E/B, pure E/B-modes. In Schneider & Kilbinger (2007), conditions for such filters were obtained. The procedure is as follows: Substitute for the two-point correlation functions, $\xi_\pm$, from Eqs. (2.65) and (2.66) in the above equations, then change the order of integration and rearrange, results in the following relations between E/B-mode statistics and their power spectra:

$$E = \frac{1}{2}\int_0^\infty \frac{\mathrm{d}\ell\,\ell}{2\pi}\{P_\mathrm{E}(\ell)[W_+(\ell) + W_-(\ell)] + P_\mathrm{B}(\ell)[W_+(\ell) - W_-(\ell)]\} \,, \tag{2.68}$$

$$B = \frac{1}{2}\int_0^\infty \frac{\mathrm{d}\ell\,\ell}{2\pi}\{P_\mathrm{E}(\ell)[W_+(\ell) - W_-(\ell)] + P_\mathrm{B}(\ell)[W_+(\ell) + W_-(\ell)]\} \,, \tag{2.69}$$

where $W_\pm(\ell)$ are the Hankel transforms of $T_\pm(\vartheta)$;

$$W_+(\ell) = \int_0^\infty \mathrm{d}\vartheta\,\vartheta\,T_+(\vartheta)\mathrm{J}_0(\ell\vartheta) \,, \tag{2.70}$$

$$W_-(\ell) = \int_0^\infty \mathrm{d}\vartheta\,\vartheta\,T_-(\vartheta)\mathrm{J}_4(\ell\vartheta) \,. \tag{2.71}$$

If the statistics E or B are to be pure E/B-modes, each of them should respectively depend only on $P_\mathrm{E}(\ell)$ or $P_\mathrm{B}(\ell)$, meaning $W_+(\ell) = W_-(\ell)$; this in turn yields a relation between the filter functions, $T_\pm$. Furthermore, in practice the correlation functions cannot be estimated over all angular ranges, which introduces a minimum and maximum allowed ϑ, basically confining the filter functions to this angular range. Taking the inverse Hankel transform of $W_+(\ell)$, and

making use of the fact that $W_+(\ell) = W_-(\ell)$, gives the following relation between the filters $T_\pm(\vartheta)$:

$$\begin{aligned} T_+(\vartheta) &= \int_0^\infty \mathrm{d}\ell\, \ell\, \mathrm{J}_0(\ell\vartheta) \int_0^\infty \mathrm{d}\theta\, \theta\, T_-(\theta)\mathrm{J}_4(\ell\theta) \\ &= \int_0^\infty \mathrm{d}\theta\, \theta\, T_-(\theta) G(\vartheta,\theta)\,, \end{aligned} \tag{2.72}$$

where $G(\vartheta,\theta)$ is:

$$\begin{aligned} G(\vartheta,\theta) &= \int_0^\infty \mathrm{d}\ell\, \ell\, \mathrm{J}_0(\ell\vartheta)\mathrm{J}_4(\ell\theta) \\ &= \left(\frac{4}{\theta^2} - \frac{12\,\vartheta^2}{\theta^4}\right) H(\theta-\vartheta) + \frac{1}{\theta}\delta_\mathrm{D}(\theta-\vartheta)\,, \end{aligned} \tag{2.73}$$

in which $H(x)$ is the Heaviside step function, and δ_D is the Dirac delta function (Schneider et al. 2002a). Inserting $G(\vartheta,\theta)$ into Eq. (2.72) and repeating the same procedure for $T_-(\vartheta)$ results in two sets of equations relating the filters to each other;

$$T_+(\vartheta) = T_-(\vartheta) + \int_\vartheta^\infty \mathrm{d}\theta\, \theta\, T_-(\theta)\left(\frac{4}{\theta^2} - \frac{12\,\vartheta^2}{\theta^4}\right)\,, \tag{2.74}$$

$$T_-(\vartheta) = T_+(\vartheta) + \int_0^\vartheta \mathrm{d}\theta\, \theta\, T_+(\theta)\left(\frac{4}{\vartheta^2} - \frac{12\,\theta^2}{\vartheta^4}\right)\,. \tag{2.75}$$

Imposing the finite angular range condition on the above equations, respectively brings forth this set of equations constraining the choice of $T_\pm$:

$$\int_{\vartheta_\mathrm{min}}^{\vartheta_\mathrm{max}} \frac{\mathrm{d}\vartheta}{\vartheta}\, T_-(\vartheta) = 0 = \int_{\vartheta_\mathrm{min}}^{\vartheta_\mathrm{max}} \frac{\mathrm{d}\vartheta}{\vartheta^3}\, T_-(\vartheta)\,, \tag{2.76}$$

$$\int_{\vartheta_\mathrm{min}}^{\vartheta_\mathrm{max}} \mathrm{d}\vartheta\, \vartheta\, T_+(\vartheta) = 0 = \int_{\vartheta_\mathrm{min}}^{\vartheta_\mathrm{max}} \mathrm{d}\vartheta\, \vartheta^3\, T_+(\vartheta)\,. \tag{2.77}$$

It is important to notice that the above conditions are only valid if $\vartheta_\mathrm{min} > 0$ in Eq. (2.76), and $\vartheta_\mathrm{max} < \infty$ in Eq. (2.77). Although these constraints were calculated using mathematical reasoning, there is a physical reason behind them as well. Using such filters makes the E/B statistics independent of constant or linear shear fields, both of which cannot be distinguished uniquely as E/B-modes. The E/B-modes can be defined as the curl-free and the divergence-free modes respectively (see Sect. 2.4.1). A uniform shear field, which produces a constant ξ_+ and zero ξ_-, resulting in $E = B$. Moreover, a linear shear field is both curl and divergence free since the u vector defined in Eq. (2.60) is zero, hence such a field cannot be distinguished as E/B-mode explicitly. In this case the correlation functions are, $\xi_+ = a + b\vartheta^2$, $\xi_- = 0$, where a and b are constants[4].

[4]The shear in this case is $\gamma = c_1\boldsymbol{\vartheta} + c_2\boldsymbol{\vartheta}^* + c_3$, where c_i are constants.

There is an infinite number of filter functions satisfying Eqs. (2.77) & (2.76). Such filters can be expanded in sets of orthogonal functions, and the E/B-statistics in Eqs (2.67) can be rewritten with subscripts indicating the filter used for them (see Schneider et al. 2010):

$$E_n = \frac{1}{2}\int_0^\infty \mathrm{d}\vartheta\, \vartheta\, [T_{+n}(\vartheta)\xi_+(\vartheta) + T_{-n}(\vartheta)\xi_-(\vartheta)]\ , \tag{2.78}$$
$$B_n = \frac{1}{2}\int_0^\infty \mathrm{d}\vartheta\, \vartheta\, [T_{+n}(\vartheta)\xi_+(\vartheta) - T_{-n}(\theta)\xi_-(\theta)]\ ,$$

in terms of real space quantities, and Eqs (2.68) which relates the E/B to the Fourier space quantities can be written as

$$E_n = \int_0^\infty \frac{\mathrm{d}\ell\, \ell}{2\pi} P_\mathrm{E}(\ell) W_n(\ell)\ , \tag{2.79}$$
$$B_n = \int_0^\infty \frac{\mathrm{d}\ell\, \ell}{2\pi} P_\mathrm{B}(\ell) W_n(\ell)\ , \tag{2.80}$$

where W_n are:

$$\begin{aligned} W_n(\ell) &= \int_{\vartheta_\mathrm{min}}^{\vartheta_\mathrm{max}} \mathrm{d}\vartheta\, \vartheta\, T_{+n}(\vartheta)\mathrm{J}_0(\ell\vartheta) \\ &= \int_{\vartheta_\mathrm{min}}^{\vartheta_\mathrm{max}} \mathrm{d}\vartheta\, \vartheta\, T_{-n}(\vartheta)\mathrm{J}_4(\ell\vartheta)\ , \end{aligned} \tag{2.81}$$

the $\pm$ are omitted from the original definition of W since they are equal as we saw above for Eq. (2.70).

Furthermore, the covariance of these statistics can be calculated for one single redshift bin;

$$\begin{aligned} C^\mathrm{X}_{mn} &\equiv \langle X_m X_n\rangle - \langle X_m\rangle\langle X_n\rangle \\ &= \int_0^\infty \frac{\mathrm{d}\ell\, \ell}{2\pi} W_m(\ell) \int_0^\infty \frac{\mathrm{d}\ell'\, \ell'}{2\pi} W_n(\ell')\langle \Delta P_\mathrm{X}(\ell)\Delta P_\mathrm{X}(\ell')\rangle \\ &= \frac{1}{\pi A}\int_0^\infty \mathrm{d}\ell\, \ell\, W_m(\ell)W_n(\ell)\left(P_\mathrm{X}(\ell) + \frac{\sigma_\epsilon^2}{2\bar{n}}\right)^2\ , \end{aligned} \tag{2.82}$$

where $\Delta P_\mathrm{X}(\ell) = P_\mathrm{X}(\ell) - \langle P_\mathrm{X}(\ell)\rangle$ and X is substituted in place of either E or B. The covariance depends on the parameters of the observations given by the survey area, A, the galaxy intrinsic ellipticity dispersion, σ_ϵ, and the mean number density of galaxies, $\bar{n}$. This can be analytically calculated assuming a Gaussian shear field as in Joachimi et al. (2008). Sato et al. (2010) have shown that the Gaussian covariance estimation in Joachimi et al. (2008) overestimates the true Gaussian covariance for surveys with small area ($A \lesssim 1000$ deg^2), and they have developed a fitting formula to correct for this discrepancy; in spite of their findings I will stick with the estimation of Joachimi et al. (2008), since the fitting formula in the latter paper depends on source redshift and is developed for a single source galaxy redshift, making it non-applicable

for this work. One can write the covariance (Eq. 2.82) in terms of $T_\pm$ and the covariance of the two-point correlation functions as follows:

$$C^{\rm E}_{mn} = \frac{1}{4}\int_{\vartheta_{\rm max}}^{\vartheta_{\rm min}} {\rm d}\vartheta\, \vartheta \int_{\vartheta_{\rm max}}^{\vartheta_{\rm min}} {\rm d}\vartheta'\, \vartheta' \sum_{s,s'=\{+,-\}} T_{sm}(\vartheta)\, T_{s'n}(\vartheta') C_{ss'}(\vartheta,\vartheta')\,, \tag{2.83}$$

$$C^{\rm B}_{mn} = \frac{1}{4}\int_{\vartheta_{\rm max}}^{\vartheta_{\rm min}} {\rm d}\vartheta\, \vartheta \int_{\vartheta_{\rm max}}^{\vartheta_{\rm min}} {\rm d}\vartheta'\, \vartheta' \sum_{s,s'=\{+,-\}} s\; s'\; T_{sm}(\vartheta)\, T_{s'n}(\vartheta') C_{ss'}(\vartheta,\vartheta')\,, \tag{2.84}$$

where $C_{\pm\pm}(\vartheta,\vartheta')$ is the covariance of the 2PCFs.

2.5.1 Methods

One of the widely used methods for E/B decomposition of the 2PCFs is called the aperture mass dispersion (Schneider et al. 1998, 2002a), although it has a number of limitations. The aperture mass was defined in Kaiser (1995) and Schneider (1996) as,

$$M_{\rm ap} = \int {\rm d}^2\vartheta \kappa(\boldsymbol{\vartheta}) U(\vartheta)\,, \tag{2.85}$$

where $U(\vartheta)$ is a compensated filter function,

$$\int {\rm d}\vartheta\vartheta U(\vartheta) = 0 \tag{2.86}$$

and is (usually) zero outside a certain maximum ϑ. The aperture mass dispersion E/B-modes follow the format given in Eq. (2.67). The $T_\pm$ filter functions for the aperture mass dispersion is non-zero up to zero angular distances. Hence to calculate the E/B-mode aperture mass dispersions one needs to know $\xi_\pm$ to arbitrarily low separations, a task not feasible since two galaxies with very small angular separation are not distinguishable in observations, and their shapes cannot be measured. This also means that the aperture mass dispersion filters do not follow the conditions given in Eq. (2.76). Typically, to fill in the gap for the small separations theoretical predictions are used, which can bias the results by more than 10% (Kilbinger et al. 2006). Another problem with this method is that during the decomposition some of the valuable information contained in $\xi_\pm$ is lost, hence lowering the signal-to-noise of the measurement.

To correct for the first flaw of the aperture mass dispersion method, Schneider & Kilbinger (2007) introduced the ring statistics, which correlate the shear in two rings avoiding small separations. The $T_\pm$ filter functions for these statistics are called $Z_\pm$ and are shown in Fig. 2.4. These filters follow the conditions in Eqts. (2.76) and (2.77). However, the signal-to-noise of ring statistics is even lower than that of aperture mass dispersion. In the next section we will go through another E-/B-mode decomposition method, COSEBIs, which conserves all of the E-mode information in the 2PCFs and thus is a superior method to the ones mentioned above.

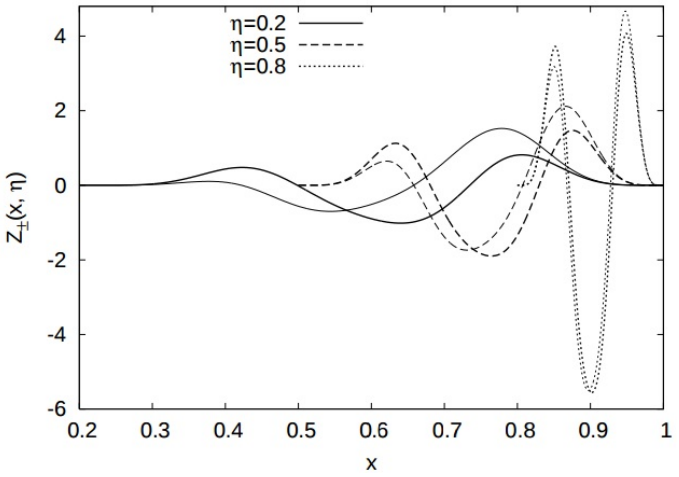

Figure 2.4: The filter functions of Ring statitics. Z_+ (thick curves) and Z_- (thin curves) are shown for three different ratio values, $\eta = \theta_{\rm max}/\theta_{\rm min}$. $\theta_{\rm max}$ and $\theta_{\rm min}$ are the maximum and minimum angular separation of two galaxies in two rings. These filters are zero outside of their range and oscillate inside (adopted from Schneider & Kilbinger 2007).

2.6 COSEBIs

In the previous section we saw that in order to build pure E-/B-mode statistical measures defined on a finite angular interval from 2PCFs (Eq. 2.67), we need filter functions that satisfy conditions in Eqs. (2.77) & (2.76). There is in fact an infinite number of filter functions satisfying these conditions. Schneider et al. (2010) have constructed two complete orthogonal sets of these functions which form the basis of "Complete Orthogonal Sets of E-/B-mode Integral", "COSEBIs".

The two sets of basis functions are the Lin- and Log-COSEBIs, which are written in terms of polynomials in θ and $\ln(\theta)$ in real space, respectively. In addition to that work, in Fu & Kilbinger (2010), other filters where constructed which maximized the signal-to-noise ratio for a specific angular range, or maximize the information content of E statistics via Fisher analysis. In this thesis I will use the Log-COSEBIs, since compared to the Lin-COSEBIs a smaller number is sufficient to essentially capture all the information.

2.6.1 Filters

In order to measure the COSEBIs modes from data we need to first know the functional form of $T_{\pm n}$, which will be then applied to $\xi_\pm$ to filter out the E-/B-modes (see Eq. 2.78). In this thesis I will only use Log-COSEBIs for the analysis and hence here we will go through the derivation of their filter functions.

To solve for $T_{+n}(\vartheta)$ Schneider et al. (2010) defined $t_{+n}(z) = T_{+n}(\vartheta)$, with

$$z = \ln\left(\frac{\vartheta}{\vartheta_{\rm min}}\right) . \tag{2.87}$$

We require $T_{+n}(\vartheta)$ to be zero outside of $[\vartheta_{\rm min}, \vartheta_{\rm max}]$ so that E_n and B_n are measured for this angular range. Then the argument of t_{+n}, would be confined to the interval between 0 and

$z_{\max} = \ln(\vartheta_{\max}/\vartheta_{\min})$. To find the Log-COSEBIs, Schneider et al. (2010) wrote $t_{+n}(z)$ as polynomials in z,

$$t_{+n}(z) \equiv \sum_{j=0}^{n+1} c_{nj} z^j = N_n \sum_{j=0}^{n+1} \bar{c}_{nj} z^j \,, \tag{2.88}$$

where c_{nj}, $\bar{c}_{nj}$ and N_n are constant coefficients and $c_{nj} = N_n \bar{c}_{nj}$ with $N_n \equiv c_{n(n+1)} \neq 0$ so that $\bar{c}_{n(n+1)} = 1$, were defined to simplify the calculations. Applying the conditions in Eq. (2.77) for t_{+n} results in,

$$\int_0^{z_{\max}} \mathrm{d}z\, e^{2z}\, t_{+n}(z) = 0 = \int_0^{z_{\max}} \mathrm{d}z\, e^{4z}\, t_{+n}(z) \,. \tag{2.89}$$

Since we want to find a set of functions T_{+n} which form a set of basis functions for all the filter functions satisfying the conditions necessary to separate E-/B-modes on a finite angular range (Eq. 2.77), it is convenient to enforce an orthonormality condition on them. Hence, Schneider et al. (2010) applied this extra condition on the $T_{+n}(\vartheta)$ functions which for $t_{+n}(z)$ yield

$$\frac{\vartheta_{\min}}{\Delta\vartheta} \int_0^{z_{\max}} \mathrm{d}z\, e^{z}\, t_{+n}(z) t_{+m}(z) = \delta_{nm} \,, \tag{2.90}$$

where δ_{nm} is the Kronecker delta and $\Delta\vartheta = \vartheta_{\max} - \vartheta_{\min}$. To write these constraints in a more compact and elegant way Schneider et al. (2010) introduced a set of coefficients of the form

$$J(k, j) \equiv \int_0^{z_{\max}} \mathrm{d}z\, e^{kz}\, z^j = \frac{\gamma(j+1, -k z_{\max})}{(-k)^{j+1}} \,, \tag{2.91}$$

where $\gamma(a, x)$ is the lower incomplete gamma function.

Inserting the definition in Eq. (2.88) into the constraints in Eq. (2.89), gives rise to the following sets of equations:

$$\begin{aligned} \sum_{j=0}^{n+1} \bar{c}_{nj} J(2, j) &= -J(2, n+1) \,, \\ \sum_{j=0}^{n+1} \bar{c}_{nj} J(4, j) &= -J(4, n+1) \,. \end{aligned} \tag{2.92}$$

The first two coefficients, namely $\bar{c}_{10}$ and $\bar{c}_{11}$, can be obtained solely from the above equations. These coefficients together with the normalization, N_1, produce the first weight function $t_{+1}(z)$. The rest of the coefficients, $\bar{c}_{nj}$, can also be constructed if all the $\bar{c}_{mj}$ with $m < n$ have been determined. Inserting $t_{+n}(z)$ from Eq. (2.88) in the orthonormality conditions in Eq. (2.90)

results in

$$\left.\begin{aligned} &\sum_{j=0}^{n+1}\sum_{i=0}^{m+1} J(1,i+j)\bar{c}_{nj}\bar{c}_{mi} = 0\ , \\ &\text{or} \\ &\sum_{j=0}^{n}\left(\sum_{i=0}^{m+1} J(1,i+j)\bar{c}_{mi}\right)\bar{c}_{nj} = -\sum_{i=0}^{m+1} J(1,i+j)\bar{c}_{mi}\ , \end{aligned}\right\} \text{for } 1 \le m \le n \tag{2.93}$$

and

$$N_n^2 \sum_{i,j=0}^{n+1} J(1,i+j)\bar{c}_{nj}\bar{c}_{ni} = \frac{\vartheta_{\rm min}}{\Delta\vartheta} = e^{z_{\rm max}} - 1\ . \tag{2.94}$$

Eqs. (2.92) and the $(n-1)$ orthonormality relations in Eq. (2.93) provide the $(n+1)$ linear equations to solve for the $(n+1)$ $\bar{c}_{nj}$ coefficients. Furthermore, the orthonormality relation for t_{+n} with itself (Eq. 2.94), fixes the normalization coefficients, N_n, up to an arbitrary sign which for definiteness are assumed to be positive.

Numerically the coefficients are not well-behaved, in the sense that one needs to have very high accuracy (high number of meaningful digits) for them to satisfy Eq. (2.93). On the other hand, as was done in Schneider et al. (2010), the representation

$$t_{+n}(z) = N_n \prod_{i=1}^{n+1}(z - r_{ni})\ , \tag{2.95}$$

is easier to handle and fewer number of digits is needed for the r_{ni} roots. Schneider et al. (2010) have shared a Mathematica code for calculating r_{ni} and N_n for a given angular range and number of COSEBIs modes. For a larger n a larger working precision is required in that code.

The Lin-COSEBIs filters have analytical forms in real (Schneider et al. 2010) and also in Fourier space (Asgari et al. 2012). However, the roots and the amplitude of the Log-COSEBIs filter functions, $T_{\pm n}$, need to be solved analytically. In summary to find these filter functions take the following steps:

- Choose an angular range for which the analysis is going to be performed.
- Solve the system of $n+1$ equations (2.92) and (2.93) to find the $n+1$ constants.
- Alternatively for a more stable result, solve the system of equations in (2.95).

Fig. 2.5 shows the Lin- and Log-COSEBIs filter functions for an angular range of $[1', 400']$ for three COSEBIs modes, $n = 1$, $n = 2$ and $n = 10$. The number of roots for each filter function in its base of support is equal to $n+1$. The Lin-COSEBIs filter functions, are polynomials

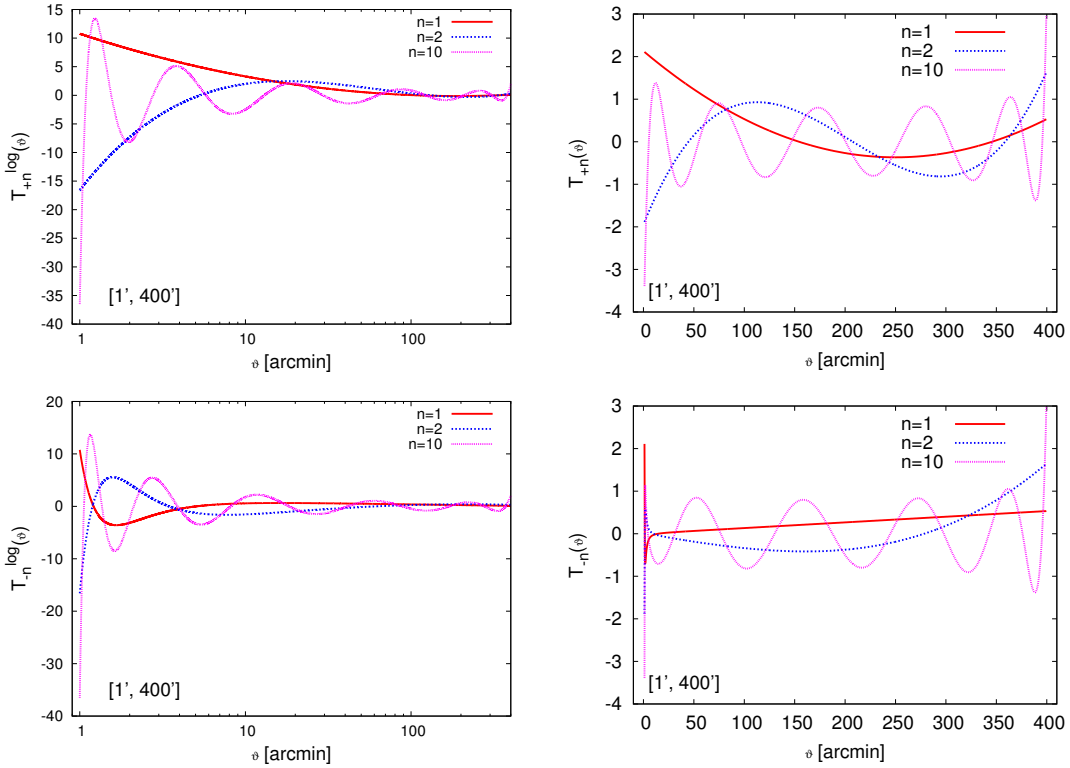

Figure 2.5: The COSEBIs filter function in real space. The $T_{\pm}^{\mathrm{Log}}$ shown in the left panels have fairly evenly spaced roots in logarithmic scales, while the Lin-COSEBIs filters shown in the right panels have their root distributed relatively evenly in their base of support. The angular range considered for these plots is $[1', 400']$.

in θ. Ergo, they oscillate fairly homogeneously in their angular range as seen in the right panels. Whereas the Log-COSEBIs filters, which are polynomials in $\ln(\vartheta)$, have their roots distributed relatively regularly in logarithmic scales. This means that the linear $T_{\pm}(\theta)$ show the same sensitivity across all angular scales in $[1', 400']$. However, since the distribution of the roots of Log-COSEBIs is more concentrated towards smaller θ they show a preference towards sampling smaller scales, which turns out to be beneficial for sampling $\xi_{\pm}$ more efficiently, since they show more features on smaller angular scales. Consequently, a much smaller number of Log-COSEBIs modes are required to reach the full information level. In this section we will go through the derivation of the Log-COSEBIs $T_{\pm}$.

COSEBIs in Fourier space

While COSEBIs data analysis is performed in real space, the quickest way to treat COSEBIs in theory is to work in Fourier space and to use Eq. (2.82) for the covariance. This allows one to use the natural grid point that the COSEBIs offer, rather than using the arbitrary angular

grid used for 2PCFs (e.g. compare Eq. 2.82 with Eq. 2.83). Hence here I will explain the $W_n(\ell)$ functions which are the Hankel transform (Eq. 2.81) of their real space counterparts. Since both J_0 and T_{+n} are oscillating functions and especially $J_0(\ell\vartheta)$ oscillations depend on the ℓ-mode considered, evaluating this integral becomes very challenging. A piece-wise integration, from one extremum to the next, is used in this thesis to evaluate $W_n(\ell)$, and the results are stored for further analysis. In Asgari et al. (2012) I showed a (semi-)analytic formula for the linear $W_n(\ell)$. However, currently there is no analytic formula for their logarithmic counterpart. In Fig. 2.6 the behavior of the Log-COSEBIs weight functions in Fourier space can be seen. There are two oscillatory modes present for each W_n which depends on the boundaries of its angular range, $\theta_{\rm min}$ and $\theta_{\rm max}$. We can see these modes of oscillations in the zoomed in panels. Here three angular ranges are considered, $[1', 400']$, $[20', 400']$ and $[1', 20']$. By comparing the plots we see that the lower frequency oscillations, shown in the the top panels, depend on $\theta_{\rm min}$ while the higher frequency ones, shown in the bottom panels, depend on $\theta_{\rm max}$. The number of oscillations for W_n with the same angular range but different n is the same. However, the amplitude of the oscillations is higher for a larger n for large ℓ values. As a result the higher COSEBIs modes have more information about the convergence power spectrum for larger ℓ, which corresponds to smaller scales in real space. The position of the first peak of W_n depends inversely on $\theta_{\rm max}$. The first peak is at a larger ℓ for a larger n.

2.6.2 Covariance

The covariance matrix of a statistical measure is required for any analysis. In Chapters 3 I will use the COSEBIs covariance to define a new optimised weak lensing statistic called CCOSEBIs and in Chapter 4 to perform cosmological analysis of the CFHTLenS survey. The covariance is used also in defining figure-of-merit quantities to predict the performance of future surveys explained in Sect. 2.1.

Eq. (2.82) shows the relation between COSEBIs' covariance and convergence power spectra for one redshift bin assuming a Gaussian shear field. In Fig. 2.7 we see the overall behaviour of such a covariance for the E-mode Log-COSEBIs, for a CFHT-like survey with parameters given in Table. 4.3 (see Asgari et al. 2012). The shot noise term- arising from the intrinsic ellipticity dispersion of the galaxies- in Eq. (2.82) is set to zero for this plot, for demonstration purposes. Note that in general this term is dominant and cannot be neglected.

The Log-COSEBIs have a band covariance matrix, where modes with close n show a higher correlation.

By dividing the galaxies into tomographic redshift bins we can better understand the evolution of large scale structures. Similar to Eq. (2.68) the E-mode COSEBIs can be written in terms of the cross power spectrum between redshift bins i and j,

$$E_n^{ij} = \int_0^\infty \frac{{\rm d}\ell\,\ell}{2\pi} P_{\rm E}^{ij}(\ell) W_n(\ell)\,, \tag{2.96}$$

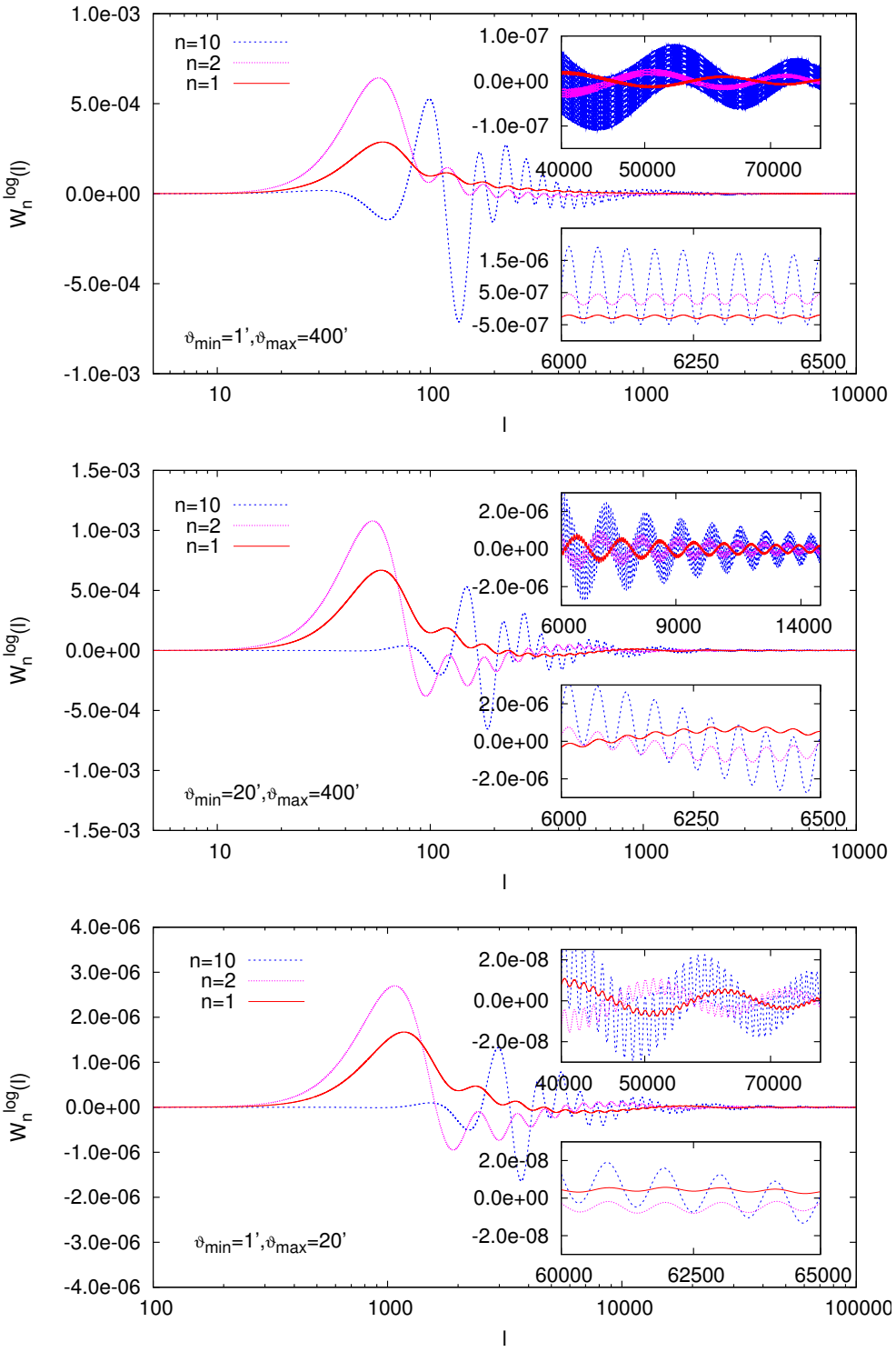

Figure 2.6: The weight functions $W_n(\ell)$ shown for three angular ranges, $[1', 400']$, $[20', 400']$ and $[1', 20']$. $W_n(\ell)$ functions are the Hankel transformations of $T_\pm(\vartheta)$ as in Eq. (2.81). The position of the first peak mostly depends on $\vartheta_{\rm max}$ and is rather insensitive to $\vartheta_{\rm min}$. There are two modes of oscillations shown in the zoomed in panels, which depend on $\vartheta_{\rm min}$ and $\vartheta_{\rm max}$.

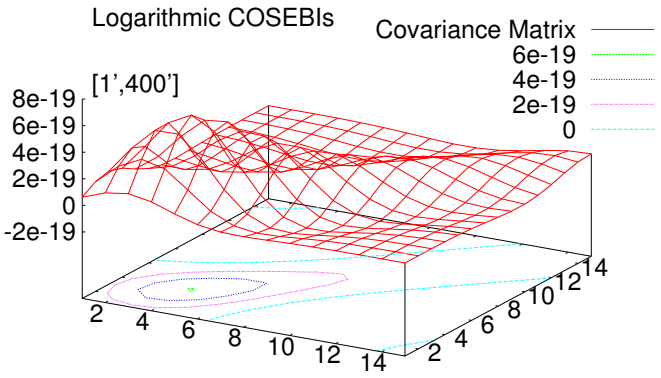

Figure 2.7: A 3D representation of non-tomographic covariance of 15 E-mode COSEBIs for an angular range of $[1', 400']$, for a CFHT-like survey. The x and y axes correspond to the elements of the covariance matrix, and the value of the vertical axis shows the value of the covariance of the corresponding element. In the making of these plots the source ellipticity dispersion, σ_ϵ is assumed to be zero. A contour representation of the covariance is shown at its base.

where the power spectrum is defined as

$$P_{\rm E}^{ij}(\ell) = \frac{9\,H_0^4\Omega_{\rm m}^2}{4c^4}\int_0^{\chi_{\rm h}} {\rm d}\chi\,\frac{g^i(\chi)\,g^j(\chi)}{a^2}\,P_\delta\left(\frac{\ell}{f_{\rm K}(\chi)},\chi\right)\,, \tag{2.97}$$

and

$$g^i(\chi) = \int_\chi^{\chi_{\rm h}} {\rm d}\chi'\,p^i_\chi(\chi')\,\frac{f_{\rm K}(\chi'-\chi)}{f_{\rm K}(\chi')}\,. \tag{2.98}$$

Similarly, the covariance of the tomographic COSEBIs, assuming Gaussian shear fields (see Joachimi et al. 2008) is,

$$\begin{aligned} C_{mn}^{X(ij,kl)} &\equiv \langle X_m^{ij}X_n^{kl}\rangle - \langle X_m^{ij}\rangle\langle X_n^{kl}\rangle \\ &= \frac{1}{2\pi A}\int_0^\infty {\rm d}\ell\;\ell\;W_m(\ell)W_n(\ell)\left(\bar{P}_{\rm X}^{ik}(\ell)\bar{P}_{\rm X}^{jl}(\ell)+\bar{P}_{\rm X}^{il}(\ell)\bar{P}_{\rm X}^{jk}(\ell)\right)\,, \end{aligned} \tag{2.99}$$

where

$$\bar{P}_{\rm X}^{ik}(\ell) := P_{\rm X}^{ik}(\ell) + \delta_{ik}\frac{\sigma_\epsilon^2}{2\bar{n}}\,, \tag{2.100}$$

and X stands for either E or B.

In general, a tomographic covariance for r redshift bins consists of $[r(r+1)/2]^2$ building blocks, each of which is a covariance matrix of (E_n^{ij}, E_m^{kl}) where i, j, k, l are fixed and $n, m = 1, 2, ..., n_{\rm max}$. This means in total the covariance matrix has $[r(r+1)\,n_{\rm max}]^2/4$ elements, where $n_{\rm max}$ is the maximum number of COSEBIs modes considered.

A covariance matrix is by definition symmetric and a tomographic covariance, for a Gaussian field, is made up of smaller covariances, therefore, only $x(x+1)/2\times n_{\rm max}(n_{\rm max}+1)/2$ elements, with $x = r(r+1)/2$, have to be calculated, the rest are equal to these. Fig. 2.8 shows the elements of such a covariance matrix which need to be calculated schematically.

The covariance of the E_n^{ij} depends on six indices; in order to apply normal matrix

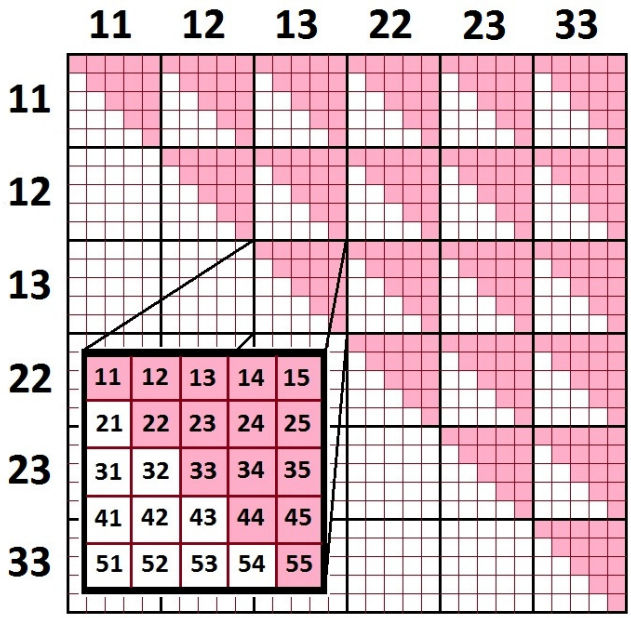

Figure 2.8: A representation of how a tomographic covariance is built. In this diagram 3 redshift-bins (1,2,3) and 5 COSEBIs modes are assumed to be present. The zoomed in panel shows an instance of one of the covariance building blocks; the numbers 1-5 show the COSEBIs mode considered, for example 15 means the covariance of E_1 and E_5 is shown in that particular block. The numbers on the sides of the matrix show which redshift bins are considered, for instance 12 means the covariance of redshift-bins 1 and 2 is relevant. Due to symmetry, only a part of the covariance elements have to be calculated which are colored pink.

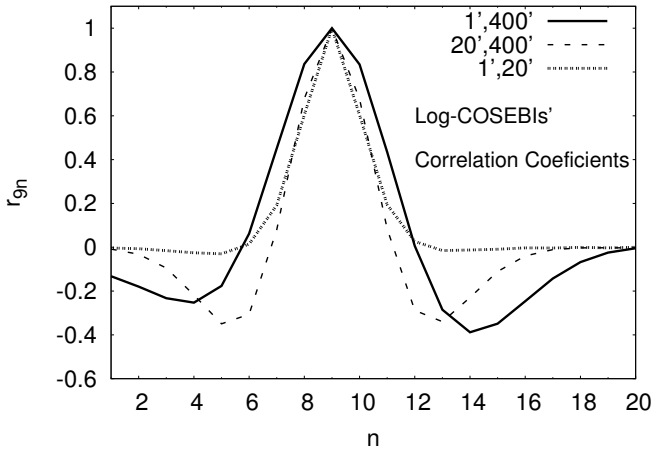

Figure 2.9: Correlation coefficients of non-tomographic COSEBIs for different angular ranges $[\vartheta_{\rm min}, \vartheta_{\rm max}]$ at $m = 9$, for a CFHT-like survey. Here M, the capital subscripts, are equal to the COSEBIs mode, m.

operations, the three indices of E_n^{ij} are combined into one 'superindex' N, given by

$$N = \left[(i-1)\times r - \frac{(i-1)(i-2)}{2} + (j-i)\right] \times n_{\rm max} + n \,, \tag{2.101}$$

where r is the total number of redshift bins and $n_{\rm max}$ is the total number of COSEBIs modes.

To better visualize the behavior of the COSEBIs covariance, in Asgari et al. (2012) I defined correlation coefficients of COSEBIs,

$$r_{MN} = \frac{C^{\rm E}_{MN}}{\sqrt{C^{\rm E}_{MM} C^{\rm E}_{NN}}} \,, \tag{2.102}$$

which is preferable for showing the behavior of the off-diagonal terms. Note that the capital

subscripts N and M can be different from the COSEBIs subscripts, if several source populations are considered.

Fig. 2.9 demonstrates a comparison between correlation coefficients of three different angular ranges, $[1', 400']$, $[20', 400']$, and $[1', 20']$, at a fixed $M = 9$ for a non-tomographic case. The behavior of the correlation coefficients is different for the three angular ranges, however, can see that for all cases r_{9N} is small for an N far from 9.

Using the new labeling, the correlation coefficients of E_7^{11} and E_7^{23} (corresponding to $N = 7$ and $N = 82$, respectively) with the other E_n^{ij} is shown in Fig. 2.10, where 15 COSEBIs modes and 4 redshift bins are considered. Each of the peaks in the figure correspond to the correlation coefficient of E_7^{11} and E_7^{ij}. The highest peak with $r = 1$ occurs for $M = N$, while the rest of the peaks are correlations between different redshift bins. The angular range considered here is $[1', 400']$. Here similar to the single redshift distribution case in Fig. 2.9 we see that the correlations are larger for $n = m$ and decrease as $|n - m|$ increases.

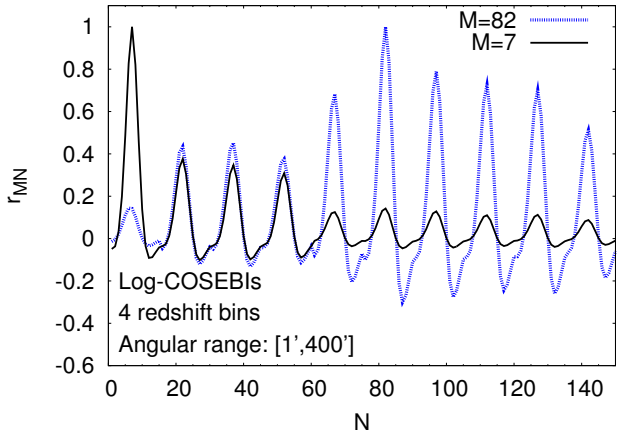

Figure 2.10: The correlation coefficients of COSEBIs for an angular range of $[1', 400']$ and 4 redshift bins. In total 15 COSEBIs modes were considered for each graph.

2.6.3 COSEBIs: A Literature Review

Schneider et al. (2010) developed the formalism for the COSEBIs. In addition, they used a figure-of-merit similar to f in Eq. (2.19) to find out the number of COSEBIs needed to capture the full information level. They assumed a single redshift distribution and two free cosmological parameters, σ_8 and Ω_m in their study. They showed that a finite number of COSEBIs is sufficient for this task. In Asgari et al. (2012) I extended Schneider et al. (2010) to a tomographic analysis with an increased set of cosmological parameters. I found that although the number of COSEBIs modes needed in the presence of tomographic bins increases, a finite number was enough to essentially reach the full information level. Furthermore, I showed the sensitivity of E_n to a few of the cosmological parameters. Fig. 2.11 shows the sensitivity of the first five Log-COSEBIs to three cosmological parameters. Since the amplitude of the COSEBIs filter functions can be set arbitrarily (remember that we enforced orthonormality on

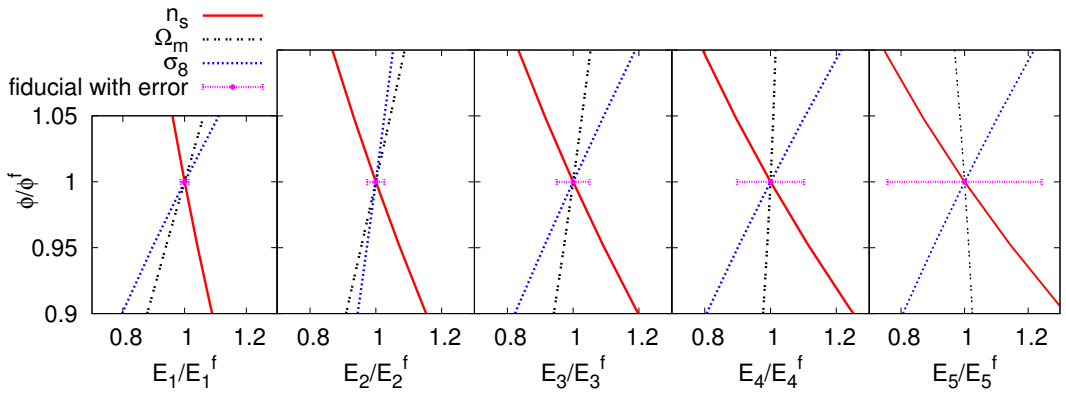

Figure 2.11: Dependence of three cosmological parameters, σ_8, Ω_m, n_s, on the first five E-mode Log-COSEBIs for a single galaxy redshift distribution. Both the parameters and the E_n values are normalized to their fiducial values. The error-bars show the normalized noise, $\sqrt{C_{nn}}/E_n^f$. The parameters of a future large scale survey such as Euclid or LSST are assumed for this figure with an angular range of $[1', 400']$ (Reproduced from Asgari et al. 2012).

them in Sect. 2.6.1), their values normalized with respect to those for the fiducial parameters are shown. For each COSEBIs mode, its normalized variance is shown at the fiducial point as well. The signal-to-noise of the COSEBIs can thus be inferred from these plots. From the plots we see that, the relative variance increases with increasing n, this is due to the stronger oscillations of the weight function T_{+n} (Fig. 2.5) or, equivalently, the W_n (Fig. 2.6). Although it provides a more intuitive way for understanding COSEBIs, this figure by itself cannot be used for constraining parameters, due to the non-diagonal covariance between the E_n's. Therefore, a likelihood or Fisher analysis is required.

Eifler (2011) used COSEBIs in a cosmic shear study of the Coyote Universe emulator, which predicts the matter power spectrum for different cosmological models. They also found that a finite number of COSEBIs is sufficient to capture all the information.

Kilbinger et al. (2013) and Fu et al. (2014) utilized COSEBIs on CFHTLenS data assuming one redshift bin. In Chapter 4 I will give more details about their studies. Huff et al. (2014a) used COSEBIs on Sloan Digital Sky Survey (SDSS) data, for a cosmic shear analysis. This study does not assume tomographic redshift bins either. Their results show constraints for σ_8 and $\Omega_\mathrm{m}h^2$.

In Chapter 4 I will use COSEBIs and compressed COSEBIs, which will be explained in Chapter 3, to constrain cosmological parameters by analysing the CFHTLenS data. Therefore, here I will use the figure-of-merit defined in Eq. (2.19) to show that a finite number of COSEBIs is sufficient to capture the full information for a CFHT-like survey, in Fig. 2.12. For one free parameter $f = \sigma$ the minimum standard deviation of the estimated parameter. σ_8 and Ω_m are the two free parameters in this analysis with a redshift range of 0.58 to 1.3, assuming either a single redshift distribution or four redshift bins. The angular range considered is $[1', 40']$. We

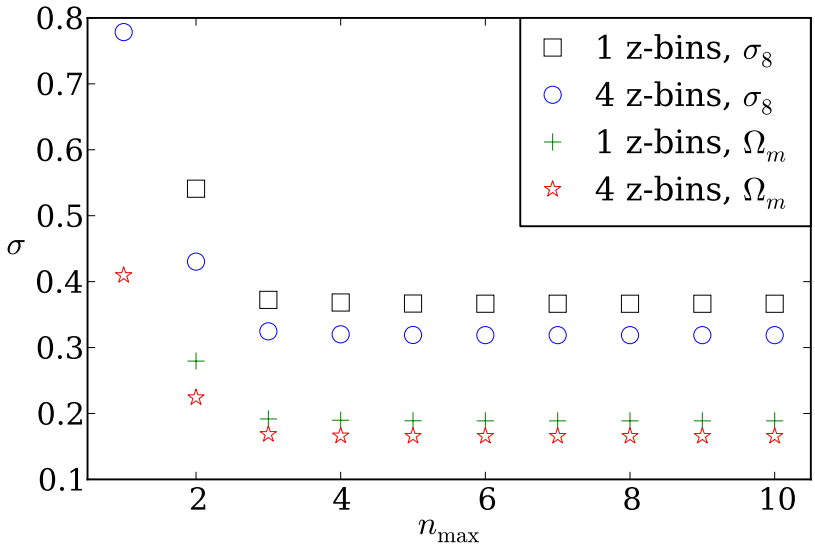

Figure 2.12: The minimum standard deviation of estimated parameters, σ versus $n_{\max}$, the number of COSEBIs modes used in the analysis. Here there are two free parameters, σ_8 and Ω_m, and the others are fixed. For each set of points one of the two parameters is marginalized over. A single redshift distribution and four redshift bins are considered between 0.58 and 1.3, in accordance to the data analysis part of this thesis in Chapter 4 (Reproduced from Asgari et al. 2012).

can compare the information level for the two cases, and between the two parameters. Since f is defined for a single parameter in this figure the other parameter is marginalized over. The value of σ shows the lowest possible error on the parameters, as was discussed in Sect. 2.1.2. From the figure we can see that after 3 COSEBIs modes the value of σ essentially remains unchanged. Note that for 4 redshift bins $n_{\max} = n$ COSEBIs means $n \times 4 \times 5/2$ modes. Also with two free parameters and a single redshift bin, $n_{\max} = 1$ results in a single equation for constraining two variables which results in a singular Fisher matrix, hence, the value of σ is not plotted for this case. We can see that the information gain with tomography is small for this particular case.

Chapter 3

Data Compression for Precision Cosmology

This chapter is dedicated to the work I have done with Peter Schneider which has been published in Asgari & Schneider (2015).

3.1 Importance of Data Compression

Future cosmological surveys are faced with the difficulty of how to extract cosmological parameters from their wealth of observables. Taking the future space-based telescope Euclid[1] as example, statistics to be obtained from the data include second-order shear statistics across several populations of source galaxies, which – using the common usage – will be termed 'redshift bins' throughout this chapter. As shown in Schneider et al. (2010) and Asgari et al. (2012), the COSEBIs (Complete Orthogonal E-/B-mode Integrals) form appropriate combinations of the shear two-point correlation functions $\xi_{\pm}(\theta)$ which cleanly separate E- and B-mode shear (see, e.g., Crittenden et al. 2002; Schneider et al. 2002d). In addition, COSEBIs are highly efficient in terms of data compression, since essentially all cosmological information is contained in a small number of COSEBIs (see, e.g., Kilbinger et al. 2013; Huff et al. 2014b, for applications of COSEBIs to cosmic shear data sets).

The efficiency of data compression decreases, however, if several populations of sources are used. For example, with $\sim$ 10 redshift bins, the total number of COSEBIs which should be used to extract cosmological information is of order 500. Furthermore, higher-order shear information contains additional, valuable information – both regarding cosmological parameters as well as for calibrating the shear data – and should be taken into account. Since third-order shear statistics depend on three variables (say, three sides of a triangle), and combinations of three redshift bins, the number of observables for third-order shear statistics

[1] http://sci.esa.int/euclid/, Laureijs et al. (2011)

which need to be considered is considerably larger than that for second-order shear statistics. Furthermore, shear-peak statistics has been shown to yield powerful constraints and should likewise be considered (see, e.g., Marian et al. 2013, and references therein). Therefore, the number of pure shear observables will be several thousands, although the number of cosmological parameters to be determined is of order a dozen.

In practice, issues are even more complicated, in that astrophysical and other systematics need to be accounted for. For example, effects of intrinsic alignments (see, e.g., Joachimi & Bridle 2010, and references therein) need to be mitigated, by including further observables, i.e., the galaxy-galaxy lensing signal and the galaxy correlation functions. Even if one uses a COSEBI-like data compression for them (e.g., Eifler et al. 2014), the number of redshift combinations will still lead to a strongly enhanced number of observables.

One of the major difficulties in analyzing this data is the accurate determination of the expectation values for these observables as a function of cosmological parameters and, in particular, the estimation of their covariance matrix. If one determines the covariance as a sample variance of different numerical realizations, one needs many more realizations than the dimension of the data vector in order to get a reliable estimate of the covariance matrix and its inverse (see, e.g., Hartlap et al. 2007). Because of this difficulty, data compression is mandatory for any analysis of survey data. Taylor & Joachimi (2014) have found relations between the accuracy of the estimated covariance, the number of simulations and observables for Gaussian distributed data and underlying parameters. Their work also suggest that data compression is essential for reducing the number of simulations required for reaching a certain accuracy in the covariance.

In this chapter, I suggest a form of data compression that is based on the sensitivity of the various observables to the parameters that are to be estimated. This idea has been explored before using Fisher matrices, where observable combinations are chosen to maximize the sensitivity to a given set of parameters (see Tegmark et al. 1997, for example). I have also considered other methods to combine cosmological observables, an example is given in Appendix A. This method lacks the flexibility of the cosmpression method explained in this Chapter and is only presented as an example to show that the compression method considered in this Chapter is supperior. The cosmological parameters currently are, and until the launch of Euclid will be even more, strongly constrained, and thus only a relatively small volume in parameter space needs to be explored.[2] I will therefore assume that the relevant parameter region is small, which allows us to define linear combinations of observables based on a low-order Taylor expansion of the dependence of these observables on parameters, which should contain almost all the cosmological information in the data.

In the following section I introduce our data compression formalism for general observables

[2]Additional parameters, needed to parametrize intrinsic alignment effects, may be less well constrained from independent data sets or theoretical models.

(statistics). I then specialize this method in Sect. 3.3 to study how this strategy works for COSEBIs compression. In Sect. 3.4 I specify the cosmological model which will be used for the results section. In Sect. 3.5 I first illustrate the weight functions for the compressed statistics made of COSEBIs, then using a Fisher formalism I explore the efficiency of the compressed versus regular COSEBIs. Section 3.6 is dedicated to mimicking a band power spectrum using linear combinations of COSEBIs. Finally I conclude in Sect. 3.7.

3.2 Formalism

Let $\hat{X}_n$ be the statistics obtained from the data, $1 \leq n \leq N$, with expectation value $\left\langle \hat{X}_n \right\rangle = X_n(\phi_\mu)$, where the ϕ_μ, $1 \leq \mu \leq P$, denote the parameters of the model, including the cosmological parameters as well as others, for example nuisance parameters. Assuming that the uncertainty in the parameters is 'small', we consider an expansion of the functions $X_n(\phi_\mu)$ around the fiducial value $X_n^{\rm f} = X_n(\phi_\mu^{\rm f})$,

$$X_n(\phi_\mu) = X_n^{\rm f} + \mathbf{D}_{n\mu} p_\mu + \frac{1}{2}\mathbf{Z}_{n\mu\nu} p_\mu p_\nu \ , \tag{3.1}$$

where $p_\mu = \phi_\mu - \phi_\mu^{\rm f}$, and

$$\mathbf{D}_{n\mu} = \left(\frac{\partial X_n}{\partial \phi_\mu}\right)_{|\phi_\kappa^{\rm f}} \ ; \quad \mathbf{Z}_{n\mu\nu} = \left(\frac{\partial^2 X_n}{\partial \phi_\mu \, \partial \phi_\nu}\right)_{|\phi_\kappa^{\rm f}} \tag{3.2}$$

are the first and second derivatives of the expectation values with respect to the model parameters, taken at the fiducial point in parameter space. Here and below, summation over repeated indices is implied, unless noted otherwise.

I assume that the likelihood $\mathscr{L}(\chi^2)$ is a monotonically decreasing function of

$$\chi^2 = \left[\hat{X}_m - X_m(\phi_\mu)\right] \left(\mathbf{C}^{-1}\right)_{mn} \left[\hat{X}_n - X_n(\phi_\mu)\right] \ , \tag{3.3}$$

where $\mathbf{C}$ is the covariance matrix of the observables $\hat{X}_n$. Maximizing the likelihood then requires to find the minimum of χ^2 with respect to the parameters; using Eq. (3.1), I obtain

$$\frac{\partial \chi^2}{\partial p_\kappa} = -2\left[\mathbf{D}_{m\kappa} + \mathbf{Z}_{m\kappa\mu} p_\mu\right] \left(\mathbf{C}^{-1}\right)_{mn} \times \left[\hat{X}_n - X_n^{\rm f} - \mathbf{D}_{n\mu} p_\mu - \frac{1}{2}\mathbf{Z}_{n\mu\nu} p_\mu p_\nu\right] = 0 \ . \tag{3.4}$$

In this equation, I have neglected the dependence of the covariance matrix on the parameters, either because $\mathbf{C}$ is determined from the data itself, or because the dependence of $\mathbf{C}$ on the parameters is assumed to be weak. From Eq. (3.4), we see that the determination of the parameters p_μ does involve the observables $\hat{X}_n$ only in the linear combinations

$$\hat{F}_\kappa := \mathbf{D}_{m\kappa} \left(\mathbf{C}^{-1}\right)_{mn} \hat{X}_n \ ; \quad \hat{S}_{\kappa\nu} := \mathbf{Z}_{m\kappa\nu} \left(\mathbf{C}^{-1}\right)_{mn} \hat{X}_n \ , \tag{3.5}$$

with expectation value

$$F_\kappa := \mathbf{D}_{m\kappa}\left(\mathbf{C}^{-1}\right)_{mn} X_n \; ; \quad S_{\kappa\nu} := \mathbf{Z}_{m\kappa\nu}\left(\mathbf{C}^{-1}\right)_{mn} X_n \; . \tag{3.6}$$

Thus, the expansion of the expectation values of the original observables $\hat{X}_n$ around a fiducial model motivates the definition of linear combinations of observables which contain all the information about the parameters ϕ_κ, provided the second-order expansion is accurate. The set in Eq. (3.5) of $P + P(P+1)/2 = P(P+3)/2$ observables thus is expected to allow for an efficient data compression (note that $\hat{S}_{\mu\nu} = \hat{S}_{\nu\mu}$).

In order to obtain the new observables $\hat{F}_\kappa$ and $\hat{S}_{\kappa\nu}$, one first needs to estimate the covariance $\mathbf{C}$ of the original observables which, due to the high dimensionality in future cosmological surveys, provides a real challenge. However, the covariance $\mathbf{C}$ is needed here for the definition of appropriate combinations of observables, and not for parameter estimates. Hence, an approximation for $\mathbf{C}$ may be expected to be sufficient for this purpose. Disregarding the parameter dependence of $\mathbf{C}$ in the derivation of Eq. (3.4) provides such an approximation which avoids the necessity to obtain a large covariance matrix for more than one cosmological model. If the approximation for $\mathbf{C}$ deviates substantially from the true covariance, I expect that the new observables do not contain the full information about the parameters, since they deviate from the 'optimal' combination of the original $\hat{X}_n$. Hence, the better the initial estimate of $\mathbf{C}$, the more efficient the new observables will be.

Thus, I propose a strategy to first obtain an approximation for the covariance $\mathbf{C}$, based on which the new observables $\hat{F}_\kappa$ and $\hat{S}_{\kappa\nu}$ are defined. The number of these observables is substantially smaller than the original ones, and hence an accurate estimation of their covariance can be obtained from fewer simulations compared to $\mathbf{C}$. On the other hand, the number of new observables is substantially larger than the number of parameters, which is expected to provide a mitigation for the choice of non-optimal combinations from an approximate form of $\mathbf{C}$. It is for this reason that I consider the second-order derivatives of the original observables; the first-order ones coincide with that of the Karhunen–Loève method for the case of known covariance (see, e.g., Tegmark et al. 1997).

We can now combine the new observables $\hat{F}_\kappa$ and $\hat{S}_{\kappa\nu}$ into the $N' = P(P+3)/2$ compressed quantities $\hat{X}^{\mathrm{c}}_i$. According to Eq. (3.5), we can write

$$\hat{\boldsymbol{X}}^{\mathrm{c}} = \mathbf{H}\mathbf{C}^{-1}\hat{\boldsymbol{X}} \equiv \mathbf{B}\hat{\boldsymbol{X}} \; , \tag{3.7}$$

where we are using vectorial notation for the $\hat{X}^{\mathrm{c}}_i$ and $\hat{X}_n$ are used. The $N' \times N$ (rows $\times$ columns) matrix $\mathbf{H}$ is given in terms of first and second partial derivatives of the functions $X_n(\phi_\kappa)$ at the fiducial point in parameter space and $\mathbf{B} = \mathbf{H}\mathbf{C}^{-1}$ is the compression matrix. Accordingly, the covariance matrix of $\hat{\boldsymbol{X}}^{\mathrm{c}}$ is given as

$$\mathbf{C}^{\mathrm{c}} = \mathbf{B}\mathbf{C}\mathbf{B}^{\mathrm{t}} \; , \tag{3.8}$$

where the superscript 't' denotes the transpose of a matrix. The χ^2-function in terms of the new observables is

$$\chi^2 = \left[\hat{\boldsymbol{X}}^{\rm c} - \boldsymbol{X}^{\rm c}\right]^{\rm t} (\mathbf{C}^{\rm c})^{-1} \left[\hat{\boldsymbol{X}}^{\rm c} - \boldsymbol{X}^{\rm c}\right] . \tag{3.9}$$

From what was discussed above, the covariance $\mathbf{C}^{\rm c}$ should be calculated from $\mathbf{C}$ only if an accurate estimate of the latter can be obtained; in general, it will be much more practical to determine $\mathbf{C}^{\rm c}$ directly, for example from simulations.

For completeness, I mention for the special case that the covariance $\mathbf{C}$ *can indeed* be determined accurately, that we can solve Eq. (3.4) for the parameters $\phi_\mu = p_\mu + \phi_\mu^{\rm f}$. Writing it in terms of the new observables, Eq. (3.4) becomes

$$\Delta F_\kappa + \Delta S_{\kappa\nu}\, p_\nu = (\mathbf{D}_{m\kappa} + \mathbf{Z}_{m\kappa\nu} p_\nu) \left(\mathbf{C}^{-1}\right)_{mn} \left(\mathbf{D}_{n\nu} p_\nu + \frac{1}{2}\mathbf{Z}_{n\mu\nu} p_\mu p_\nu\right) , \tag{3.10}$$

with $\Delta F_\kappa = \hat{F}_\kappa - F_\kappa^{\rm f}$, $\Delta S_{\kappa\nu} = \hat{S}_{\kappa\nu} - S_{\kappa\nu}^{\rm f}$. If we then expand $p_\mu = p_\mu^{(1)} + p_\mu^{(2)}$, where $p_\mu^{(1)}$ ($p_\mu^{(2)}$) is first (second) order in the ΔF_κ, $\Delta S_{\kappa\nu}$, we obtain to first order

$$\Delta F_\kappa = \mathbf{D}_{m\kappa} \left(\mathbf{C}^{-1}\right)_{mn} \mathbf{D}_{n\nu} p_\nu^{(1)} \equiv U_{\kappa\nu} p_\nu^{(1)} , \tag{3.11}$$

from which we can easily obtain $p_\nu^{(1)}$ from the inverse of the symmetric matrix $\mathbf{U}$, $p_\nu^{(1)} = \left(\mathbf{U}^{-1}\right)_{\nu\mu} \Delta F_\mu$. The second-order terms lead to the equation

$$\Delta S_{\kappa\nu}\, p_\nu^{(1)} = \mathbf{U}_{\kappa\nu}\, p_\nu^{(2)} + \left(\frac{1}{2}\mathbf{G}_{\kappa\mu\nu} + \mathbf{G}_{\mu\kappa\nu}\right) p_\mu^{(1)} p_\nu^{(1)} , \tag{3.12}$$

where $\mathbf{G}$ is defined as

$$\mathbf{G}_{\kappa\mu\nu} := \mathbf{D}_{m\kappa} \left(\mathbf{C}^{-1}\right)_{mn} \mathbf{Z}_{n\mu\nu} . \tag{3.13}$$

With the foregoing solution for $p_\mu^{(1)}$ and the inverse of $\mathbf{U}$, this can be immediately solved for $p_\mu^{(2)}$.

3.3 Application to COSEBIs

I will now apply the method of the previous section to a specific statistics for cosmic shear measurements, the COSEBIs (see Sect. 2.6). They provide a complete representation of the shear two-point correlation functions (2PCFs) in a given finite interval of angular scales, chosen such that they cleanly separate between E- and B-modes (Crittenden et al. 2002; Schneider et al. 2002d). The E-mode COSEBIs are related to the 2PCFs via

$$E_n^{(ij)} = \frac{1}{2} \int_{\theta_{\rm min}}^{\theta_{\rm max}} {\rm d}\vartheta\, \vartheta \left[T_{+n}(\vartheta)\, \xi_+^{(ij)}(\vartheta) + T_{-n}(\vartheta)\, \xi_-^{(ij)}(\vartheta)\right] , \tag{3.14}$$

where $1 \leq i, j \leq r$ label the redshift bins considered. Their relation to the convergence power spectra is given in Chapter 2, Eq. (2.96).

In Asgari et al. (2012) I showed that COSEBIs already provide an efficient means of data compression, since the full cosmological information contained in the 2PCFs can be recovered with a small number of COSEBIs. However, in the case of several redshift bins for the source galaxies, the number of components grows with the number of tomographic redshift bins, r, by a factor of $r(r+1)/2$. In this section I will use the formalism explained in Sect. 3.2 to obtain a method to compress the number of relevant statistical quantities and compare the results with a full COSEBIs analysis.

In order to apply the method of the past section for obtaining a compressed version of COSEBIs, we need to find their (approximate) covariance matrix for a given cosmology, in addition to their first- and second-order derivatives with respect to the cosmological parameters. The new set of statistics are related to the COSEBIs via the compression matrix, **B**, defined in Eq. (3.7),

$$E^{\rm c}_{\mu} = \mathbf{B}_{\mu\mathscr{N}} E_{\mathscr{N}} = \mathbf{B}_{\mu n i j} E^{(ij)}_{n} \,, \tag{3.15}$$

where the new index

$$\mathscr{N} = \left[(i-1)r - \frac{(i-1)(i-2)}{2} + (j-1)\right] n_{\rm max} + n \tag{3.16}$$

is a combination of the three indices i, j and n, the parameter $n_{\rm max}$ is the maximum order of COSEBIs considered, and r is the total number of redshift bins.

3.4 Cosmological Model, Survey Parameters and Covariance

A cold dark matter (CDM) cosmological models with a dynamical dark energy, characterized by its equation-of-state parameter, w_0, is used throughout this Chapter (for references to wCDM models, see Peebles & Ratra 2003, and references therein). Table. 3.1 contains the two sets of parameter values considered here. The fiducial model is used for obtaining the compressed COSEBIs (CCOSEBIs hereafter), while the assumed 'true' underlying cosmology is slightly different. That means, we calculate the CCOSEBIs according to the equations of Sect. 3.2, using the covariance and parameter derivatives of the COSEBIs, **C**, **D** and **Z**, for the fiducial cosmology; but, these new observables $E^{\rm c}_{\mu}$ are applied using the 'true' cosmological model. The linear matter power spectrum is calculated using the Bond & Efstathiou (1984) transfer function and a primordial power-law power spectrum with spectral index n_s. For non-linear scales, the halo fit formula of Smith et al. (2003) is adopted.

For a cosmic shear analysis, we need the survey parameters and the redshift distribution of

	σ_8	Ω_m	Ω_Λ	w_0	n_s	h	Ω_b
Fiducial	0.8	0.27	0.73	−1.0	0.97	0.70	0.045
True	0.83	0.31	0.68	−1.1	0.96	0.67	0.049

Table 3.1: The fiducial cosmological parameters consistent with the WMAP 7-years results (Komatsu et al. 2011), and the underlying true parameters consistent with Planck (Planck Collaboration et al. 2014b). The normalization of the power spectrum, σ_8, is the standard deviation of perturbations in a sphere of radius $8h^{-1}$Mpc today. Ω_m, Ω_Λ, and Ω_b are the matter, the dark energy and the baryonic matter density parameters, respectively. w_0 is the dark energy equation of state parameter, which is equal to the ratio of dark energy pressure to its density. The spectral index, n_s, is the power of the initial power spectrum. The dimensionless Hubble parameter, h, characterizes the rate of expansion today.

z-distribution parameters					survey parameters		
α	β	z_0	z_{min}	z_{max}	A	σ_ϵ	$\bar{n}$
2.0	1.5	0.71	0.0	2.0	20000	0.3	35

Table 3.2: The parameters of a fiducial large future survey. α, β, and z_0 characterize the total redshift distribution of sources, while z_{min} and z_{max} indicate the minimum and the maximum redshifts of the sources use for the cosmic shear analysis. Here A is the survey area in units of deg^2, σ_ϵ is the galaxy intrinsic ellipticity dispersion, and $\bar{n}$ is the mean number density of sources per square arcminute in the field.

the galaxies. The latter is characterized by

$$p(z) \propto \left(\frac{z}{z_0}\right)^\alpha \exp\left[-\left(\frac{z}{z_0}\right)^\beta\right], \tag{3.17}$$

for $z_{min} \leq z \leq z_{max}$ where the parameters, α, β, z_0, z_{min} and z_{max} depend on the survey (Brainerd et al. 1996). Table. 3.2 summarizes the survey and redshift parameters assumed in our analysis.

I assume Gaussian shear fields to find the covariances needed for obtaining CCOSEBIs and also for the Fisher analysis which is explained in Sect. 2.6.2.

3.5 Results

In this section we will first see the filter functions of the CCOSEBIs for the fiducial cosmology, followed by a figure-of-merit analysis. I will then compare the figure-of-merit values for cases where the covariance is known versus the use of a wrong covariance in constructing the CCOSEBIs.

3.5.1 Weight functions of Compact COSEBIs

Inserting Eqs. (3.14) and (2.96) into Eq. (3.15) results in relations between $\boldsymbol{E}^{\rm c}$ and the COSEBIs filters,

$$E_\mu^{\rm c} = \frac{1}{2}\int_{\theta_{\rm min}}^{\theta_{\rm max}} {\rm d}\vartheta\,\vartheta[\mathbf{B}_{\mu nij}T_{+n}(\vartheta)\,\xi_+^{(ij)}(\vartheta) + \mathbf{B}_{\mu nij}T_{-n}(\vartheta)\,\xi_-^{(ij)}(\vartheta)]\;, \tag{3.18}$$

and

$$E_\mu^{\rm c} = \int_0^\infty \frac{{\rm d}\ell\,\ell}{2\pi}\mathbf{B}_{\mu nij}W_n(\ell)\,P_{\rm E}^{(ij)}(\ell)\;. \tag{3.19}$$

For each redshift bin pair, ij, a set of $N' = P(P+3)/2$ (P is the number of free parameters) filters exist. The new filters in real and Fourier space, respectively, are

$$\begin{aligned} T_{\pm\mu ij}^{\rm c}(\vartheta) &= \mathbf{B}_{\mu nij}T_{\pm n}(\vartheta) \\ W_{\mu ij}^{\rm c}(\ell) &= \mathbf{B}_{\mu nij}W_n(\ell)\;. \end{aligned} \tag{3.20}$$

With the above definitions we can rewrite the compressed statistics, $\boldsymbol{E}^{\rm c}$, in terms of the compressed filter functions,

$$E_\mu^{\rm c} = \frac{1}{2}\int_{\theta_{\rm min}}^{\theta_{\rm max}} {\rm d}\vartheta\,\vartheta[T_{+\mu ij}^{\rm c}\,\xi_+^{(ij)}(\vartheta) + T_{-\mu ij}^{\rm c}\,\xi_-^{(ij)}(\vartheta)]\;, \tag{3.21}$$

and

$$E_\mu^{\rm c} = \int_0^\infty \frac{{\rm d}\ell\,\ell}{2\pi}W_{\mu ij}^{\rm c}(\ell)\,P_{\rm E}^{(ij)}(\ell)\;. \tag{3.22}$$

Multiplying each $E_\mu^{\rm c}$ by a constant has no effect on the information level. I can therefore normalize the filter functions for each compressed statistic separately, so that

$$\sum_{ij}\frac{1}{\Delta\vartheta}\int_{\theta_{\rm min}}^{\theta_{\rm max}} {\rm d}\vartheta\,[T_{+\mu ij}^{\rm c}(\vartheta)]^2 = 1\;. \tag{3.23}$$

Fig. 3.1 shows the first-order filter functions, $T_{+\mu ij}^F(\vartheta)$ and $W_{+\mu ij}^F(\ell)$, with $1 \le \mu \le P = 7$ for the fiducial cosmology described in Sect. 3.4. Here I assume 3 redshift bins and seven free fiducial parameters, using 20 COSEBIs filters defined between $\theta_{\rm min} = 1'$ and $\theta_{\rm max} = 400'$. The redshift bins were chosen such that they contain an equal number of galaxies. Since the filters are designed to maximize the information obtained, their shape shows where most of the information in $\xi_+(\vartheta)$ or $P_{\rm E}(\ell)$ lies. Here we choose to only show the first-order filters, although later on I will use the first and second order as well as the combination of both to obtain the figure-of-merit. The general trend of $T^{\rm F}$ shows that there is more information about all of the parameters in the higher-redshift bins and on smaller angular scales.

Each individual parameter shows a different pattern for each of the redshift pairs.

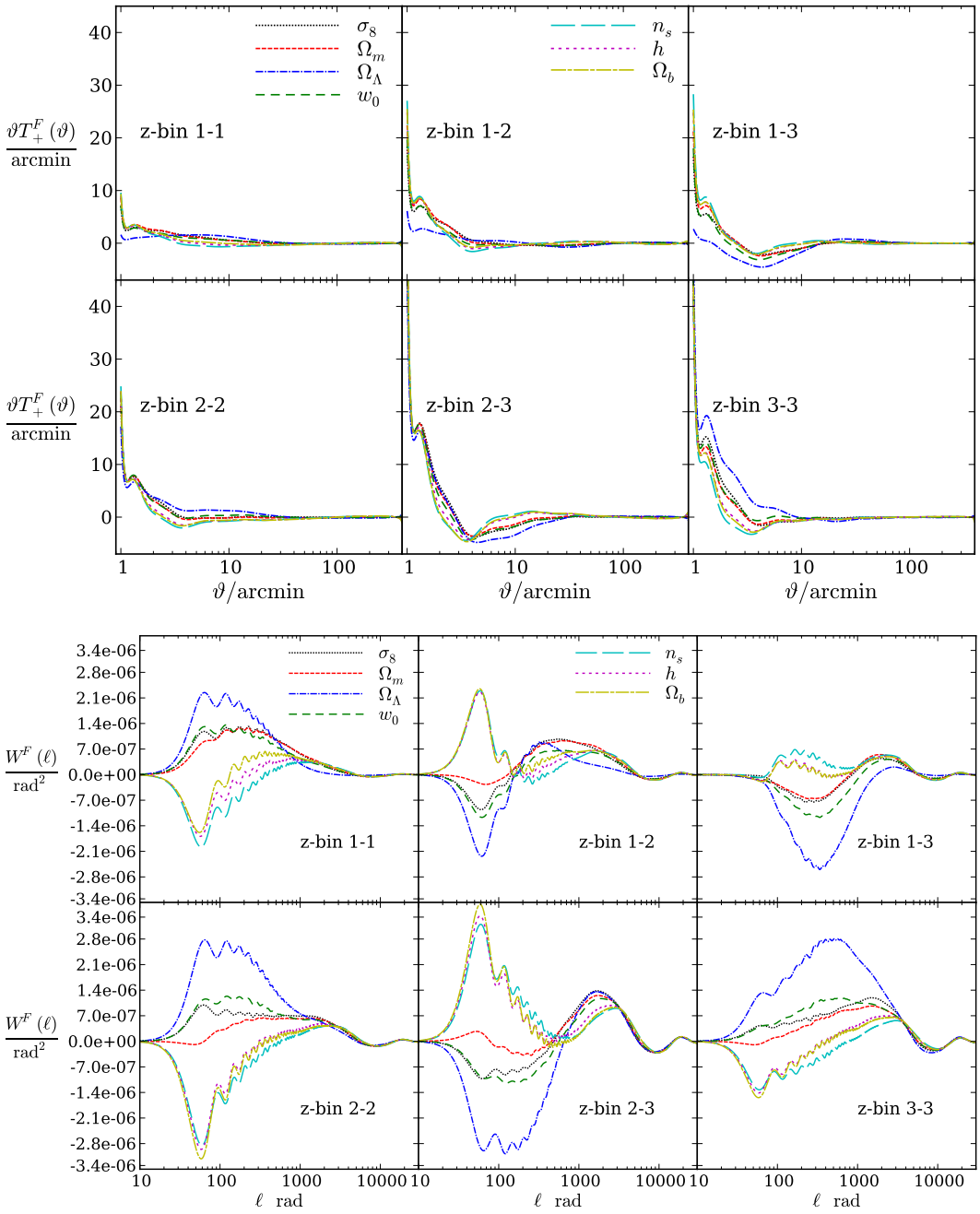

Figure 3.1: The filter functions $T^F_{+\mu ij}(\vartheta)$ (top panel) and $W^F_{\mu ij}(\ell)$ (bottom panel) are related to their progenitors via Eq. (3.20). For clarity I show $\vartheta T^F_{+\mu ij}(\vartheta)$ on a logarithmic ϑ-scale. 3 redshift bins and 20 COSEBIs filters defined between $\theta_{\rm min} = 1'$ and $\theta_{\rm max} = 400'$ are considered. Each parameter has a different filter function for each redshift pair bin (z-bin). A comparison between these functions and their progenitors (see Schneider et al. 2010 and Asgari et al. 2012) show that they have considerably fewer structures and oscillations. The filters for each parameter are normalized according to Eq. (3.23). Hence one can compare how effective each z-bin pair is for constraining the parameters.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
σ_8	43.7	−59.2	53.4	−36.0	17.2	−4.8	0.7	−1.9	4.5	−5.9	5.8	−4.6	2.8	−1.3	0.3
Ω_m	42.1	−57.8	53.5	−37.9	20.1	−7.2	1.8	−1.9	4.2	−5.9	6.0	−4.8	3.0	−1.3	0.3
Ω_Λ	48.5	−61.4	51.0	−31.3	13.4	−3.1	0.3	−1.8	4.3	−5.7	5.6	−4.4	2.7	−1.2	0.3
σ_8, σ_8	46.2	−61.2	52.8	−32.5	12.5	−1.1	−1.0	−1.7	4.7	−6.0	5.6	−4.2	2.5	−1.1	0.3
$\sigma_8, \Omega_\mathrm{m}$	38.2	−54.5	53.9	−41.9	25.4	−11.3	3.7	−2.2	3.9	−5.8	6.2	−5.2	3.3	−1.5	0.4
σ_8, Ω_Λ	44.9	−61.0	53.8	−33.2	11.8	0.9	−3.2	−0.3	4.5	−6.3	5.8	−4.1	2.3	−0.9	0.2
$\Omega_\mathrm{m}, \Omega_\mathrm{m}$	47.6	−62.0	52.1	−30.8	10.9	−0.3	−1.2	−1.7	4.7	−5.8	5.4	−4.1	2.5	−1.1	0.3
$\Omega_\mathrm{m}, \Omega_\Lambda$	49.8	−63.7	51.1	−26.2	4.5	5.2	−3.8	−1.7	5.4	−6.1	4.9	−3.3	1.9	−0.8	0.2
$\Omega_\Lambda, \Omega_\Lambda$	48.9	−62.8	51.5	−28.7	8.4	1.8	−2.3	−1.4	4.7	−5.9	5.3	−3.8	2.3	−1.0	0.2

Table 3.3: The elements of the normalized compression matrix in percentage, $100 \times \mathbf{B}_{\mu n}$, for 1 redshift bin and three parameters. The first column shows the subscript of E^c_μ, where one parameter subscripts belong to the first order statistics, $\boldsymbol{F}$, and double subscripts belong to the second order statistics $\boldsymbol{S}$. The first row shows the value of n.

For example, the real space filters, T^F, for Ω_Λ have significantly higher amplitudes for combinations of redshift bins 2 and 3 compared to combinations with the lowest redshift bin. σ_8 and Ω_m filters closely follow each other, although in Fourier space, the differences are more pronounced for W^F. A closer inspection of the plots shows that, since these curves are not exactly the same and also evolve with redshift, it is possible to break their degeneracies which are present in non-tomographic studies (e.g., Van Waerbeke et al. 2001; Hoekstra et al. 2002; Jarvis et al. 2003; Hetterscheidt et al. 2007; Kilbinger et al. 2013). The oscillations of the W^F are a real feature of the CCOSEBIs weights and do not vanish when more COSEBIs are incorporated to calculate them.

The Tables 3.3 and 3.4 show the elements of the compression matrix, $\mathbf{B}$, for CCOSEBIs. According to Eq. (3.15), each row of $\mathbf{B}$ corresponds to the coefficients for making one of the CCOSEBIs statistics, E^c_μ, by linearly combining the COSEBIs E_n. The value of the elements of $\mathbf{B}$ show how important each COSEBIs mode is for building a CCOSEBIs mode. In both tables, the element values are much smaller for large n compared to smaller n. As a result we can safely (conservatively) take just the first 20 COSEBIs to build the compressed statistics.

3.5.2 Fisher analysis

We will now use Eq. (2.10) to find the Fisher matrix for COSEBIs and via Eq. (2.19) define a figure of merit.

I have shown in Asgari et al. (2012) that for a sufficiently large survey one can neglect the first term in Eq. (2.11) since it does not depend on the survey area (see Eq. 2.99), while the second term is proportional to the survey area. Therefore, I neglect the first term in this study.

Fig. 3.2 shows the dependence of f on the number of COSEBIs modes, n_max, for 8 redshift bins and 7 free parameters. The constraints get tighter as n_max increases and reach a saturation level for all cases. The solid curve shows how much information can be gained

	z-bin 1-1						z-bin 1-2						z-bin 2-2					
	n=1	n=2	n=3	n=4	n=5	n=6	n=1	n=2	n=3	n=4	n=5	n=6	n=1	n=2	n=3	n=4	n=5	n=6
σ_8	17.8	−17.4	10.4	−4.4	1.3	−0.2	22.2	−33.3	31.3	−20.3	8.4	−1.8	36.8	−48.2	41.4	−25.8	10.8	−2.3
Ω_m	19.4	−19.9	12.8	−6.1	2.2	−0.4	25.7	−37.9	35.7	−23.5	10.1	−2.2	33.2	−43.4	37.1	−23	9.6	−2.1
Ω_Λ	13.1	−8.2	1.7	0.2	0.1	−0.2	1.0	−13.6	20.4	−16.2	7.3	−1.6	50.7	−59.7	45.0	−24.3	8.9	−1.7
σ_8, σ_8	12.1	−10.6	4.5	−0.5	−0.5	0.2	13.3	−20.8	20.5	−14.1	6.2	−1.4	43.1	−56.7	48.5	−29.6	12.1	−2.5
σ_8, Ω_m	24.9	−26.2	17.9	−9.4	3.7	−0.8	37.4	−51.4	44.5	−26.5	10.0	−1.8	15.0	−21.6	21.1	−15.3	7.5	−1.9
σ_8, Ω_Λ	4.3	−4.0	1.1	1.1	−1.2	0.4	5.0	−8.0	10.2	−9.5	5.7	−1.7	44.9	−60.7	52.5	−31.9	12.7	−2.5
Ω_m, Ω_m	13.6	−12.8	6.3	−1.3	−0.3	0.2	17.7	−25.2	23.5	−15.7	6.9	−1.6	41.7	−54.8	46.4	−28	11.2	−2.3
Ω_m, Ω_Λ	9.4	−9.5	5.1	−1.1	−0.4	0.3	13.7	−19.4	18.1	−12.3	5.6	−1.3	44.1	−57.9	49.1	−29.5	11.6	−2.3
$\Omega_\Lambda, \Omega_\Lambda$	9.5	−7.5	2.1	1.0	−1.1	0.3	7.8	−14.2	15.8	−11.9	5.7	−1.4	46.5	−59.8	49.4	−29.1	11.3	−2.2

Table 3.4: The normalized compression matrix elements in percentage, $100 \times \mathbf{B}$, for 2 redshift bins and three parameters. The first column shows the subscript of E^c_μ, where one parameter subscripts belong to the first order statistics, $\boldsymbol{F}$, and double subscripts belong to the second order statistics $\boldsymbol{S}$. The absolute values of the columns of $\mathbf{B}$ for each redshift pair decreases rapidly after n=3.

if the $8n_{\rm max} \times 9/2 = 36n_{\rm max}$ COSEBIs are used, which means, the maximum information, or minimum f value. The points show the amount of information in the first- ($\boldsymbol{F}$) and second-order ($\boldsymbol{S}$) CCOSEBIs, as well as the combination of both, denoted by $\boldsymbol{E}^c$ as before. The parameters used for calculating the covariance matrix to build $\boldsymbol{F}$, $\boldsymbol{S}$ and $\boldsymbol{E}^c$ are that of the fiducial cosmology which are slightly different from the assumed true parameters (see Table. 3.1). Nevertheless, the first-order CCOSEBIs are sufficient to reach a similar Fisher information level. Notice here that the $\boldsymbol{F}$ data vector for this case has 7 components, $\boldsymbol{S}$ has 28 components, and $\boldsymbol{E}^c$ have 35, while for $n_{\rm max} = 10$ the COSEBIs have 360 components, i.e., there is at least an order of magnitude difference between the number of observables for CCOSEBIs and COSEBIs. Hence, we can obtain the same accuracy of derived parameters with a highly significant reduction of observables.

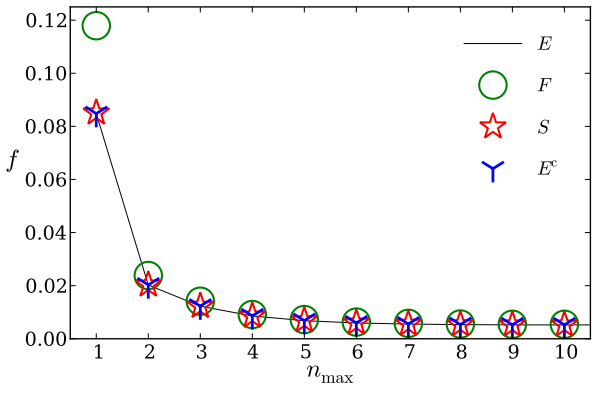

Figure 3.2: The figure-of-merit, f, as a function of the number of COSEBIs, $n_{\rm max}$, used. 7 free parameters listed in Table. 3.1, and 8 tomographic redshift bins are considered here. The solid line shows the result for using Log-COSEBIs with the true underlying cosmology. It also represents the maximum information level for a given $n_{\rm max}$. The circles, stars and the Y-shaped symbols represent the f-values for First order, Second order, and their combination $\boldsymbol{E}^c$, where $n_{\rm max}$ COSEBIs modes with the fiducial cosmological parameters are utilized in making them.

The strong reduction of observables needed to cover all the cosmological information is of great interest with regards to obtaining accurate covariances, and thus reliable confidence regions for cosmological parameters. Whereas analytical methods may be able to obtain

approximate covariances (see, for example, Takada & Jain 2009; Sato et al. 2009; Pielorz et al. 2010; Hilbert et al. 2011a; Takahashi et al. 2011, 2014, and references therein), an accurate covariance accounting for the complex survey geometry will probably require extensive simulations. Obtaining the covariance as sample variance from independent realizations of the simulated cosmology requires a number of realizations roughly proportional to the number of observables (see Hartlap et al. 2007; Taylor & Joachimi 2014). Hence, even a modest reduction in the number of relevant observables is useful. As we have seen above, the CCOSEBIs serve this purpose very well.

Whereas the construction of the CCOSEBIs requires information about the covariance, this does not have to be very accurate. In order to show how using a substantially wrong covariance in defining the compressed data vector impacts on the constraints, I artificially change the value of σ_ϵ which affects the diagonals of the covariance matrix according to Eq. (2.99). Fig. 3.3 shows f (defined in Eq. 2.19) for 7 free parameters, 5 redshift bins and 20 COSEBIs modes, as a function of the change in the parameter σ_ϵ. f is normalized by its minimum value, i.e., using COSEBIs with their true covariance, while σ_ϵ is normalized by its true value. The first-order statistics, $\boldsymbol{F}$, which has the same dimension as the parameter space, rapidly diverges from the true Fisher information limit, while the second order, $\boldsymbol{S}$, and $\boldsymbol{E}^{\rm c}$, which span a larger dimensional space, are much less sensitive to the errors of the COSEBIs covariance, used for constructing the CCOSEBIs. Even for a 16 times larger σ_ϵ the fractional difference between the optimal f and the measured one from $\boldsymbol{E}^{\rm c}$ is small. Hence, the consideration of the second-order statistics indeed provides a powerful mitigation for inaccurate covariances.

At this point, I should stress again that the accuracy with which the original covariance can be determined affects the efficiency of the data compression, in the sense that what fraction of the cosmological information contained in the original COSEBIs is preserved in the compressed COSEBIs. However, it does not bias the parameter determination, as the CCOSEBIs are linear combinations of the original ones.

Hence we conclude that the method proposed here – constructing new observables using an approximate covariance, and employing these for cosmological parameter studies – yields a very promising tool for an effective reduction in the necessary efforts for constructing accurate covariances. This data compression will also be of great help if the covariances are to be obtained from the data themselves, by subdividing the survey region, calculating the sample variance on each sub-survey, and scaling the result with the ratio of sub-survey to survey area.

3.6 Band power

As mentioned in Sect. 3.3, any filter function defined on a finite angular interval which satisfies the constraints in Eq. (2.77) can be expressed in terms of the COSEBIs filters. A particular filter one might be interested in is a top-hat function in Fourier space, corresponding to a band

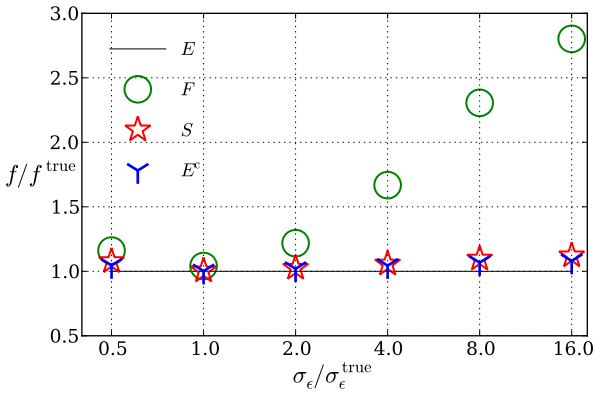

Figure 3.3: The figure-of-merit, f, as a function of σ_ϵ. f is normalized by its minimum value which corresponds to using COSEBIs with the correct covariance (the solid line). The intrinsic ellipticity dispersion of galaxies, σ_ϵ, is varied with respect to its true value, 0.3, to show the effects of using a wrong covariance. The markers show the value of f for first order, $\boldsymbol{F}$, second order, $\boldsymbol{S}$ and the combination of both $\boldsymbol{E}^{\rm c}$ CCOSEBIs.

power (for example, Brown et al. 2003; Hikage et al. 2011). In this section I will study how well band powers can be approximated from correlation functions measured on a finite interval with clean E-/B-mode separation.

Thus, let $\hat{W}(\ell)$ be a target filter function in Fourier space, and let us design a filter that approximates $\hat{W}(\ell)$ as closely as possible. That means we want to find a filter which minimizes

$$\Delta = \int \mathrm{d}\ell\, \ell\, \left[W(\ell) - \hat{W}(\ell)\right]^2 , \tag{3.24}$$

where $W(\ell)$ is a linear combination of the $W_n(\ell)$,

$$W(\ell) = \sum_n c_n W_n(\ell) . \tag{3.25}$$

Note here that the Δ integral can be defined with a different weighting of ℓ, for example, one can replace $\mathrm{d}\ell\, \ell$ in Eq. (3.24) with $\mathrm{d}\ln\ell$ or simply $\mathrm{d}\ell$. Doing so does not affect the final estimation accuracy of $\hat{W}(\ell)$ significantly, but may be numerically advantageous.

I want to find the coefficients c_n that minimize Δ; setting the derivatives of Δ with respect to c_m to zero, we find

$$\sum_n c_n \int \mathrm{d}\ell\, \ell\, W_n(\ell) W_m(\ell) = \int \mathrm{d}\ell\, \ell\, W_m(\ell) \hat{W}(\ell) . \tag{3.26}$$

By defining the matrix

$$\mathbf{Q}_{nm} \equiv \int \mathrm{d}\ell\, \ell\, W_n(\ell) W_m(\ell) \tag{3.27}$$

and the vector

$$V_n \equiv \int \mathrm{d}\ell\, \ell\, W_n(\ell) \hat{W}(\ell) , \tag{3.28}$$

we can rewrite Eq. (3.26) in matrix form,

$$\mathbf{Q}\boldsymbol{c} = \boldsymbol{V} \implies \boldsymbol{c} = \mathbf{Q}^{-1}\boldsymbol{V} . \tag{3.29}$$

The minimum of Δ for this solution is

$$\Delta_{\rm min} = \int \mathrm{d}\ell\, \ell\, \hat{W}^2(\ell) - \boldsymbol{V}\mathbf{Q}^{-1}\boldsymbol{V}\,, \tag{3.30}$$

and we quantify the relative deviation of the closest filter W to $\hat{W}$ by

$$\delta_{\rm min} = \sqrt{\Delta_{\rm min}\left[\int \mathrm{d}\ell\, \ell\, \hat{W}^2(\ell)\right]^{-1}}\,. \tag{3.31}$$

Filters which satisfy Eq. (2.77) and vanish outside of the angular range $[\theta_{\rm min}\,, \theta_{\rm max}]$, are the only ones that can be represented by COSEBIs. Hence, applying any filter that does not satisfy these conditions, on either power spectra or 2PCFs, results in spillage from outside the measured angular range. A top-hat function in Fourier space is an example of a filter which cannot be well represented by weight functions which correspond to a finite range in real space. A top-hat function in Fourier is defined as $\hat{W}(\ell) = 1$ between $\ell_{\rm min}$ and $\ell_{\rm max}$ and zero otherwise. The real space version, $\hat{T}_+(\vartheta)$, of such a function is

$$\hat{T}_+(\vartheta) = \int_0^\infty \mathrm{d}\ell\, \ell\, \mathrm{J}_0(\ell\vartheta)\hat{\mathrm{W}}(\ell) = \int_{\ell_{\rm min}}^{\ell_{\rm max}} \mathrm{d}\ell\, \ell\, \mathrm{J}_0(\ell\vartheta) = \frac{1}{\vartheta}[\ell_{\rm max}\,\mathrm{J}_1(\ell_{\rm max}\vartheta) - \ell_{\rm min}\,\mathrm{J}_1(\ell_{\rm min}\vartheta)]\,. \tag{3.32}$$

This equation shows that the function $\hat{T}_+(\vartheta)$ is manifestly non-zero for $\vartheta < \theta_{\rm min}$ and $\vartheta > \theta_{\rm max}$, which implies that the band power cannot be represented in terms of E-/B-mode separating combinations of correlation functions over a finite angular interval.

Using Parseval's theorem we can find a lower bound for Δ,

$$\begin{aligned}\Delta &= \int_0^\infty \mathrm{d}\ell\, \ell \left[W(\ell) - \hat{W}(\ell)\right]^2 = (2\pi)^2 \int_0^\infty \mathrm{d}\vartheta\, \vartheta \left[T_+(\vartheta) - \hat{T}_+(\vartheta)\right]^2 \\ &= (2\pi)^2 \left\{ \int_0^{\theta_{\rm min}} \mathrm{d}\vartheta\, \vartheta\, \hat{T}_+^2(\vartheta) + \int_{\theta_{\rm max}}^\infty \mathrm{d}\vartheta\, \vartheta\, \hat{T}_+^2(\vartheta) + \int_{\theta_{\rm min}}^{\theta_{\rm max}} \mathrm{d}\vartheta\, \vartheta \left[T_+(\vartheta) - \hat{T}_+(\vartheta)\right]^2 \right\},\end{aligned} \tag{3.33}$$

where $T_+(\vartheta)$ is the real space form of $W(\ell)$. The sum of the first two integrals in Eq. (3.33) is the absolute lower bound on Δ, since the last integral in that equation is non-negative, irrespective of the choice of $T_+(\vartheta)$. Hence, the lower bound for $\delta_{\rm min}$ is

$$\delta_{\rm min} \geq \delta_{\rm LB} \equiv \sqrt{\frac{8\pi^2}{\ell_{\rm max}^2 - \ell_{\rm min}^2} \left\{ \int_0^{\theta_{\rm min}} \mathrm{d}\vartheta\, \vartheta\, \hat{T}_+^2(\vartheta) + \int_{\theta_{\rm max}}^\infty \mathrm{d}\vartheta\, \vartheta\, \hat{T}_+^2(\vartheta) \right\}}\,. \tag{3.34}$$

In order to reach the absolute lower bound, the last integral in Eq. (3.33) should vanish. It is necessary and sufficient for $\hat{T}_+(\vartheta)$ to satisfy the conditions in Eq. (2.77) for that to happen, since then $\hat{T}_+(\vartheta)$ in the range $\theta_{\rm min} \leq \vartheta \leq \theta_{\rm max}$ can be represented as a sum over the COSEBIs weights T_{+n}. Inserting the analytic form of $\hat{T}_+(\vartheta)$ from Eq. (3.32) into Eq. (2.77) results in the

following two conditions:

$$\mathrm{I} := \mathrm{J}_0(\ell_{\min}\theta_{\max}) - \mathrm{J}_0(\ell_{\min}\theta_{\min}) - \mathrm{J}_0(\ell_{\max}\theta_{\max}) + \mathrm{J}_0(\ell_{\max}\theta_{\min}) = 0 \,, \tag{3.35}$$

and

$$\begin{aligned}\mathrm{II} := {} & \theta_{\min}^2 \mathrm{J}_0(\ell_{\min}\theta_{\min}) - \theta_{\max}^2 \mathrm{J}_0(\ell_{\min}\theta_{\max}) + \frac{2}{\ell_{\min}}[\mathrm{J}_1(\ell_{\min}\theta_{\max}) - \mathrm{J}_1(\ell_{\min}\theta_{\min})] \\ & - \theta_{\min}^2 \mathrm{J}_0(\ell_{\max}\theta_{\min}) + \theta_{\max}^2 \mathrm{J}_0(\ell_{\max}\theta_{\max}) - \frac{2}{\ell_{\max}}[\mathrm{J}_1(\ell_{\max}\theta_{\max}) - \mathrm{J}_1(\ell_{\max}\theta_{\min})] = 0 \,,\end{aligned} \tag{3.36}$$

which should be simultaneously true for $\delta_{\min} = \delta_{\mathrm{LB}}$. However, most combinations of $\theta_{\min}$, $\theta_{\max}$, $\ell_{\min}$ and $\ell_{\max}$ do not satisfy these conditions. As a result in general $\delta_{\min} > \delta_{\mathrm{LB}}$ for most cases.

Once the COSEBIs, E_n, are measured from the data, the band power can be estimated by linearly combining them,

$$\begin{aligned}\hat{E} &= \frac{1}{2\pi}\int \mathrm{d}\ell\, \ell\, \hat{W}(\ell) P_{\mathrm{E}}(\ell) \approx \frac{1}{2\pi}\int \mathrm{d}\ell\, \ell\, W(\ell) P_{\mathrm{E}}(\ell) \\ &= \sum_n c_n \frac{1}{2\pi}\int \mathrm{d}\ell\, \ell\, W_n(\ell) P_{\mathrm{E}}(\ell) = \sum_n c_n E_n \,,\end{aligned} \tag{3.37}$$

Fig. 3.4 depicts the convergence to an estimated $W(\ell)$ by increasing the number of COSEBIs modes. The number of COSEBIs needed for convergence is substantially higher than the number needed for constraining parameters with one redshift bin.

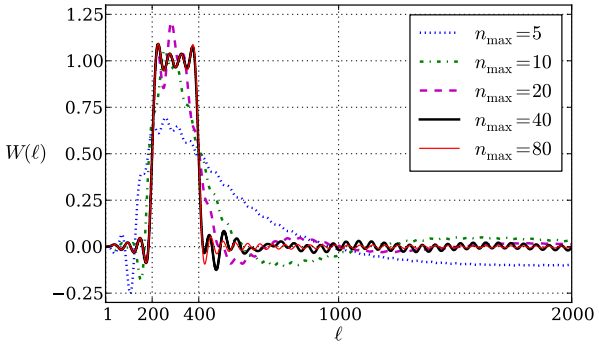

Figure 3.4: The estimated top-hat filter function with $\ell_{\min} = 200$ and $\ell_{\max} = 400$, from $n_{\max}$ COSEBIs filters defined on $1' < \vartheta < 400'$. The changes between using 40 and 80 COSEBIs filters are small, so that no better representation is obtained by using an even higher value of $n_{\max}$.

Fig. 3.5 shows the dependence of $\delta_{\min}$ on the number $n_{\max}$ of COSEBIs. Here I demonstrate that for all ranges in ℓ considered, a saturation level is reached, as a result, adding more COSEBIs filters will not lead to a smaller difference between the estimated W and the top hat filter. Furthermore, in Table. 3.5 I show the value of $\delta_{\min}$ for 80 COSEBIs which can be compared to its lower bound, δ_{LB} (see Eq. 3.34), and the relative difference between the estimated band power and its true value, $\delta_{\mathrm{band}} = |(\hat{E} - E)/\hat{E}|$. The saturated $\delta_{\min}$ values are larger than but close to δ_{LB}. The difference between the two arises from violating conditions

(3.35) and (3.36). In Table. 3.5, I use three ℓ-weighting schemes, which do not change the $\delta_{\rm min}$ values significantly. However, the estimated band-power deviations, $\delta_{\rm band}$, can vary by more than a few percent between the cases. This is due to the spillage of the estimated band power and the fact that the ℓ-weighting scheme decides which way the spillage is directed. The $\delta_{\rm band}$ values are cosmology dependent and can be very different for a power spectrum with more features.

It is interesting to note that the deviations $\delta_{\rm band}$ of the estimated band powers from their true values are in some cases *smaller* than the relative deviation $\delta_{\rm min}$ between the top-hat filter and the best representation of the top hat by COSEBIs weight functions. This, however, is an effect of the properties of the power spectrum in the assumed cosmological model: the power spectrum is sufficiently smooth that the spilling caused by the effective weight $W(\ell)$ out of, and into the range of the top hat, can compensate each other (see Schneider et al. 2002b, for a related discussion on band powers in cosmic shear analysis). Hence, the fact that $\delta_{\rm band}$ is relatively small is *not* a statement about the accuracy of the method of band-power estimates, but rather a consequence of the properties of the power spectrum. But the latter should be probed by estimating the band power. Thus, it would be strongly misleading to judge the accuracy of the method on presumed properties that rather ought to be investigated. Indeed, it is the quantity $\delta_{\rm min}$, which yields an estimate on the accuracy with which band powers can be obtained.

Becker & Rozo (2014) constructed slightly different band powers, with a weight function in ℓ-space of a log-normal form. As expected, their smoother weight $\hat{W}$, which does not impose strict ℓ cuts in Fourier space, compared to the top-hat considered here, leads to less spilling of the best-fitting weight function in ϑ-space. Or in other words, the log-normal weight $\hat{W}$ can be more accurately represented by the Fourier transform of the shear correlation function on a finite interval, at the prize of a non-diagonal covariance of band powers even in the case of Gaussian shear fields.

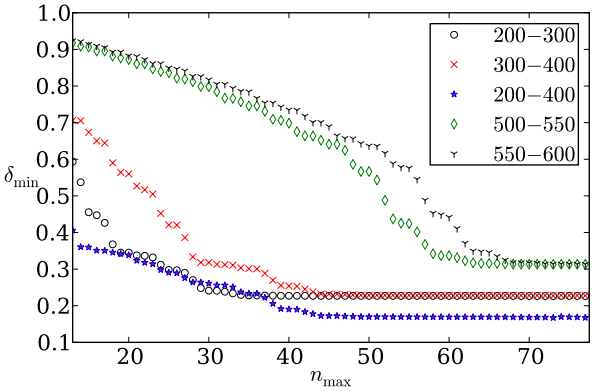

Figure 3.5: The relative difference, $\delta_{\rm min}$, between the estimated top hat and the input as a function of the number of COSEBIs filters, $n_{\rm max}$, utilized for a few ℓ-ranges. In all cases the saturation level is reached before $n_{\rm max} = 77$. The minimum value of $\delta_{\rm min}$ is shown in Table. 3.5. In general, a higher number of modes are needed for a narrower top-hat filter, which is due to the spillage beyond the observed angular range (see Eq. 3.34).

$l_{\rm min} - l_{\rm max}$	$\delta_{\rm min}$	$\delta_{\rm LB}$	$\delta_{\rm band}, {\rm d}\ell\,\ell$	$\delta_{\rm band}, {\rm d}\ell$	$\delta_{\rm band}, {\rm d}\ell/\ell$
200 – 300	2.28	2.27	3.31	0.72	2.29
300 – 400	2.30	2.29	6.78	1.74	2.99
200 – 400	1.72	1.71	4.94	1.20	2.62
500 – 550	3.14	3.14	10.24	3.03	4.27
550 – 600	3.14	3.14	14.90	7.83	9.14

Table 3.5: Examples of band-power estimation from 80 COSEBIs for $\xi_{\pm}(\vartheta)$ in the interval $\vartheta \in [1', 400']$. The first column shows the ℓ-range of the top-hat function $\hat{W}(\ell)$. The rest of the columns show percentage values for minimum relative difference between the estimated and the top hat, $\delta_{\rm min}$, the absolute minimum (Lower Bound) value for $\delta_{\rm min}$, $\delta_{\rm LB}$ (see Eq. 3.34), and the relative difference between the estimated band power and its true value, $\delta_{\rm band} = |(\hat{E} - E)/\hat{E}|$, for different ℓ weightings, respectively. $\delta_{\rm min}$ and $\delta_{\rm LB}$ values correspond to the ${\rm d}\ell\,\ell$ weighting. The values of these quantities for other cases are similar. The cosmological model used here is the fiducial model from Table. 3.1, with one redshift bin.

3.7 Summary and Discussion

Data compression is an important challenge to tackle for future cosmological surveys, it is essential for estimating accurate covariances. Current cosmological surveys such as Planck provide us with tight constraints on most cosmological parameters. This was my motivation to define combinations of statistics inspired by their low-order Taylor expansion around a fiducial cosmological model. The strategy for finding the compressed statistics involves first- and second-order derivatives of a parent statistics with respect to the free parameters as well as their covariance. The statistics corresponding to the first order derivatives, $\boldsymbol{F}$, have the same dimension as the parameter space, while the statistics derived from second-order derivatives, $\boldsymbol{S}$, provide a possibility to span a larger-dimensional space. Consequently, $\boldsymbol{F}$ is more sensitive to the choice of the fiducial cosmology and covariance. The combination of $\boldsymbol{F}$ and $\boldsymbol{S}$, enables one to use well-defined and motivated sets of statistics which alleviate many of the data analysis problems. In total the number of compressed statistics is $P(P+3)/2$, where P is the number of free parameters in the model.

In the case of a cosmic shear analysis, the COSEBIs already provide an effective compression compared to other two-point statistics, for example, the shear two-point correlation functions. However, adding tomographic bins, which is necessary for intrinsic alignment corrections (see King & Schneider 2003; Heymans et al. 2004), substantially increases the number of observables. As a result, further data compression is required. I applied the compression formalism to Log-COSEBIs to study its properties. I found that for a well-estimated COSEBIs covariance matrix, the first-order compressed statistics are sufficient. However, as mentioned above, the accuracy of covariance estimations from simulations depend on the number of observables incorporated. The higher this number is, the more simulations

are needed which rapidly becomes too expensive. Consequently, I used highly inaccurate covariances for defining the compressed COSEBIs (CCOSEBIs), to test their efficiency for such cases. I found that the figure-of-merit obtained from the first-order CCOSEBIs deviates substantially from the optimal information level as the difference between the assumed COSEBIs covariance and their true covariance increases. In contrast, the set of second-order CCOSEBIs is far less sensitive to the choice of covariance, owing to its larger dimensionality. The combination of both is basically insensitive to the accuracy of the covariance, at least in the framework of the simple model that I have tested here. Consequently, I propose that this strategy is applicable for the future data analysis, which I present in Chapter 4. Note that the first-order CCOSEBIs is equivalent to the Karhunen–Loève data compression (with parameter-independent covariance) in the case where the covariance is accurately known (Tegmark et al. 1997).

In this chapter I used a Fisher analysis, which assumes the parameters have a normal distribution, to compare the constrains from COSEBIs and CCOSEBIs. Both Fisher matrix and $\boldsymbol{F}$, the first order compressed statistics depend only on the first order derivatives and the covariance. If the fiducial cosmology coincides with the truth and the covariance is exact, then the $\boldsymbol{F}$ is equivalent to a Fisher formalism, since in this case the derivative matrix of $\boldsymbol{F}$ is equal to its covariance matrix which is consequently equal to the Fisher matrix. However, when the covariance deviates from the truth the differences become visible. For my future studies I plan to use likelihood analysis which does not make assumptions about the Gaussianity of the likelihood with respect to the model parameters.

The COSEBIs filter functions form a complete basis for any filter that satisfies Eqs. (2.77) which are necessary and sufficient conditions for a clean E-/B-separation on a finite interval, together with the condition that the filters should also vanish outside of the finite angular range. Consequently, any filter that satisfies these conditions can be represented by a linear combination of the COSEBIs filters. In this chapter I showed how any given weight function can be mimicked by COSEBIs weights. In particular, I tried to represent top-hat filters in Fourier space using this strategy. I found that, due to the infinite support of a Fourier top-hat in real space, an accurate representations of them is impossible. This task becomes harder as the top hat and the angular range get narrower. Consequently, band convergence power spectra estimated from finite angular range information will suffer from spillage, hence they will be inaccurate and biased, in a way that is dependent on the power spectrum – the quantity to be probed. Hence, we caution against using narrow-band power spectra for cosmic shear analysis. In the final chapter of this thesis, I will use band power spectra in the form of pseudo Cls, for cosmic shear analysis and show that this effect causes biases in the estimated cosmological parameters.

Chapter 4

A Cosmic Shear Analysis of CFHTLenS

The Canada-France Hawaii Telescope Lensing Survey (CFHTLenS[1]) is the largest completed survey optimised for cosmic shear analysis, to date. In the first chapter we saw the cosmological results derived from the analysis of this survey, which I will build upon throughout this Chapter using COSEBIs (see Sect. 2.6) and CCOSEBIs (see Sect. 3.3) as statistical tools. The first step taken is to set up and test the pipelines analytically (Sect. 4.2) and then with simulations in Sect. 4.3. The predictions for the values of COSEBIs and CCOSEBIs are compared to their measured values from the simulations in Sect. 4.3.1 and Sect. 4.3.2 respectively. An MCMC algorithm is used to sample the likelihood for σ_8, the standard deviation of perturbations in a sphere of radius $8h^{-1}$Mpc today, and Ω_m, the matter density parameter, from the simulations in Sect. 4.3.3 and found to recreate the input cosmological parameters. Finally in Sect. 4.4 the pipelines are applied to the CFHTLenS data blindly to find the likelihood of σ_8 and Ω_m for tomographic and non-tomographic cases.

4.1 CFHTLenS: An Introduction

CFHTLenS spans 154 square degrees in 5 optical bands (u^*, g', r', i', z'). The team has already published several papers on the data analysis. Kilbinger et al. (2013) performed a two dimensional analysis of the data using several cosmic shear methods, including the COSEBIs, the statistics that forms the focus of this work. The aim of their work was to use a large angular range to estimate cosmological parameters, but they faced difficulties estimating the COSEBIs from their mock data, known as the Clone simulations (see Sect. 4.3). The main reason for their difficulties was the fact that the accuracy of the simulations for very large angular scales is questionable, due to the limited box size. Consequently, they did not use COSEBIs for

[1] http://www.cfhtlens.org/

z-bin 1	z-bin 2	z-bin 3	z-bin 4
[0.58, 0.72]	[0.72, 0.87]	[0.86, 1.02]	[1.02, 1.30]

Table 4.1: The redshift bins for tomography. Although CFHTLenS contains information about galaxies with lower redshifts, only galaxies between redshifts 0.58 and 1.3 are considered. The reason behind this choice is that the galaxy intrinsic ellipticity contamination between these redshift bins is insignificant in the analysis (see Heymans et al. 2013).

σ_8	Ω_m	Ω_{tot}	w_0	n_s	h	Ω_b
[0.2, 1.5]	[0.05, 1.0]	1	−1.0	0.96	0.701	0.046

Table 4.2: The prior on the free cosmological parameters and the value of the fixed cosmological parameters. Flat ΛCDM are the only models considered. The free parameters are σ_8 and Ω_m with flat priors given above. Here $\Omega_{tot} = \Omega_m + \Omega_\Lambda$.

their final analysis of the data. However, Fu et al. (2014) used COSEBIs in conjunction with third order aperture mass moments to constrain cosmological parameters for a single redshift distribution.

Heymans et al. (2013) performed a tomographic analysis of the CFHTLenS data using only shear two point correlation functions (2PCFs). As a result they did not incorporate any E-/B-mode decomposition methods in their analysis. They explored the effects of intrinsic alignments on the estimated shear correlation functions and found that for their highest four redshift bins the intrinsic-intrinsic and intrinsic-shear contribution was negligible. As a result, in this thesis I will limit the analysis to these redshift bins (see Table. 4.1). This limitation results in a weaker constraining power in the analysis. For the purposes of this thesis where I apply COSEBIs and CCOSEBIs to tomographic data for the first time, I will fix all the cosmological parameters to their best fit values from the previous CFHTLenS analysis (see Table. 4.2) and only constrain σ_8 and Ω_m. Fixing the parameters to their best fit values from the previous CFHTLenS analysis provides the chance to compare this work with them in a straightforward fashion. My future analysis with include the full redshift range and a larger set of free cosmological parameters.

An overview of the CFHTLenS data set is given in Heymans et al. (2012). Erben et al. (2013) detail the data reduction while Hildebrandt et al. (2012) and Miller et al. (2013) discuss the photometric redshift and galaxy shape measurements, respectively. According to Heymans et al. (2012), 75% of the fields pass the systematic tests. Hence in the following I will only use these remaining 129 fields in the analysis.

The galaxy shear calibration in Heymans et al. (2012) accounts for additive and multiplicative biases between the observed, ϵ_{obs}, and the true ellipticities, ϵ_{true},

$$\epsilon_{obs} = (1 + m)\, \epsilon_{true} + c \,, \tag{4.1}$$

In general, a perfect data reduction and PSF correction results in $m = c = 0$, however, so far none of the available pipelines have been able to reach this ideal. The multiplicative bias comes from inaccuracies in shape measurements, due to shape noise. It is estimated from galaxy image simulations (see Miller et al. 2013). The value of the multiplicative bias is significant and depends on signal-to-noise as well as galaxy size. Miller et al. (2013) model m as a function of signal-to-noise ratio and size with two parameters fitted from image simulations. The multiplicative bias on average results in a 6% correction to the measured ellipticities. The additive bias found from the image simulations is equal to zero, unlike its value measured from the data. The CFHTLenS team estimated c is zero for ϵ_1 and 2×10^{-3} for ϵ_2 on average. The origin of the additive bias is unknown and its value is calibrated from the data empirically. Heymans et al. (2012) found that the value of the additive bias is independent of the PSF size, PSF ellipticity and galaxy type. However, they note that it depends on the signal-to-noise ratio and galaxy size. The smallest and brightest galaxies have the largest contribution to the additive bias as they have shown. They modelled c_2 as a function of signal-to-noise ratio and galaxy size with 4 free parameters which are fitted from a likelihood analysis of the data.

While the additive bias is subtracted from the observed ϵ_2 directly, the effect of the multiplicative bias is applied globally as explained in Miller et al. (2013). The measured 2PCFs (see Eq. 2.53 for their definition) are divided by the calibration function,

$$1 + K(\vartheta) = \frac{\sum_{ab} w_a w_b (1 + m_a)(1 + m_b)}{\sum_{ab} w_a w_b} \,, \tag{4.2}$$

where w_a and m_a are the weight and the multiplicative bias associated with the galaxy at position a. The sum is carried out over all pairs of galaxies with a separation within a bin around ϑ. Each galaxy has an inverse variance weight associated with it. Less noisy galaxy shapes have a larger weight value, ergo they are more important in the analysis. The definition of w can be found in Miller et al. (2013).

4.2 A Blind Analysis: Tests on COSEBIs

Before applying COSEBIs to CFHTLenS we will go through a series of tests on COSEBIs to understand the effects of noise and discrete integration on the measurements and the analysis. In my previous paper (Asgari et al. 2012) we tested COSEBIs numerically, assuming a perfect knowledge of the input quantities. In Asgari et al. (2012) I used

$$E_n^{(ij)} = \int_0^\infty \frac{\mathrm{d}\ell\, \ell}{2\pi} P_{\rm E}^{(ij)}(\ell) W_n(\ell) \,, \tag{4.3}$$

to find E-mode COSEBIs which is more convenient to use for a theoretical analysis, since most theories provide us with an input power spectrum. However, in practice shear 2PCFs, are more

straightforward to measure from data, from which COSEBIs can be inferred via,

$$E_n^{(ij)} = \frac{1}{2}\int_{\theta_{\rm min}}^{\theta_{\rm max}} {\rm d}\vartheta\,\vartheta\,[T_{+n}(\vartheta)\,\xi_+^{(ij)}(\vartheta) + T_{-n}(\vartheta)\,\xi_-^{(ij)}(\vartheta)]\,, \tag{4.4}$$

$$B_n^{(ij)} = \frac{1}{2}\int_{\theta_{\rm min}}^{\theta_{\rm max}} {\rm d}\vartheta\,\vartheta\,[T_{+n}(\vartheta)\,\xi_+^{(ij)}(\vartheta) - T_{-n}(\vartheta)\,\xi_-^{(ij)}(\vartheta)]\,, \tag{4.5}$$

where $E_n^{(ij)}$ and $B_n^{(ij)}$ are the E and B-mode COSEBIs for redshift bins i and j , $T_{\pm n}(\vartheta)$ are the COSEBIs filter functions which are zero outside the angular range $[\theta_{\rm min}, \theta_{\rm max}]$ and guarantee a clean E/B-mode separation of the 2PCFs, $\xi_\pm^{(ij)}(\vartheta)$, for the redshift bin pair ij (see Sect. 2.6). As a result the first test is to check if the two equations above result in the same E_n. In principle, the E_n estimated from Eq. (4.3) and Eq. (4.4) should be identical. However, since these integrals are estimated numerically and the fact that the integral in Eq. (4.3) has an infinite upper limit, the results may differ. For this test I choose an angular range of $[1', 400']$. Fig. 4.1 shows the residual ratio of E_n calculated from Eq. (4.3) and Eq. (4.4) for $n = 1 - 7$ and the fiducial parameters of Table. 3.1 in Chapter 3 and Table. 4.3. No tomography is assumed here. The redshift distribution follows,

$$p(z) \propto \left(\frac{z}{z_0}\right)^\alpha \exp\left(-\frac{z}{z_0}\right)^\beta\,, \tag{4.6}$$

which is a generalization of Brainerd et al. (1996), with parameters given in Table. 4.3. As we can see in this figure, the values of E_n from the two methods are in good agreement. The difference between them is significantly lower than percent level.

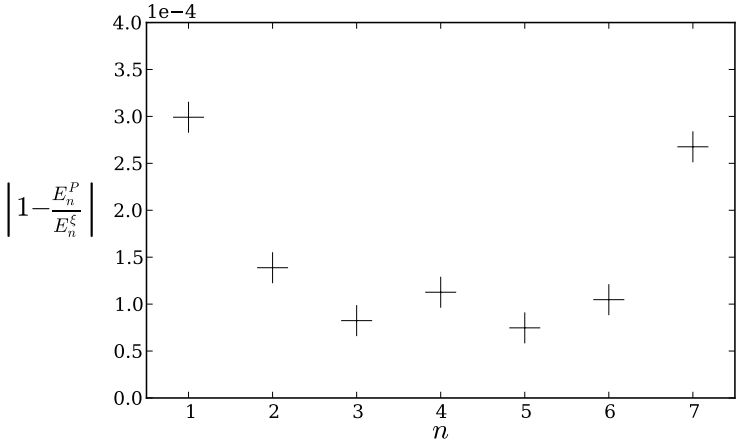

Figure 4.1: A comparison between two methods of finding E-COSEBIs. E_n^P is calculated from Eq. (4.3), while E_n^ξ estimated from Eq. (4.4). E_n with $n = 1 - 7$ are shown here for the fiducial cosmology in Table. 3.1, the redshift parameters of a CFHTLenS-like survey in Table. 4.3 and an angular range of $[1', 400']$.

4.2.1 From Smooth Integration to Noisy Trapezoidal

Since so far we have assumed to have a perfect knowledge of the 2PCFs over the angular range considered, a Gaussian integration method (see Press et al. 2002) between two extrema of the integrand is employed to evaluate E_n in both cases. Nevertheless, in practice we only

z-distribution parameters					survey parameters		
α	β	z_0	$z_{\rm min}$	$z_{\rm max}$	A	σ_ϵ	$\bar{n}$
0.836	3.425	1.171	0.2	1.5	154	0.279	11

Table 4.3: The parameters of CFHTLenS-like survey. The true parameters of the CFHTLenS survey will be discussed later in this Chapter. α, β, and z_0 characterize the total redshift distribution of sources, while $z_{\rm min}$ and $z_{\rm max}$ indicate the minimum and the maximum redshifts of the sources used for the cosmic shear analysis. Here A is the survey area in units of deg^2, σ_ϵ is the galaxy intrinsic ellipticity dispersion, and $\bar{n}$ is the effective mean number density of sources per square arcminute in the field.

have the values of 2PCFs in angular bins or at certain θ values. Consequently, we need to use a different integration routine to evaluate E_n from Eq. (4.4) for real data. The most straightforward integration method is the trapezoidal method for a linearly binned data. Our second test focuses on finding how many linear angular bins are needed to reach a certain accuracy in determining E_n.

Fig. 4.2 shows the relative value of E_n to the number of angular bins used in the trapezoidal integration. All the E_n values are normalized by their true value, calculated from the convergence power spectrum using Eq. (4.3). In Asgari et al. (2012) I have shown the sensitivity of the first five E_n to σ_8, Ω_m and n_s, also shown in Fig. 2.11. If σ_8 changes by 1% then E_1 will change by about 2%. As a result I choose to have an accuracy of 0.5% for E_n which translates to less than 10000 bins for E_7 which is the highest COSEBIs mode I will use in this analysis. As can be seen in Fig. 4.2 a higher number of angular bins are required for the higher COSEBIs modes, to reach the same amount of accuracy. The reason for this behavior is that, the $T_{+n}(\vartheta)$ functions have $n+1$ roots in their range of support and oscillate around them. Consequently, the higher modes are more sensitive to the number of θ bins incorporated in their integral (see Fig. 2.5).

The next test is to find the effects of noise on the estimated E_n. Since the number of bins needed to reach the accuracy desired is high, the shot noise term of the correlation functions covariance dominates the other terms in this case. We will therefore choose to ignore other sources of noise for this test, using only the shot noise in the covariance. at this stage. We can make a noisy $\xi_\pm(\theta)$ mock data set from

$$\xi_\pm(\theta)^{\rm Noisy} = \xi_\pm(\theta) + N \times R \,, \tag{4.7}$$

where

$$N = \sqrt{\frac{\sigma_\epsilon^4}{4\pi A \bar{n}^2 \Delta\theta\, \theta}} \,, \tag{4.8}$$

is the square root of the shot noise term in the 2PCFs covariance (see Joachimi & Schneider 2008, for example), with σ_ϵ the intrinsic ellipticity dispersion of the galaxies, A the area of the survey, $\bar{n}$ the effective mean number density of galaxies and $\Delta\theta$ the width of the angular

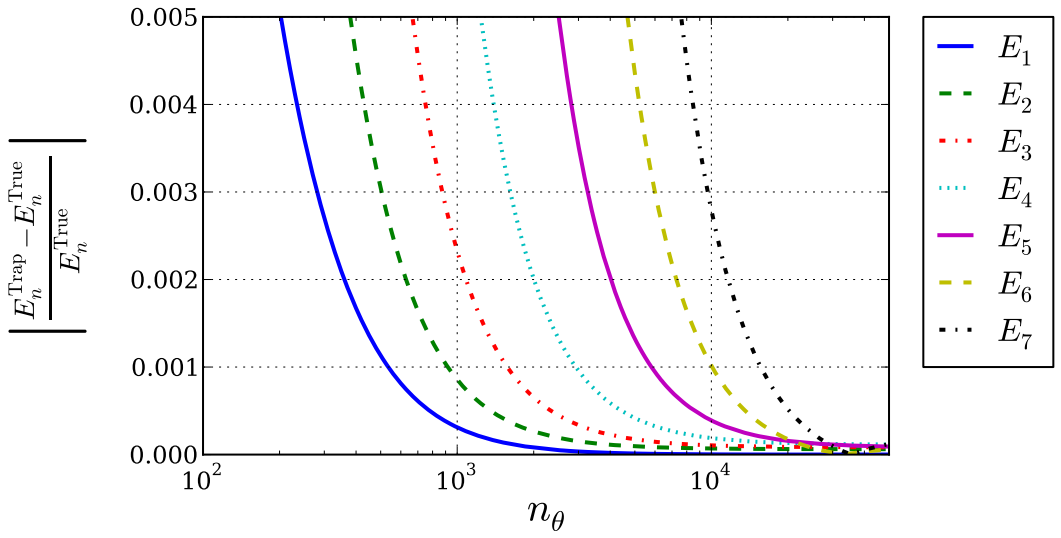

Figure 4.2: The convergence of trapezoidal integration to the true value. As n increases due to more oscillations in T_{+n} a higher number of bins is required to reach the same level of accuracy. The angular bins are linearly spaced. The true value of E_n is calculated using Eq. (4.3), while the trapezoidal (Trap) value is calculated from Eq. (4.4).

bins. R is a randomly generated number from a Gaussian distribution with a variance of 1 and a mean of 0. With the above definitions the covariance of $\xi_\pm(\theta)^{\rm Noisy}$ is equal to the desired covariance. Fig. 4.3 shows the ensemble average estimate of E_n from 50 $\xi_\pm(\theta)^{\rm Noisy}$ realizations with respect to the number of angular bins. The errors shown are the standard deviation of the mean value of E_n over all the realizations. The presence of the random errors do not change the conclusions drawn from the previous test (Fig. 4.2). By comparing Fig. 4.3 and Fig. 4.2 we can also conclude that the random noise due to the intrinsic ellipticity dispersion of galaxies does not hinder the estimation of COSEBIs.

4.3 Clone Simulations

So far we have seen analytical tests of COSEBIs. In this section I test the methods on the numerical mock data from the Clone simulations which resemble the CFHTLenS data. For this thesis I use the first version of the Clone catalogues which are based on the TCS simulation suite (see Harnois-Déraps et al. 2012, 2013, for details) and consist of 184 independent lines of sight, 12.84 square degrees each. The Clone catalogues are specifically made for the CFHTLenS data, taking into account its redshift distribution and masks. However, the positioning of the galaxies in the fields is random, hence clustering effects are ignored (see Heymans et al. 2012, for more details). The limited box size of the simulations dictates limitations to the maximum scale that can be trusted. Hence, I choose to only go up to 40′

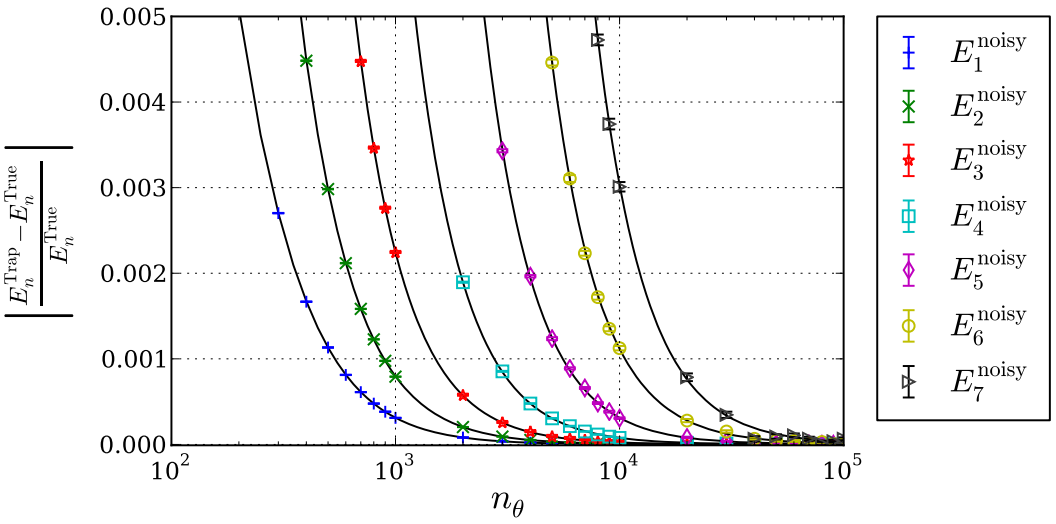

Figure 4.3: The effect of noise on the estimated E_n for the first 7 modes. Here we only consider uncorrelated noise between the angular bins. n_θ is the number of angular bins for the 2PCFs. The black solid lines are the curves in Fig. 4.2, over plotted here for comparison. The error bars are calculated from the variance between the noise realizations.

in this test. We will see in the following sections that the signal is very small for $[20', 40']$. Heymans et al. (2013) summarize the $\sigma_8(\Omega_{\rm m}/0.27)^\alpha$ constraints for several cases (see Table 2 in their paper). This parametrisation characterizes the degeneracy between σ_8 and Ω_m. They find that using larger angular scales increases the constraints on this parameter. However, in all of the cases considered by Heymans et al. (2013), other cosmological parameters are not fixed to their fiducial values and are marginalized over. Consequently, higher angular ranges add more information for those cases, while they do not have a significant impact in the analysis presented here.

The resolution of the simulations puts limits on the small scales. Combining this with the fact that on small scales baryonic (not included in the simulations) and clustering effects become important, I choose a lower angular limit of $1'$ (see Semboloni et al. 2011, for the effects different baryon feedback models have on structure formation and $\xi_\pm$). For the true data analysis presented in Sect. 4.4 the covariance matrix from the Clone simulations, is needed. Ergo, the same angular range is chosen in that analysis as well. Another advantage of using a $\theta \in [1', 40']$ is that a flat sky approximation provides a satisfactory result.

The redshift distribution of the CFHTLenS data is measured using photometric redshift estimates as explained in Hildebrandt et al. (2012). In the Clone catalogues the best fit value of each galaxy's redshift, $z_{\rm B}$, is given, which is in general different from its spectroscopic value. To mimic the real data we use $z_{\rm B}$ to choose which redshift bin a galaxy belongs to, while using the underlying redshift distribution of all the galaxies in each tomographic bin to

find the theory values for our observables. Fig. 4.4 shows the overall redshift distribution of galaxies with $z_B \in [0.58, 1.3]$ and for the four redshift bins in Table. 4.1. The tomographic bins have no overlap in z_B, however, as is clear from the figure this is not the case for the underlying true redshift distribution. In the analysis I will ignore the error on photometric redshift uncertainties, similar to other CFHTLenS studies.

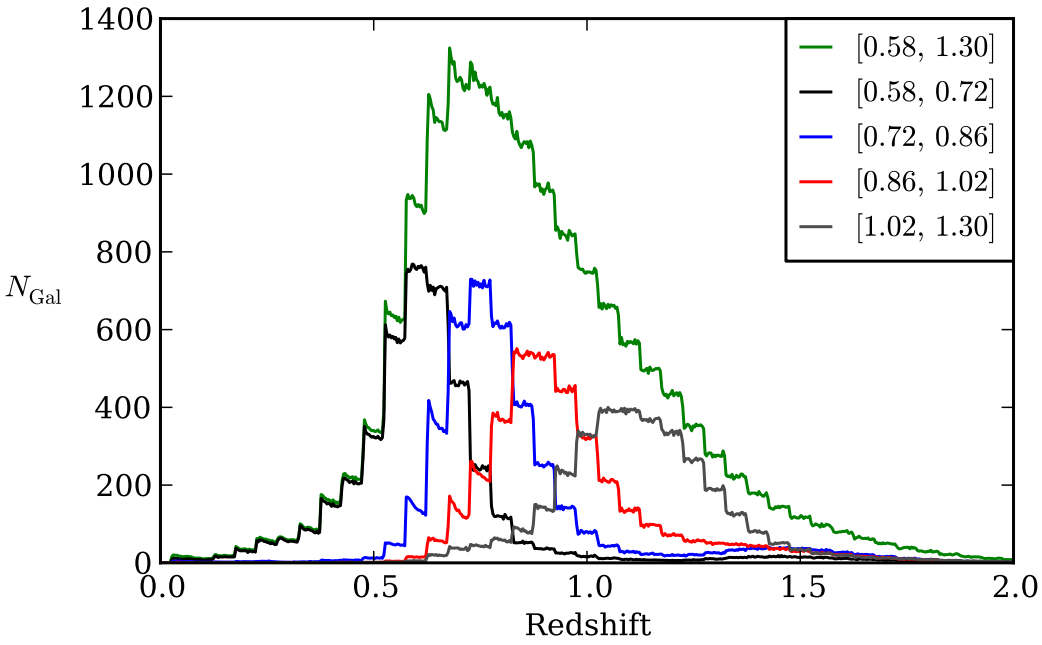

Figure 4.4: Redshift distribution for the galaxies in redshift range $z_B \in [0.58, 1.3]$. The curves show the spectroscopic redshift of the galaxies, with a z_B that belongs to the redshift range given in the plot legends. The y-axis shows the average weighted number of galaxies in a simulated field. The galaxies in the Clone simulations have the same distribution as in the CFHTLenS data. The step like shape of this plot is due to the discrete redshift slices in the simulations.

The shear given in the Clone catalogues has no intrinsic ellipticity noise associated with it. Hence to explore the effects of galaxy shape noise, it has to be added manually to the shear in the catalogues. The intrinsic ellipticity noise can be generated by a random number picked from a Gaussian distribution, with mean 0 and variance σ_ϵ^2. The σ_ϵ reported for the galaxies in the CFHTLenS fields is 0.279 per ellipticity component.

4.3.1 Measuring COSEBIs and their Covariance

In Sect. 2.6 we saw the formalism for COSEBIs. In this Chapter we saw Eq. (4.4) which shows the relation between E_n and B_n modes with the 2PCFs. The estimated 2PCFs, from the input

ellipticities and the weights, w, associated to them, for redshift bins i and j, are

$$\hat{\xi}^{ij}_{\pm}(\vartheta) = \frac{\sum w_a w_b \left[\epsilon^i_{\rm t}(x_a)\epsilon^i_{\rm t}(x_b) \pm \epsilon^i_{\rm x}(x_a)\epsilon^i_{\rm x}(x_b)\right]}{\sum w_a w_b} \,, \tag{4.9}$$

where $\epsilon_{\rm t/x}(x_a)$ are the tangential/cross ellipticities at position x_a, with respect to the reference frame connecting the pairs of galaxies involved. The sum is taken over all pairs of galaxies with separation within a bin around ϑ.

To estimate the 2PCFs, I use Athena[2] a tree code that calculates second-order correlation functions from input galaxy catalogues. The algorithm for the tree code in this case divides the area of a field of galaxies into cells of different sizes. If two galaxies are in two different cells with a large distance then they are assumed to be in the center of the cell. This method enables a faster 2PCFs calculation, at the loss of accuracy. However, the algorithm accepts an input parameter called the opening angle which controls the cell size. The smaller the opening angle the more accurate and slower the calculations. Here I use an opening angle of 0.02 radians. I have also compared the results with a brute force calculation (opening angle = 0) and found no significant differences.

As explained above the only angular ranges considered belong to $\theta \in [1', 40']$. However, I divide this range into two equal parts, hence compare the results for $[1', 40']$ with $[1', 20']$ and $[20', 40']$. Dividing the full angular range in this way allows for diagnosing measurement errors, which can affect these angular ranges in different ways. Consistency between the results for different angular ranges, is a great way to assure that analysis is reliable. In addition, for each angular distance range a single redshift distribution is compared to the tomographic case with 4 redshift bins will be considered.

The covariance matrix of the COSEBIs is measured by finding the covariance between the estimated COSEBIs for each Clone field. However, the estimated covariance from a finite number of simulations is noisy which causes biases in the inverse covariance (see Hartlap et al. 2007). Assuming Gaussian errors on the estimated covariance matrix, $\hat{\mathbf{C}}$, the inverse covariance matrix is

$$\mathbf{C}^{-1} = \frac{n_{\rm sim} - n_{\rm obs} - 2}{n_{\rm sim} - 1}\hat{\mathbf{C}}^{-1} \tag{4.10}$$

where $n_{\rm sim}$ and $n_{\rm obs}$ are the number of simulations and observables respectively. For a $n_{\rm obs}/n_{\rm sim} < 0.8$, the above formula produces an unbiased inverse covariance according to Hartlap et al. (2007). Although, this correction guarantees that the resulting $\mathbf{C}^{-1}$ is unbiased, it will still have noise associated with it, which depends on the ratio of the number of observables to the number of simulations. In this section I will use 7 COSEBIs for each redshift bin, which means that for the tomographic case with 4 redshift bins and 10 redshift bin combinations, there are 70 COSEBIs modes in total. As a result the ratio $n_{\rm obs}/n_{\rm sim} = 0.38$, which can cause about

[2]http://www.cosmostat.org/software/athena/

7% errors in the estimated inverse covariance. This value for the error on the covariance matrix is acceptable for analysing CFHTLenS data. However, for precision cosmology, these errors are not acceptable. Hence, in the following section I will redo the analysis with compressed COSEBIs, which have a smaller number of observables. Taylor & Joachimi (2014) showed that the correction in Eq. (4.10) will still result in a biased parameter covariance matrix. They provide this correction for the parameter covariance matrix,

$$\mathbf{C}_{\rm par} = \frac{n_{\rm sim} - n_{\rm obs} - 2}{n_{\rm sim} - n_{\rm obs} + n_{\rm par} - 1}\hat{\mathbf{C}}_{\rm par} \,, \tag{4.11}$$

where $\mathbf{C}_{\rm par}$ is the parameter covariance matrix and $n_{\rm par}$ is the number of parameters to be estimated. Applying this correction to $\mathbf{C}_{\rm par}$ results in a slightly smaller covariance matrix. This estimate is unbiased, however, there is still noise associated with it. This correction is small compared to the inaccuracies in the estimations for the covariances in this work, hence, it will be neglected.

The COSEBIs covariance for a single redshift distribution from the simulations is shown in Fig. 4.5, for the three angular ranges discussed above. The right panels show the covariance for E_n while the left panels show the same for B_n. In Sect. 2.6.2 we saw the theoretical prediction of the covariance assuming that the shear follows a Gaussian distribution. The covariances we see in Fig. 4.5 show a similar band matrix form, however non-Gaussian effects are important here, since the matter power spectrum for these scales becomes non-linear which causes mode mixing. This effect differentiates between the theory and simulations.

Fig. 4.6 shows the covariance for the four redshift bins in Table. 4.1. Each covariance matrix has 7^2 elements for each combination of redshift bin pairs. The x/y-axis in the plots show the redshift bin pairs considered. In total the covariance has 70^2 elements. The left panels show the covariance for E_n while the right ones are for B_n.

We see for all cases that, the value of the covariance drops for the off diagonal elements, but that, for both redshift binning cases the $[20', 40']$ range has a more diagonal covariance. This is due to the fact that non-Gaussian effects are less important for this angular range. Therefore, any analysis which only uses this angular range is less likely to be biased because of poor modelling of non-linear scales. However, as we will see later on in this chapter the cosmic shear information in this angular range is significantly lower than the lower angular scales.

The theoretical Gaussian covariance used in the previous sections is significantly different from the covariances estimated from the simulations due to non-linear growth of structure- as expected. As a result, I repeat the test for the number of θ bins (described in Sect. 4.2) required for the estimated 2PCFs to provide an accurate E_n. Repeating this test with the simulated data has the advantage that we can study the possible effects of correlated noise in $\xi_\pm$, ignored in the shot-noise only test. Fig. 4.7 shows the value of E_n measured from the Clone simulations versus the number of bins in $\theta \in [1', 40']$, for a single redshift distribution. Here the binning is

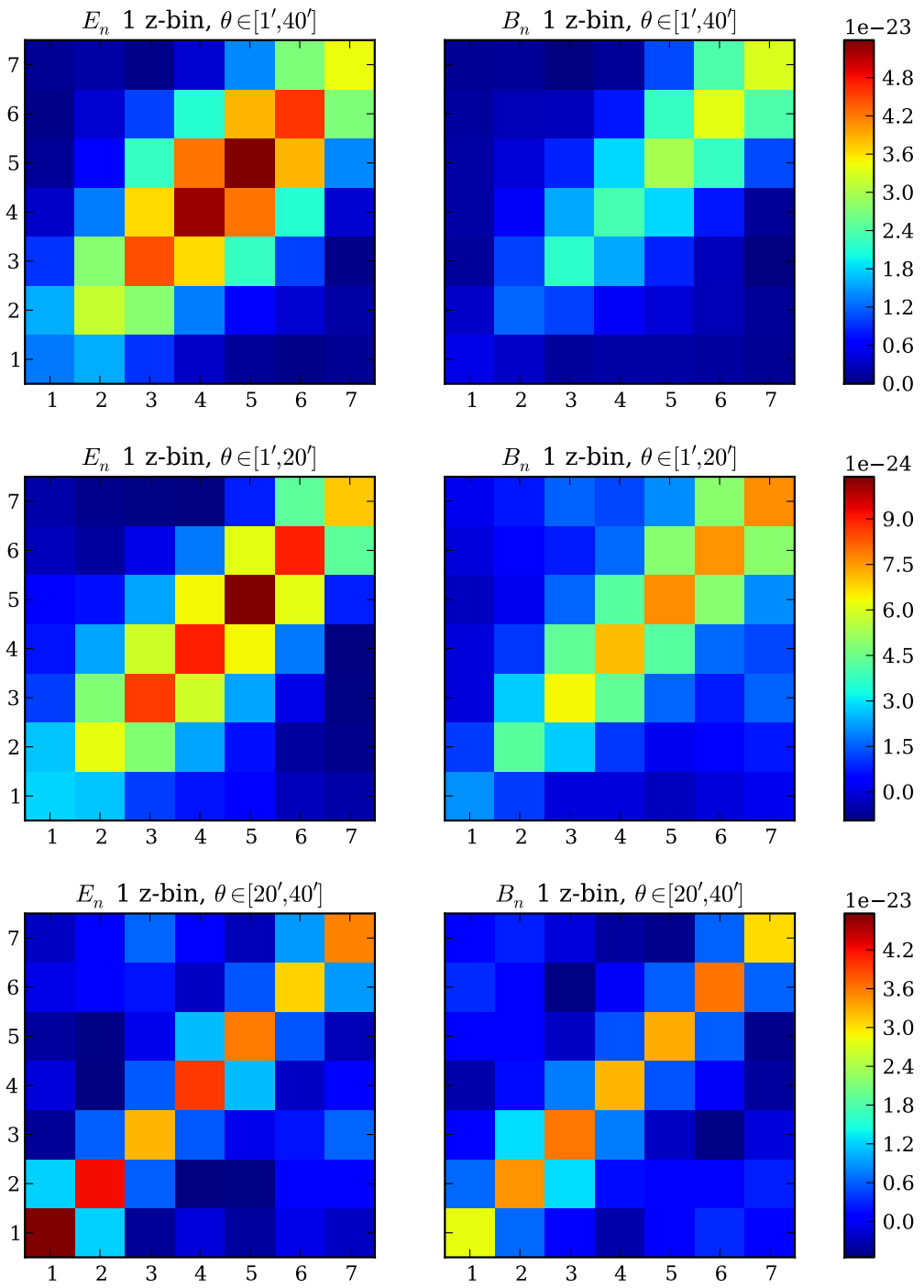

Figure 4.5: Covariance matrices of COSEBIs for a single redshift distribution. Three angular ranges, $[1', 40']$, $[1', 20']$ and $[20', 40']$ are considered here. The x/y-axis show the COSEBIs mode considered. The COSEBIs are less correlated for $[20', 40']$ compared to the other cases.

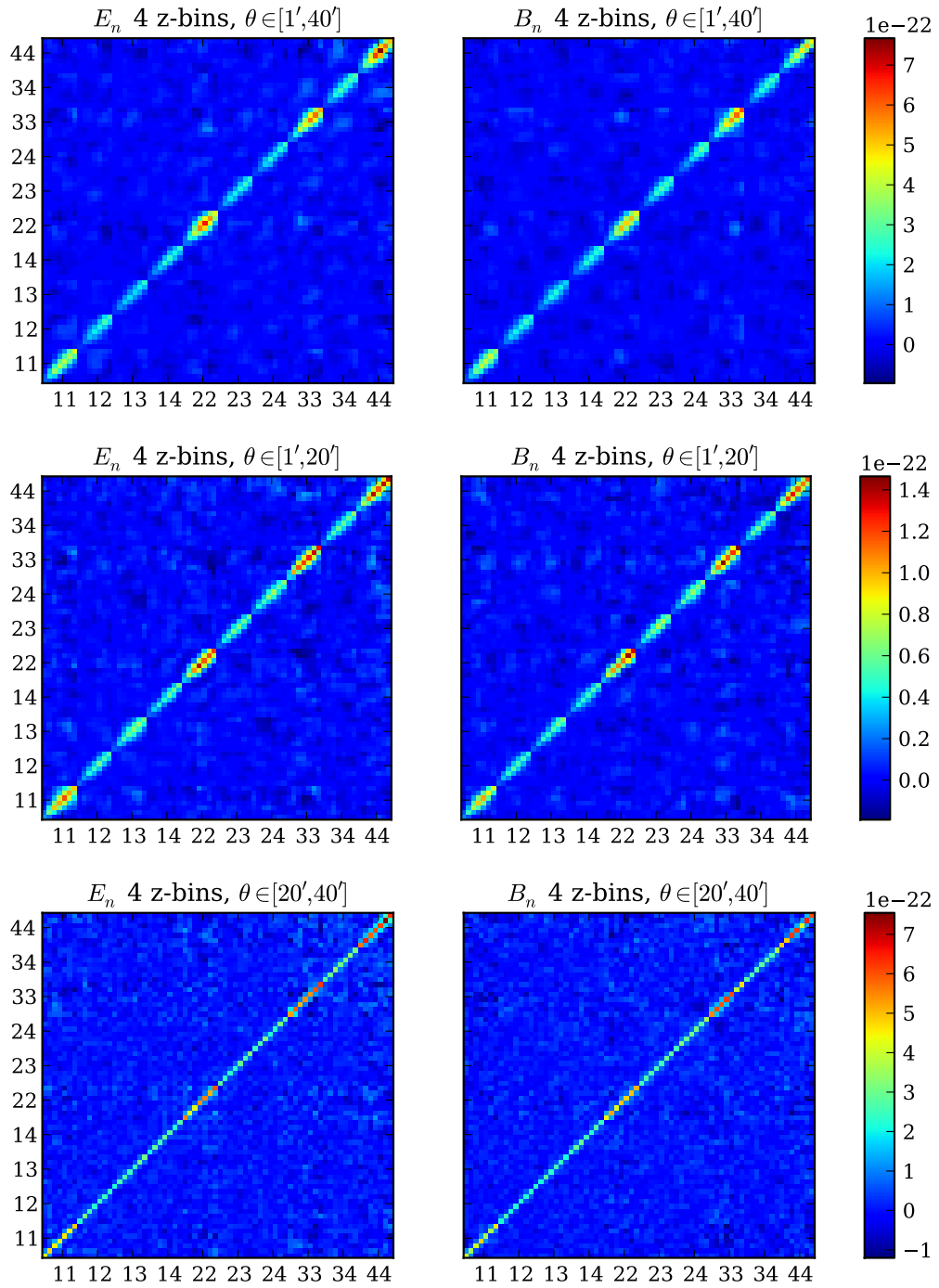

Figure 4.6: Covariance matrices of COSEBIs for 4 redshift bins. Three angular ranges, $[1', 40']$, $[1', 20']$ and $[20', 40']$ are considered here. The x/y-axis show the redshift bin combination, for example '13' means redshift bins 1 and 3 are relevant. There are 7 COSEBIs modes for each combination.

linear which is more convenient for the integral in Eq. (4.4). With a narrower angular binning scheme the number of galaxies in each bin decreases. Hence the $\xi_{\pm}$ estimate is noisier. In this analysis I have made sure that all the bins are populated with galaxies. As we can see from the plot, the error on the estimated E_n is much larger in this case. The differences between the binning schemes is smaller than the statistical errors. Nevertheless, for the rest of the analysis I choose to use 1.6×10^5 angular bins with linear spacing between $[1', 40']$ which is larger than the previous value. Note that adding more bins cannot reduce the accuracy of the estimated COSEBIs, as long as there is at least a pair of galaxies which reside in that bin.

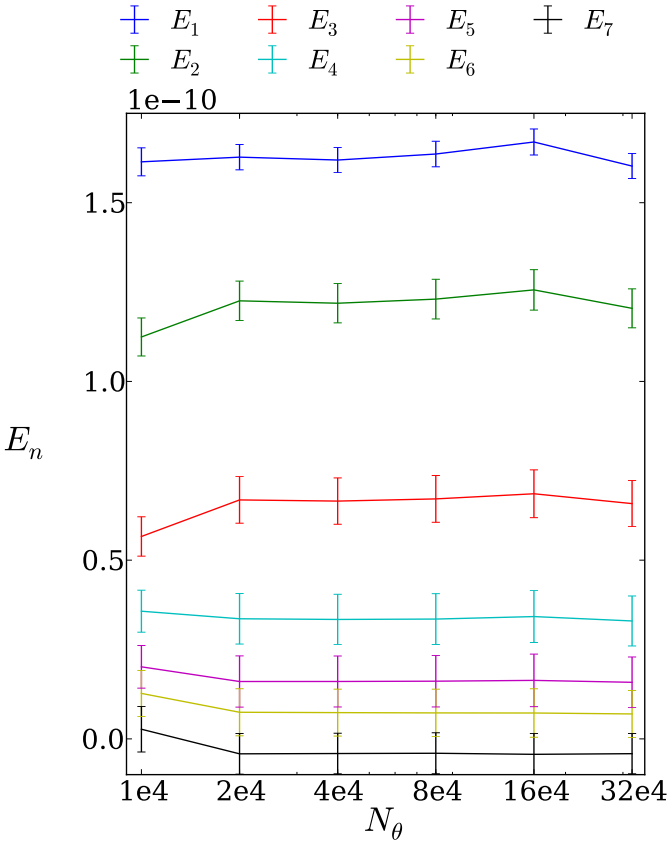

Figure 4.7: E_n versus the number of angular bins, N_θ, for Clone simulations. The angular range here is $[1', 40']$. The error bars come from the covariance measured from the simulations. We can see here that using more than about 10^5 angular bins does not contribute to a better convergence to the true value of E_n, since the fluctuations in the measurements are within the statistical errors.

The covariance matrices shown in Figs. 4.5 and 4.6 are measured from the full area of the Clone simulations which is $12.84 \times 184 = 2362.56$ square degrees. For the purposes of this thesis, I perform the analysis for the CFHTLenS area, to test the accuracy of the method and the pipeline for the statistical noise of the real data. As a result the area of the Clone has to be scaled by $184 \times 0.9 \times 16/129 = 20.54$. This scaling factor comes from the fact that each of the lines of sight of the 184 Clone simulations cover 0.9×16 pointings of the CFHT camera (Megacam) and a total of 129 Megacam pointings are used here. Furthermore, to have the same statistical fluctuations as for the CFHTLenS data only $129/(0.9 \times 16) \approx 9$ Clone fields are going to be considered for estimating the mean value of the COSEBIs.

σ_8	Ω_m	Ω_Λ	w_0	n_s	h	Ω_b
0.817	0.279	0.721	−1.0	0.96	0.701	0.046

Table 4.4: The value of the cosmological parameters for the Clone simulations (Harnois-Déraps et al. 2012).

In Sect. 2.4.1 we saw the importance of E/B-mode decomposition and that the presence of B-modes can signal remaining systematic errors. Therefore, let us first look at the measured B_n from the Clone fields. Fig. 4.8 shows the B-mode COSEBIs for the single redshift distribution case. The goodness-of-fit of the B_n to zero is estimated via a reduced χ^2, which is equal to a χ^2 normalized by the number of free parameters. In this case we have 7 free parameters (the 7 B_n modes). For all the cases the B-mode contribution is consistent with zero.

Fig. 4.9 shows the value of the B_n for the 4 redshift bins in Table. 4.1 and an angular range of $[1', 40']$. The B-mode is also negligible in this case and for the other two angular ranges, which are not shown. This is a test for the method and the simulations. The B-modes being consistent with zero does not guarantee a correct E-mode estimate, however, if they were non-zero it would indicate an error in the simulations or the method. The value of the χ^2 in some cases is rather small. This can be due to the fact that the covariance of the COSEBIs is estimated using all of the 184 simulations, while the B_n are measured using only 9 simulations. This result can change if another 9 simulations are used, which will be investigated in future work.

The E_n COSEBIs measured from the Clone simulations are shown in Fig. 4.10. In this figure a single redshift distribution with the three angular ranges discussed above are considered. The solid curve shows the value of E_n expected from theory with the underlying cosmological parameters of the Clone as listed in Table. 4.4. The COSEBIs modes are only defined for natural values of n, hence connecting them with a line has no meaning. Nevertheless, I do so to provide a visually easier comparison between the theory and data. The goodness-of-fit values given in the plots, show the reduced χ^2 values for the 7 E_n modes considered. Note that this is not a fit to the data, as a result the degrees of freedom is equal to the number of observables. From Fig. 4.10 we can see that for the area of CFHTLenS the theory values provide a good match to the simulated ones.

Figs. 4.11, 4.12 and 4.13 show the values of E_n for angular ranges $[1', 40']$, $[1', 20']$ and $[20', 40']$ and 4 redshift bins. The goodness-of-fit of the theory model to the estimated E_n is shown in the legends in the form of a reduced χ^2 for E_n. The overall χ^2 values show a good match between the theory and the simulated E_n. However, the χ^2 value for the autocorrelation of the first redshift bin and angular ranges $[1', 20']$ and $[1', 40']$ is slightly high. This mismatch is not seen for the higher angular range, which suggests that for the angular range $[1', 20']$ the first redshift bin prefers a slightly different cosmological model. The reason behind this behavior can be that the 9 Clone fields chosen for this test have a statistically significant

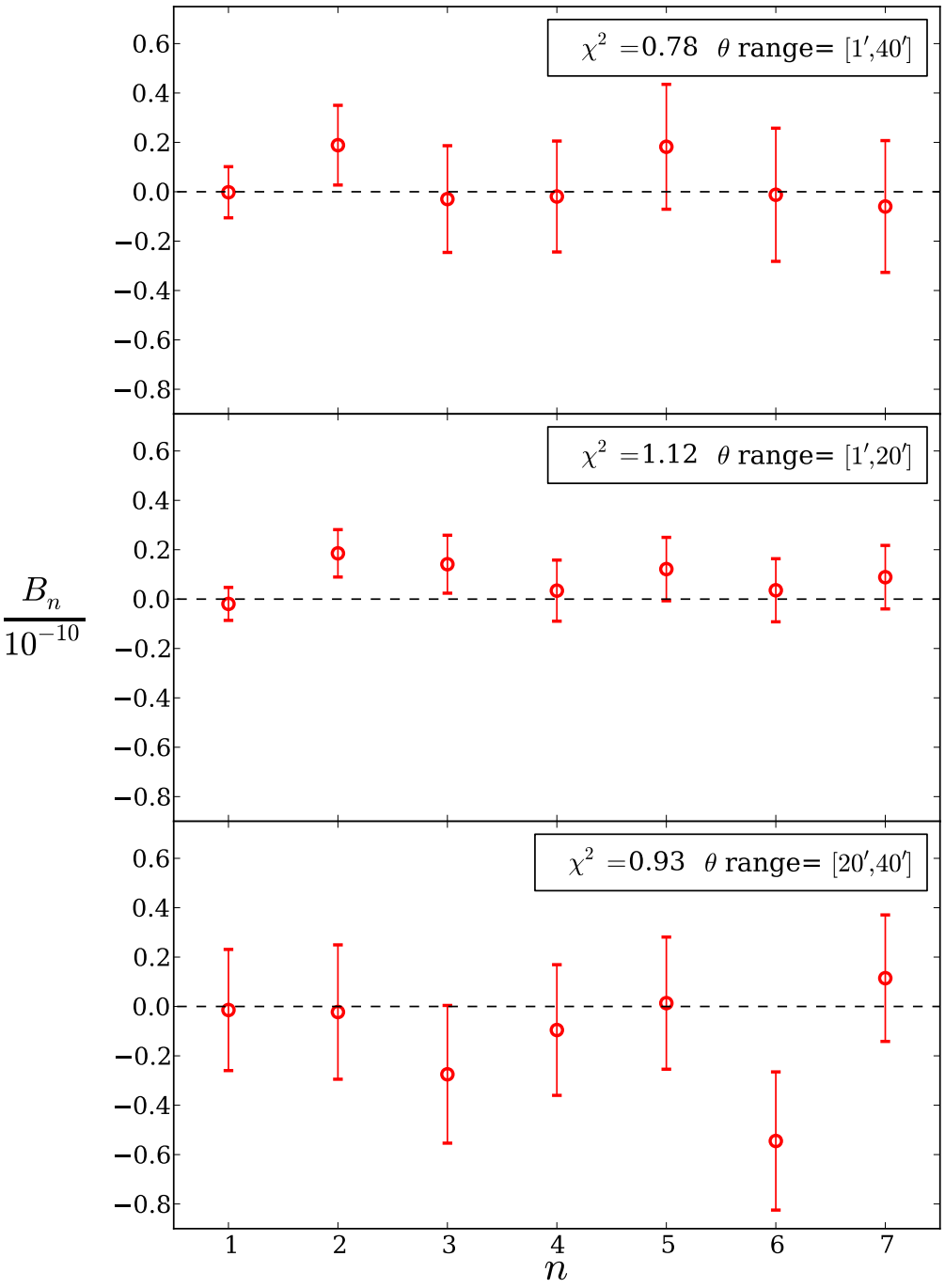

Figure 4.8: Measured B_n from the Clone simulations for a single redshift bin. Three angular ranges are considered here. For each plot the reduced χ^2 for the B-modes and the angular range is stated, showing that the simulated B-mode is consistent with zero for all the cases tested here.

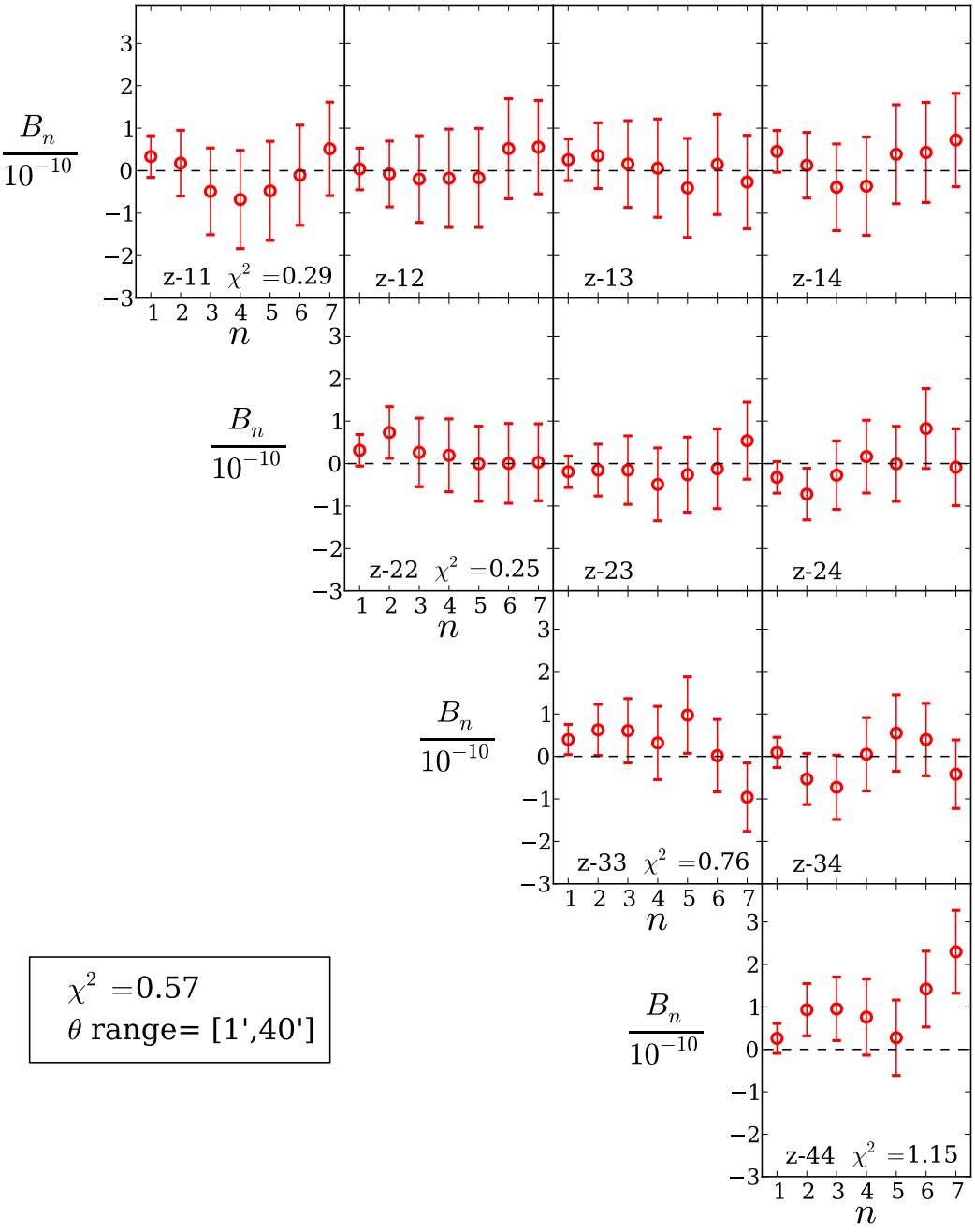

Figure 4.9: Measured B_n from Clone simulations for 4 redshift bins. An angular range of $[1', 40']$ is considered here. The reduced χ^2 for the B-modes is stated in the legends, using the full data set with 70 degrees of freedom. Each plot shows the B_n for the combination of redshift bins written on its corner. The reduced χ^2 is also stated for the autocorrelation cases, where only the covariance of the B_n in the respective redshift bin is considered.

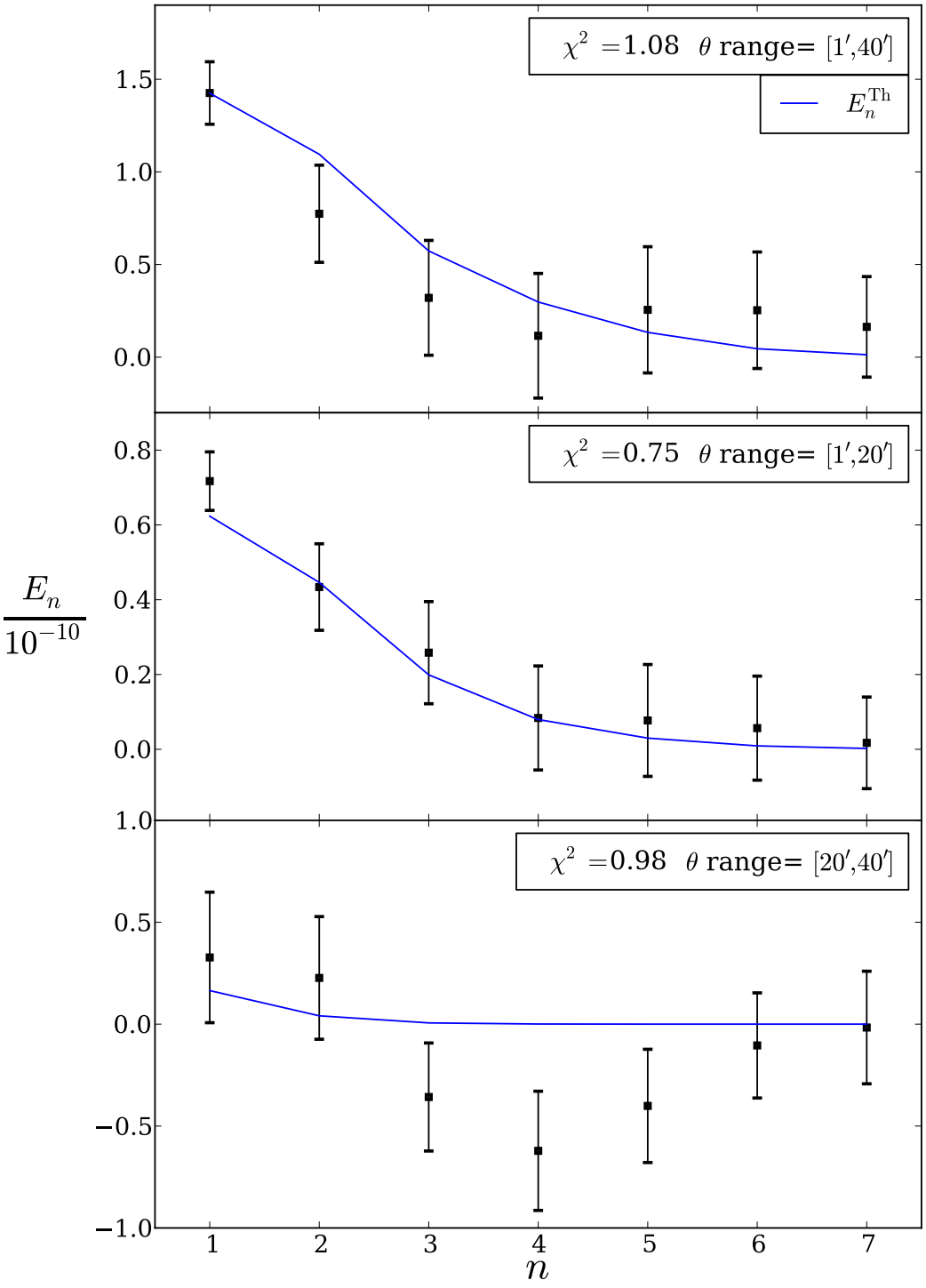

Figure 4.10: E-mode COSEBIs measured from Clone simulations. A single redshift distribution is assumed here. The error bars correspond to the field to field variation of the simulations. The solid lines show the theory values of the E_n for the underlying parameters of the simulations, given in Table. 4.4. Note that the COSEBIs modes are discrete and the theory lines are drawn only for convenience. The reduced χ^2 correspond to the difference between the theory and the simulated values.

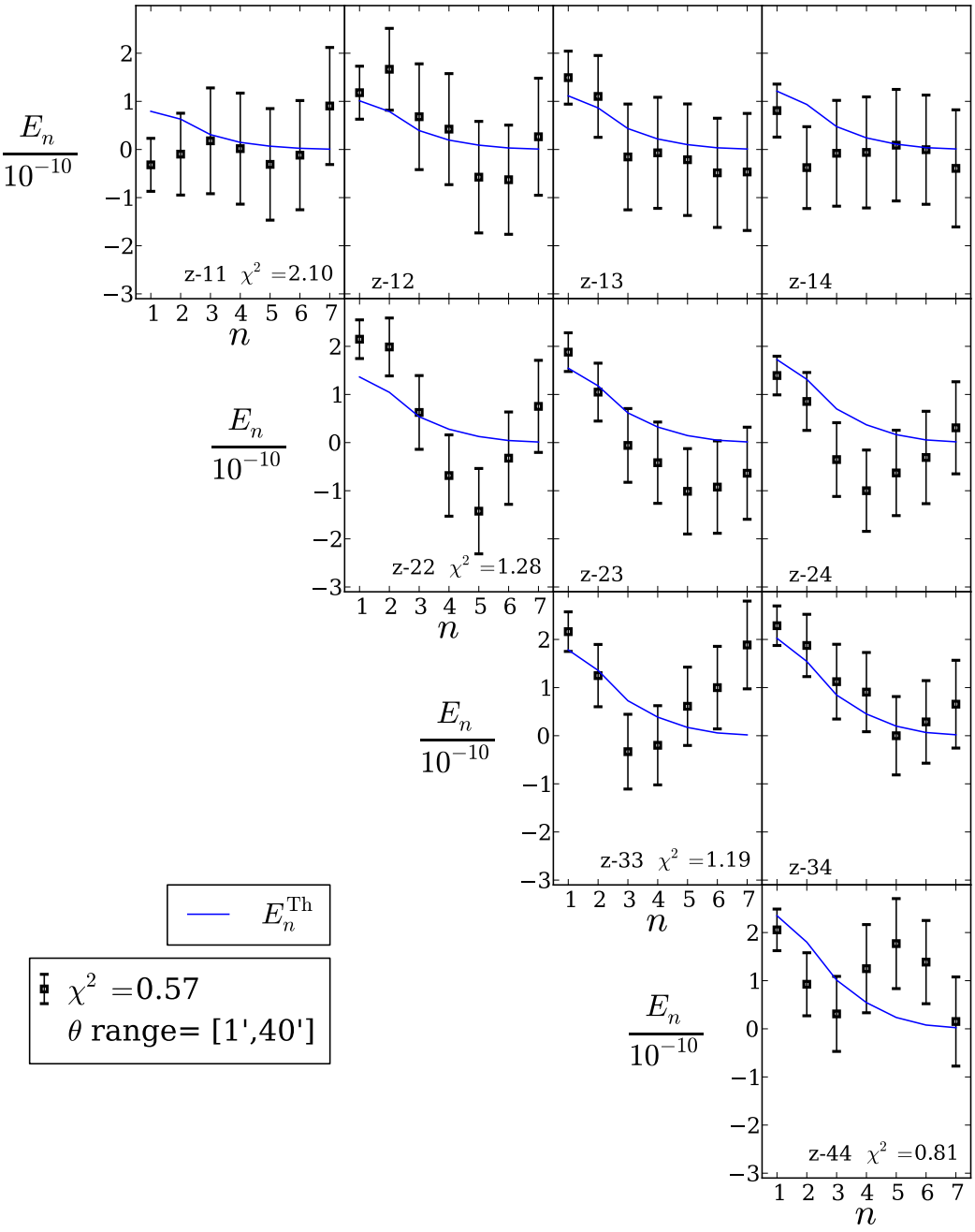

Figure 4.11: E-mode COSEBIs measured from Clone simulations for 4 redshift bins and $[1', 40']$. The error bars correspond to the field to field variation of the simulations. The solid lines show the theory values of the E_n for the underlying parameters of the simulations, given in Table. 4.4. Note that the COSEBIs modes are discrete and the theory lines are drawn only for convenience. The reduced χ^2 shows the correspondence between the theory and the simulations. The value of the reduced χ^2 is also shown for the autocorrelations between the tomographic bins.

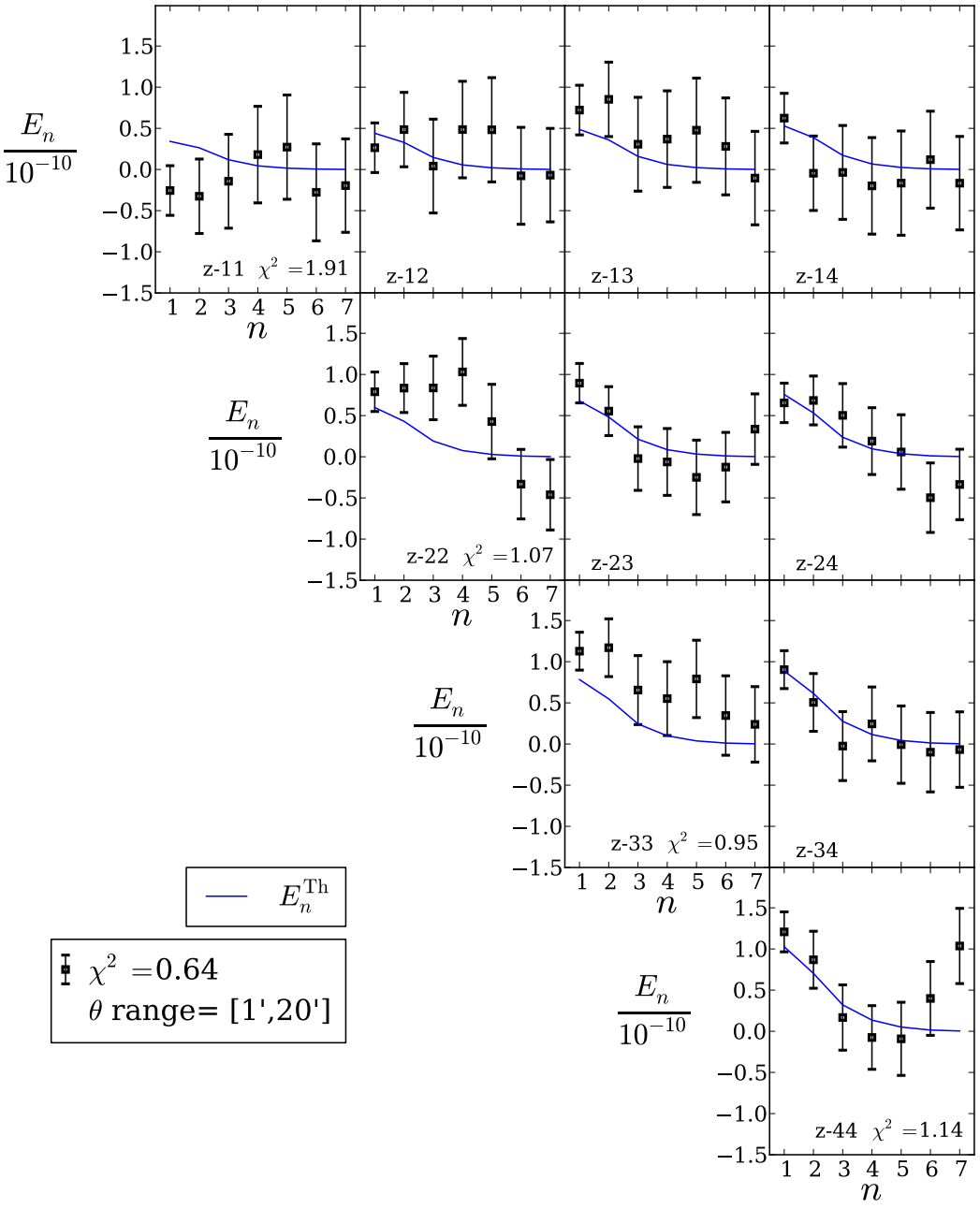

Figure 4.12: E-mode COSEBIs measured from Clone simulations for 4 redshift bins and $[1', 20']$. See the caption for Fig. 4.11 for more details.

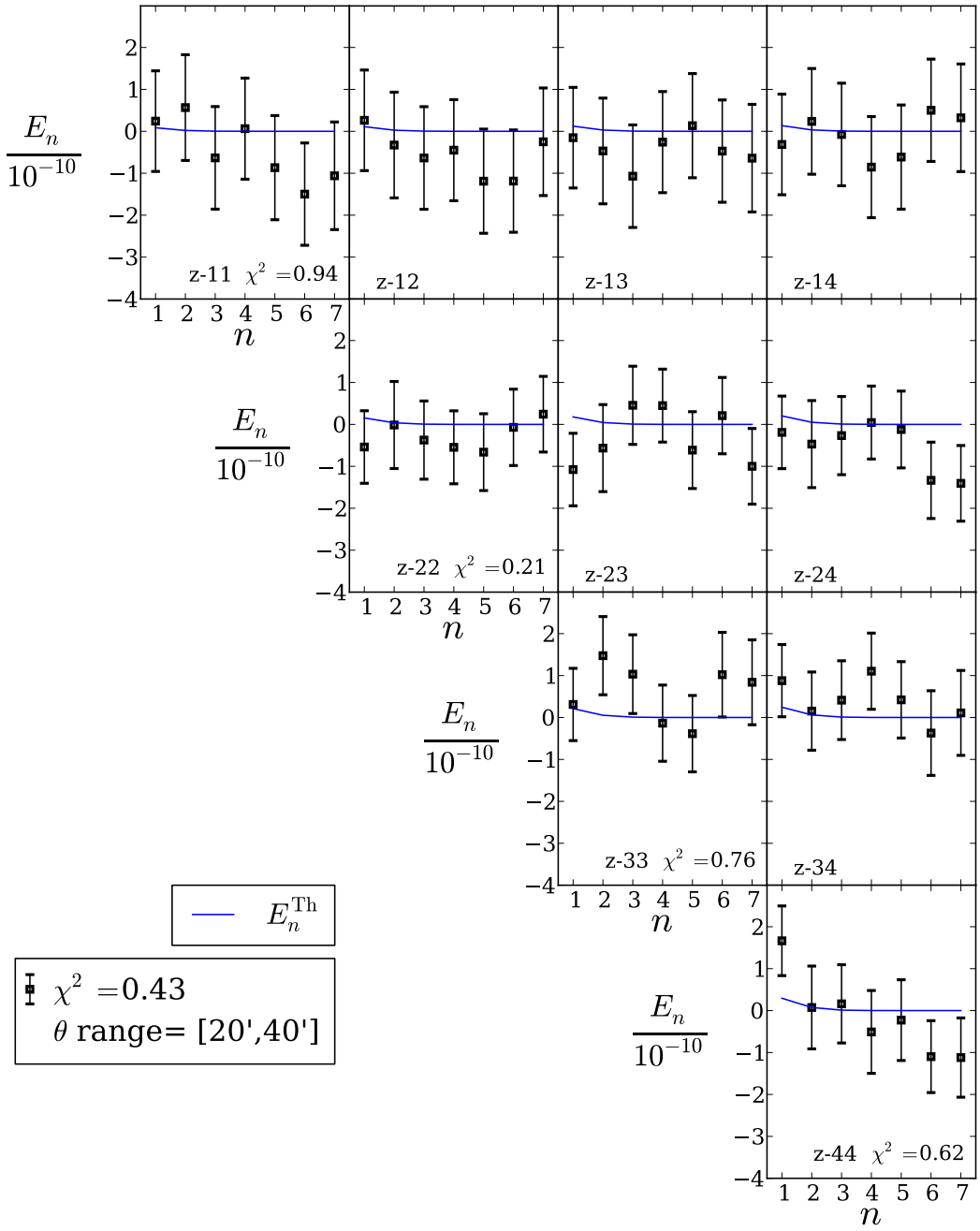

Figure 4.13: E-mode COSEBIs measured from Clone simulations for 4 redshift bins and $[20', 40']$. See the caption for Fig. 4.11 for more details.

difference with other Clone fields for this angular range and redshift bin. In my future work I will test this by using other sets of simulations. The general trend we see in these three figures is that for a higher redshift bin the signal is higher specially for $[1', 20']$ and $[1', 40']$.

4.3.2 Measuring Compressed COSEBIs

In the previous subsection we saw the estimated COSEBIs from the Clone simulations and the predications from the theory. Here I will show similar results for the compressed COSEBIs (CCOSEBIs from here on). In Chapter 3 we went through a method which is able to compress any set of observables according to their sensitivity to the parameters to be estimated. Here I use this method for two free cosmological parameters, σ_8 and Ω_m, and compress the COSEBIs to 2 first order and 3 second order CCOSEBIs. All other parameters will be fixed to the fiducial values in Table. 4.4. Sect. 3.3 gives a detailed account of how the CCOSEBIs are made from the COSEBIs.

To find the CCOSEBIs a covariance matrix has to be assumed. As is shown in Chapter 3 the CCOSEBIs are not very sensitive to the choice of the covariance. Hence we use a theory covariance matrix for the Clone parameters in Table. 4.4 and the redshift distribution in Fig. 4.4, with Gaussian distributed shear fields (see Sect. 2.6.2). Furthermore, the first and second order partial derivatives of the COSEBIs are required for defining the CCOSEBIs. I numerically calculated these using the theory COSEBIs in Eq. (4.3).

Similar to the analysis with COSEBIs we need to find the covariance of the CCOSEBIs first. For both tomography and the single redshift distribution cases only 5 CCOSEBIs are present. In the previous subsection we saw that for each redshift bin 7 COSEBIs were used in the analysis. This means that for the single redshift distribution the compression does not provide a significant reduction to the number of observables (from 7 to 5). However, in the presence of tomography the number of COSEBIs quickly increases. For the 4 redshift bin case, 70 COSEBIs modes are used in the analysis. With the compression method this number is decreased to only 5 for the two free cosmological parameters. This compression guarantees a more accurate covariance estimation from the Clone simulations, as discussed in Sect. 4.3.1. The covariance matrix of the CCOSEBIs is measured from the field to field variance between the measured CCOSEBIs from the Clone simulations.

Fig. 4.14 shows the covariance matrix for the E-mode CCOSEBIs. The right panels show the covariance for a single redshift distribution while the left panels show the same for 4 redshift bins. The angular range considered for each plot is written above it. From the plots we can see that the overall amplitude of the covariance for the 4 redshift bin case is higher and there is significant difference between the covariance for different angular ranges. Nevertheless, the covariance matrices with the same angular range share a similar shape. Comparing Fig. 4.14 with Fig. 4.6 shows the strength of this compression method. Recall that a lower number of observables results in a more accurate estimation of their covariance and a smaller correction

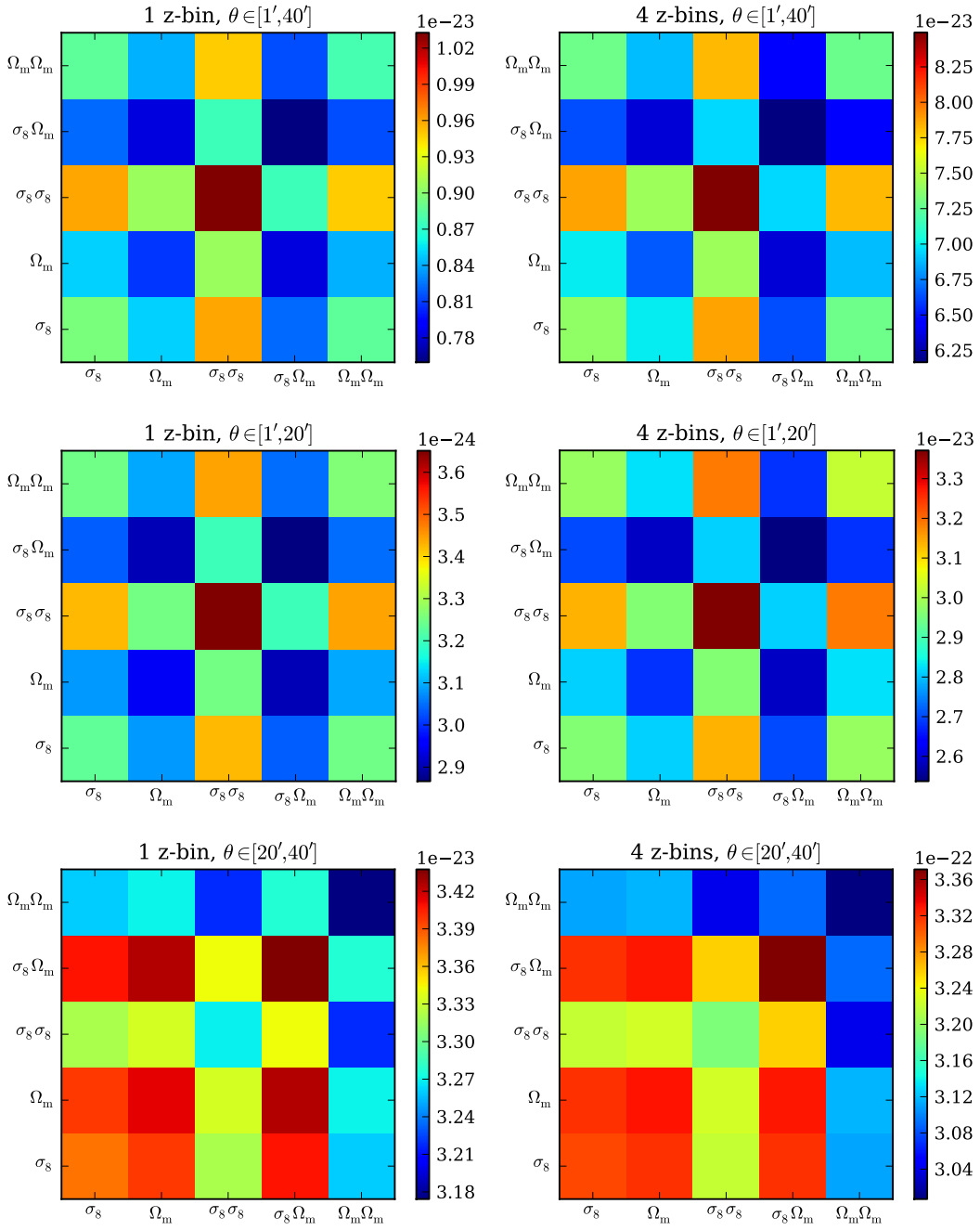

Figure 4.14: The covariance matrix of CCOSEBIs. The elements are measured from the field to field variance of the Clone simulations. Three angular ranges and two redshift binning schemes are considered here. The left panels show the covariance for a single redshift distribution, while the right panels belong to the 4 redshift bin case. The angular range for each plot is shown above it. Comparing the right column of this figure with Fig. 4.6 which shows the COSEBIs covariance for 4 redshift bins, shows how much the compression has reduced the dimensions of the data.

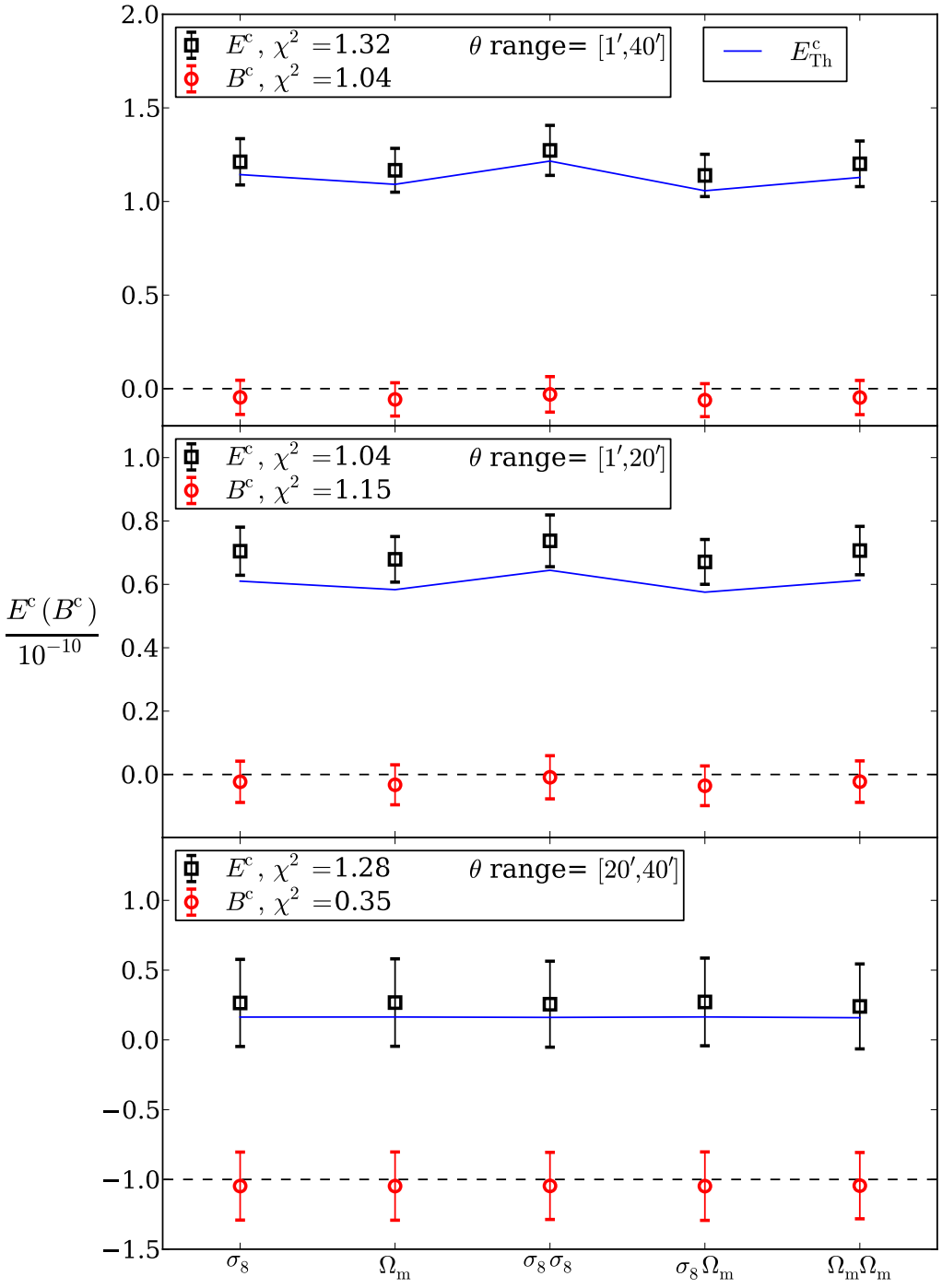

Figure 4.15: CCOSEBIs measured from the Clone simulations. The CCOSEBIs are linear combinations of the COSEBIs. A single redshift distribution is assumed here. The errors bars correspond to the field to field variation of the simulations, scaled to CFHTLenS area. The solid curves show the theory values of the E^c_μ for the true parameters of the simulations, given in Table. 4.4. Note that the CCOSEBIs modes are discrete and the theory lines are drawn only for convenience. The reduced χ^2 values with five degrees of freedom, are shown for the theory values for each angular range. The B^c_μ and their goodness-of-fit to zero are shown as well. The x-axis shows which derivatives are used for making the CCOSEBIs mode. The B^c_μ for $[20', 40']$ are offset from zero by -1 for visual purposes.

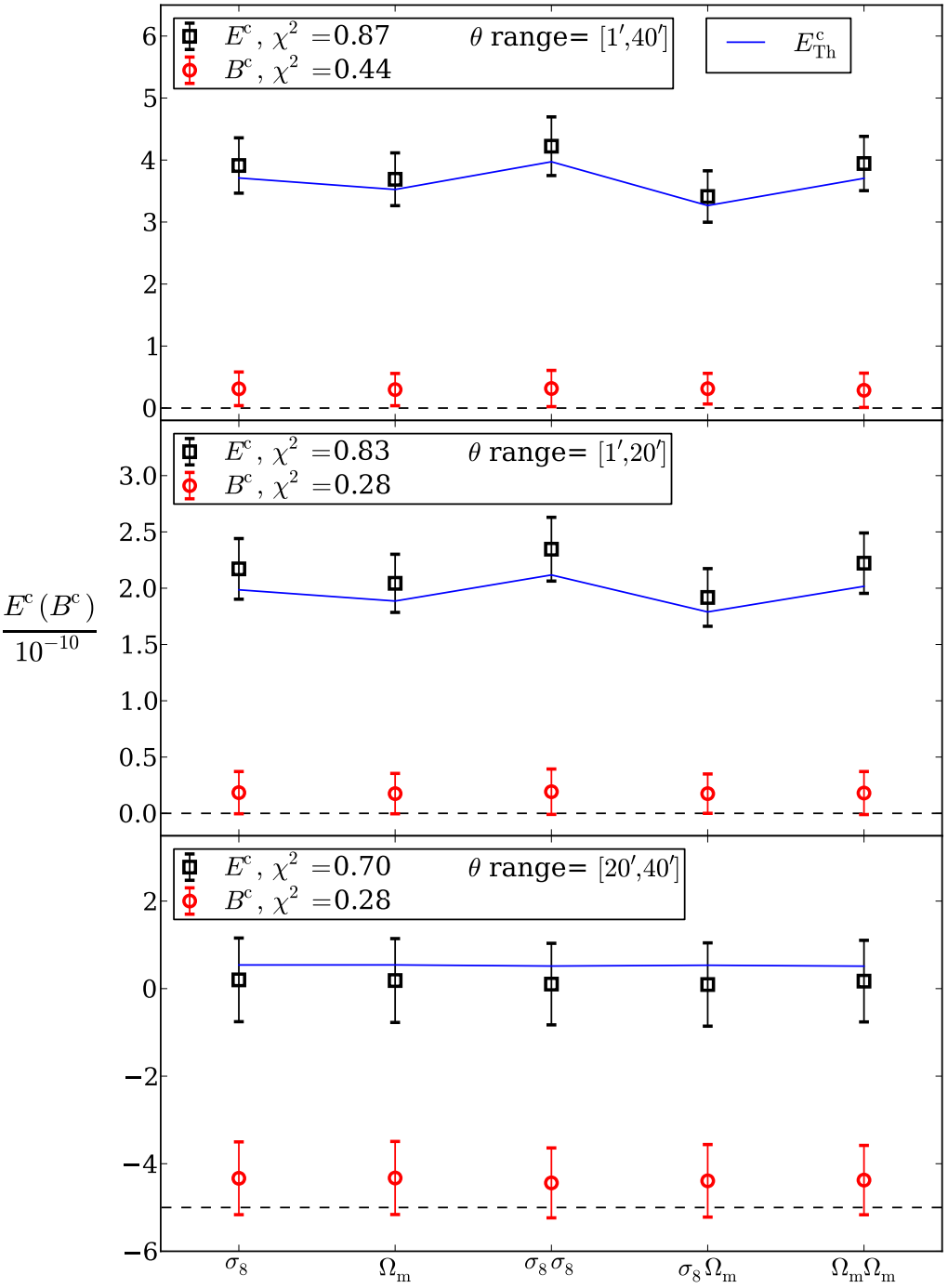

Figure 4.16: CCOSEBIs measured from the Clone simulations for 4 redshift bins. The CCOSEBIs are linear combinations of the COSEBIs. The errors bars correspond to the field to field variation of the simulations, scaled to CFHTLenS area. The curves show the theory values of the E^c_μ for the true parameters of the simulations, given in Table. 4.4. Note that the CCOSEBIs modes are discrete and the theory lines are drawn only for convenience. The reduced χ^2 values are shown for the goodness-of-fit between the theory and the measured E^c_μ for each angular range. The B^c_μ and their goodness-of-fit to zero are shown as well. The x-axis shows which derivatives are used for making the CCOSEBIs mode. The B^c_μ for $[20', 40']$ are offset from zero by -5 for visual purposes.

factor (Eq. 4.10).

To measure the CCOSEBIs from the data I first find the compression matrix which depends on the covariance and derivatives of the COSEBIs (see Sect. 3.3), then apply this compression to the measured COSEBIs from each field. Alternatively, the CCOSEBIs can be measured directly from the shear correlation functions, using their own filter functions, which are shown in Fig. 3.1 for the first order CCOSEBIs. However, since I have already measured the COSEBIs from the data I use the former method for estimating the CCOSEBIs.

Fig. 4.15 shows the resulting E-/B-mode CCOSEBIs (E^c_μ/B^c_μ) from the Clone simulations, for a single redshift distribution. The goodness-to-fit of the E-modes to their theory values and the B-modes to zero are shown in terms of the reduced χ^2. The solid curves show the theory values of the E^c_μ. Note that similar to the COSEBIs the CCOSEBIs only have discrete modes and the lines are only connecting the modes together to aid with visual inspections of the plots. The B-modes are consistent with zero which is not surprising because the CCOSEBIs are linear combinations of the COSEBIs. The x-axis in Fig. 4.15 shows which derivatives are used for defining the CCOSEBIs mode. The first two modes where only one parameter is stated belong to the first order CCOSEBIs, which depend on the first order derivatives of the COSEBIs with respect to a parameter (either σ_8 or Ω_m). The last three cases belong to the second order CCOSEBIs. These modes depend on the second order partial derivatives of the E_n with respect to the parameters written under them.

Fig. 4.16 shows the CCOSEBIs corresponding to the 4 redshift bin case. The curves show the theory CCOSEBIs. Similar to before the reduced χ^2 with 5 degrees of freedom are shown for the E-/B-modes. In general the $[20', 40']$ angular range does not provide much information, and the error bars are significantly larger for it. Consequently, in the next section I will use the angular range $[1', 40']$ as the default case and compare the other cases with it.

4.3.3 MCMC

In the previous two sections we saw the estimated COSEBIs and CCOSEBIs from the Clone simulations and their comparison to their theory values for the true cosmological parameters of the simulations. In this section I will use a Monte Carlo Markov Chain (MCMC) likelihood analysis to determine the best fit cosmology to the measurements. For the purposes of this thesis I will only allow for two free parameters, σ_8 and Ω_m, fixing all the other parameters. The fiducial values of the other cosmological parameters and the constant prior on σ_8 and Ω_m are shown in Table. 4.2. The details of the MCMC algorithm are explained in Sect. 2.1.3. The proposal distribution used to select values in the chain, is based on the Fisher matrix of σ_8 and Ω_m. As we have seen in Sect. 2.1.2 the inverse of the Fisher matrix is equal to the expected covariance of the parameters. Consequently, around the best fit value of the parameters, where the distribution is closer to a Gaussian, a multivariate Gaussian distribution with the inverse of the Fisher matrix as its covariance, provides a relatively good fit to the likelihood distribution.

For a survey with the CFHTLenS area, the confidence regions are relatively large and deviate from a normal distribution. As a result the tails of the likelihood distribution will deviate from the Gaussian form, which leads to a slower convergence of the MCMC chains. Nevertheless, when the overall shape of the likelihood is unknown the aforementioned proposal distribution is the best candidate.

Since I know where the true parameters lie in my analysis of the Clone, I can choose a starting point for the chains which is already in the high probability region in the parameter space. As a result the burn-in chain is very short. For each case that will be discussed in this section at least about 20,000 points exist in the chains and about 2,000 points are subtracted from the start of the chains as the burn-in. The convergence of the chains are tested by dividing each chain into a number of parts with equal number of points and taking the variance between them and comparing that to the total variance of the chain (see Gelmann & Rubin 1992, for example). The tests for all the cases show an acceptable convergence.

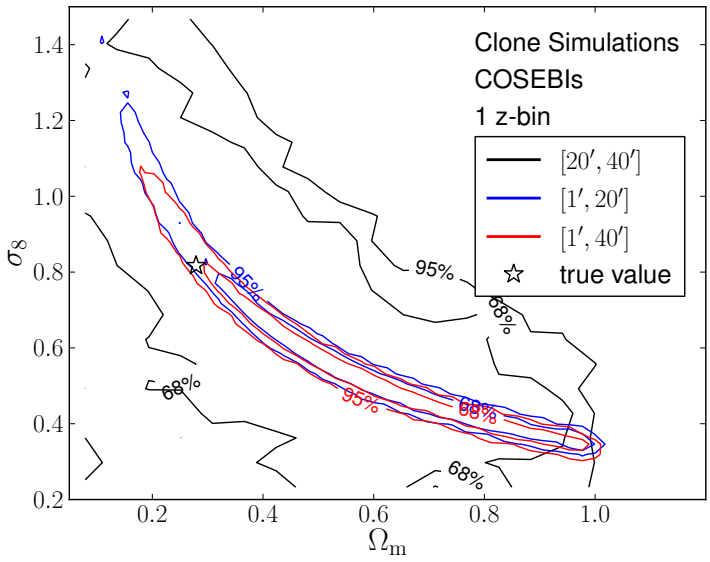

Figure 4.17: Likelihood contours for σ_8 and Ω_m measured from the Clone simulations. A single redshift distribution and 7 E-mode COSEBIs are assumed here. An MCMC algorithm is used for sample the likelihood. The 68% and 95% confidence regions are shown here. The true underlying parameters are shown with the star.

We will first compare the likelihood for the three angular ranges $[1', 20']$, $[1', 40']$ and $[20', 40']$ using the first 7 E-COSEBIs, E_n, for a single redshift distribution. Fig. 4.17 shows the 68 and 95 percentile likelihood contours for this case. The true underlying value of the parameters σ_8 and Ω_m are shown with a star in the plot. All the cases show a good agreement with the true parameter values, as it resides in the 95% likelihood contours. Note that the 9 random simulations used for this analysis are realization of the same underlying cosmology. For this kind of analysis we expect the true parameters to lie outside of the 68% contour for more than $1/3$ of the times. As expected the angular range $[20', 40']$ has very little information about the two free parameters, compared to $[1', 20']$. Ergo, the likelihood distribution for the $[1', 40']$ range is comparable to the $[1', 20']$ case. In the following I will therefore, fix the angular range to $[1', 40']$ and to compare COSEBIs and CCOSEBIs for a tomographic and

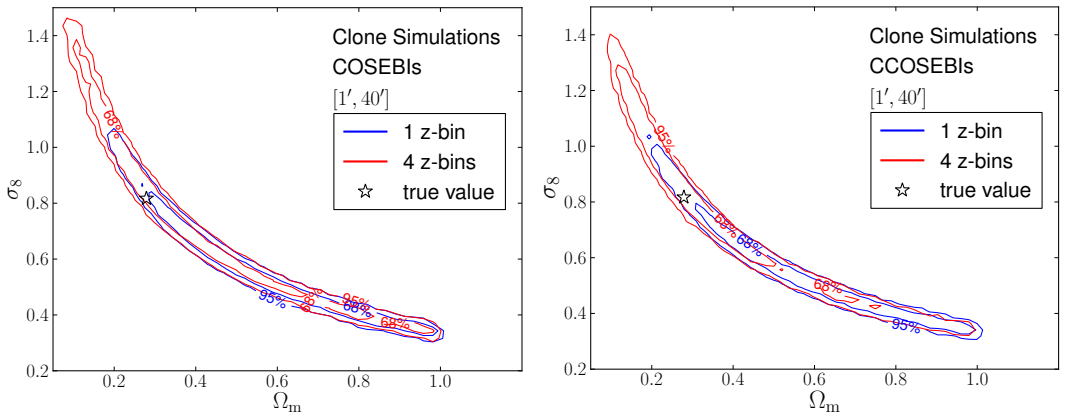

Figure 4.18: Likelihood contours for σ_8 and Ω_m measured from the Clone simulations for $\theta \in [1', 40']$. The left panel shows the likelihood when COSEBIs is the measured quantity, while the right one shows the same for CCOSEBIs. An MCMC algorithm is used to sample the likelihood. The 68% and 95% confidence regions are shown here. The true underlying parameters are shown with the star.

non-tomographic redshift binning.

Fig. 4.18 shows a comparison between the likelihood distributions of the cosmological parameters for a single redshift distribution and 4 redshift bins. The left and right panels show the results for COSEBIs and CCOSEBIs, respectively. As was shown in Fig. 2.12 the information gain with tomography for this specific case is very small. And here we can see that the width of the likelihood contour remains basically the same as tomography is considered. This is due to the fact that σ_8 and Ω_m are not very sensitive to redshift, when other parameters are fixed. However, the shape parameter, Γ, puts upper bound constraints on the estimated Ω_m when tomography is considered, which can be seen in the figure as a shift of the contours to smaller Ω_m values. Comparing the two panels we see that the likelihood regions are of similar size for the non-tomographic case. However, when tomography is considered the CCOSEBIs contours get smaller as is seen in Fig. 4.19. The reason for this behavior is that when the correction for the inverse covariance (Hartlap et al. 2007) is done the confidence regions get larger. For CCOSEBIs the number of observables is considerably smaller compared to the COSEBIs with 4 redshift bins (5 versus 70 modes). Hence this correction is much smaller for CCOSEBIs.

4.4 CFHTLenS Analysis

In the previous sections the COSEBIs and CCOSEBIs pipelines were rigorously tested, using analytic methods and simulated data. Here I will use the exact same pipelines on the CFHTLenS data and estimate the observables. Then I will use the MCMC code explained before to find the confidence regions and the best fit values for σ_8 and Ω_m, and compare to

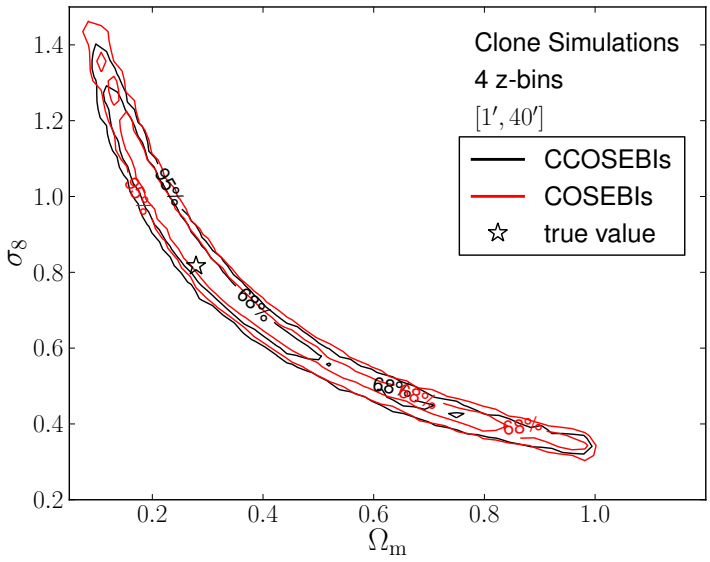

Figure 4.19: Likelihood contours for σ_8 and Ω_m measured from the Clone simulations with 4 redshift bins. An MCMC algorithm is used to sample the likelihood distribution. The 68% and 95% confidence regions are shown here. The true underlying parameters are shown with the star. The red contours show the results for COSEBIs versus the black contours for CCOSEBIs.

other published analysis of CFHTLenS with different statistical tools.

4.4.1 Measured Quantities

In Sect. 4.3.1 and Sect. 4.3.2 the methods for estimating the COSEBIs and the compressed COSEBIs (CCOSEBIs) are explained. Since the analysis here is blind, the method and the pipeline are unchanged for the analysis of the real data.

Fig. 4.20 shows the values of B-mode COSEBIs, B_n, and their goodness-of-fit to zero in the form of a reduced χ^2 with 7 degrees of freedom. The best fit value of the redshift of the galaxies used for making these plots belong to a single redshift distribution between [0.58, 1.3]. We can see that for all of the angular ranges shown here the B-modes are more or less consistent with zero. In particular, the angular range $[1', 20']$ shows a very small B-mode contribution. There is some B-mode residual present in the angular range $[20', 40']$. Regardless, this should not raise an alarm, since, most of the information, as we have seen already, comes from the lower angular range.

Figs. 4.21, 4.22 and 4.23 show the B_n modes for 4 redshift bins and the three angular ranges discussed before. Similar to the non-tomographic case the $[1', 20']$ range shows a very good agreement with zero, while there is a small disagreement for the $[20', 40']$ angular range which is also seen for the $[1', 40']$. In particular, the lower redshift bins for the autocorrelations show a higher inconsistency with zero for these angular ranges.

The CFHTLenS team has already estimated the cosmological parameters using cosmic shear analysis, as mentioned at the start of this Chapter. For example Kilbinger et al. (2013) used a single redshift distribution and combined their analysis with other cosmological probes to find more stringent constraint on the parameters. The values of the parameters they find is consistent with the true parameters of the Clone simulations. For this analysis I have fixed the

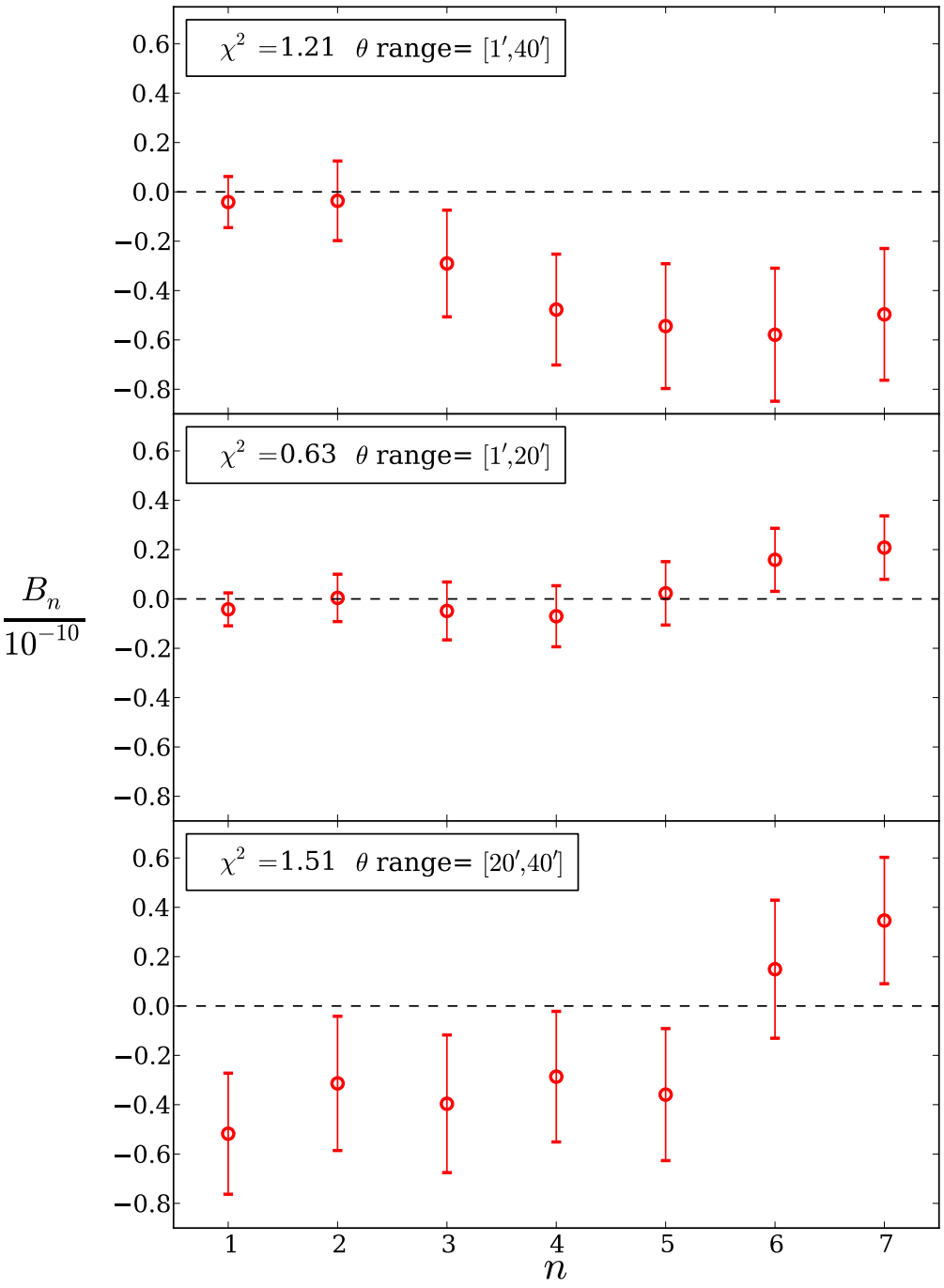

Figure 4.20: Measured B_n from the CFHTLenS data for a single redshift bin. Three angular ranges are considered here. For each plot the reduced χ^2 for the estimated B_n and the angular range are stated. Note that the B_n modes are correlated (see Fig. 4.5).

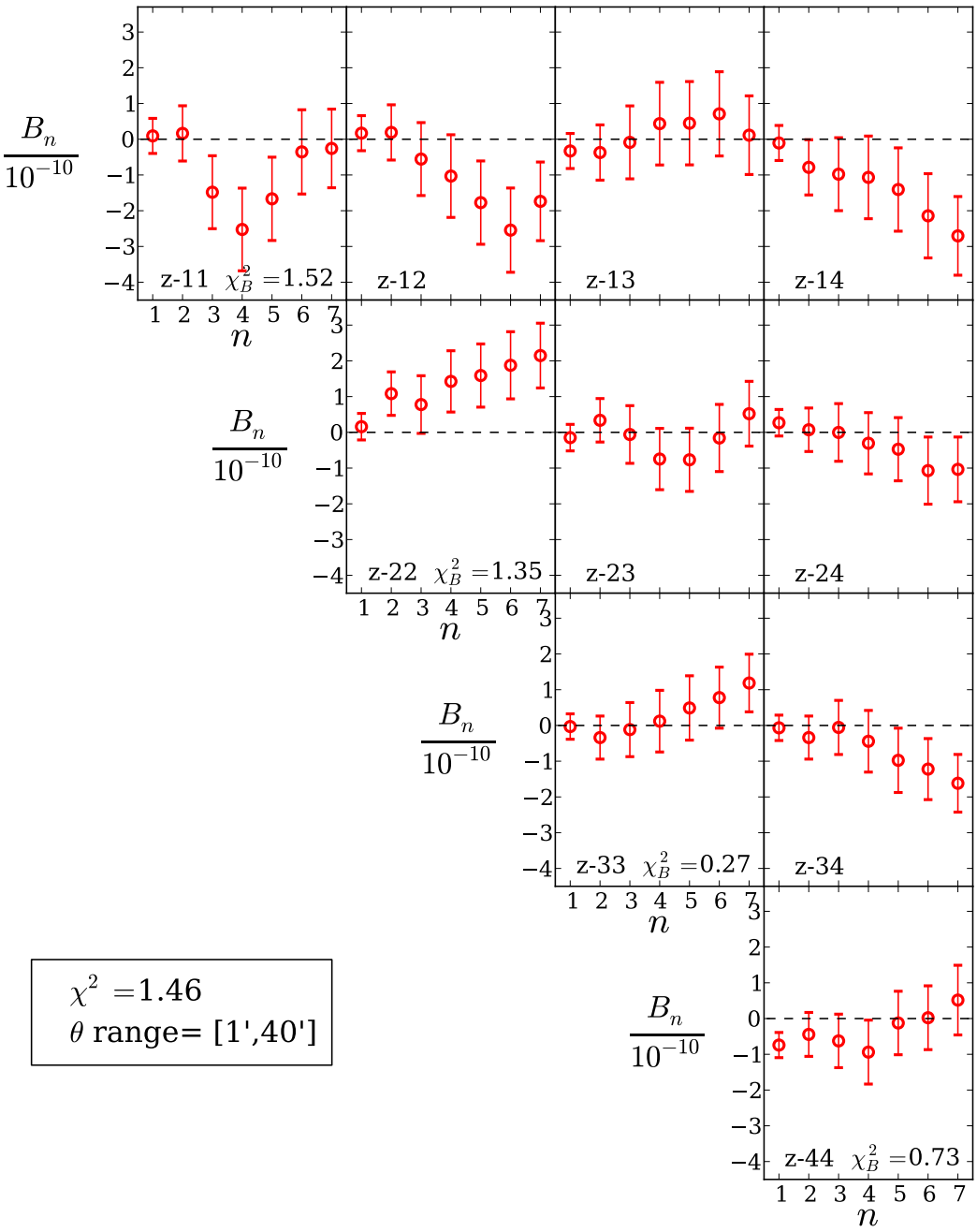

Figure 4.21: Measured B_n from the CFHTLenS data for 4 redshift bins. An angular range of $[1', 40']$ is considered here. The reduced χ^2 for the B-modes with respect to zero is declared in the legends. Each plot shows the B_n for the combination of redshift bins written on its corner. The reduced χ^2 is also declared for the autocorrelation cases. The data points shown here are correlated as seen in Fig. 4.6.

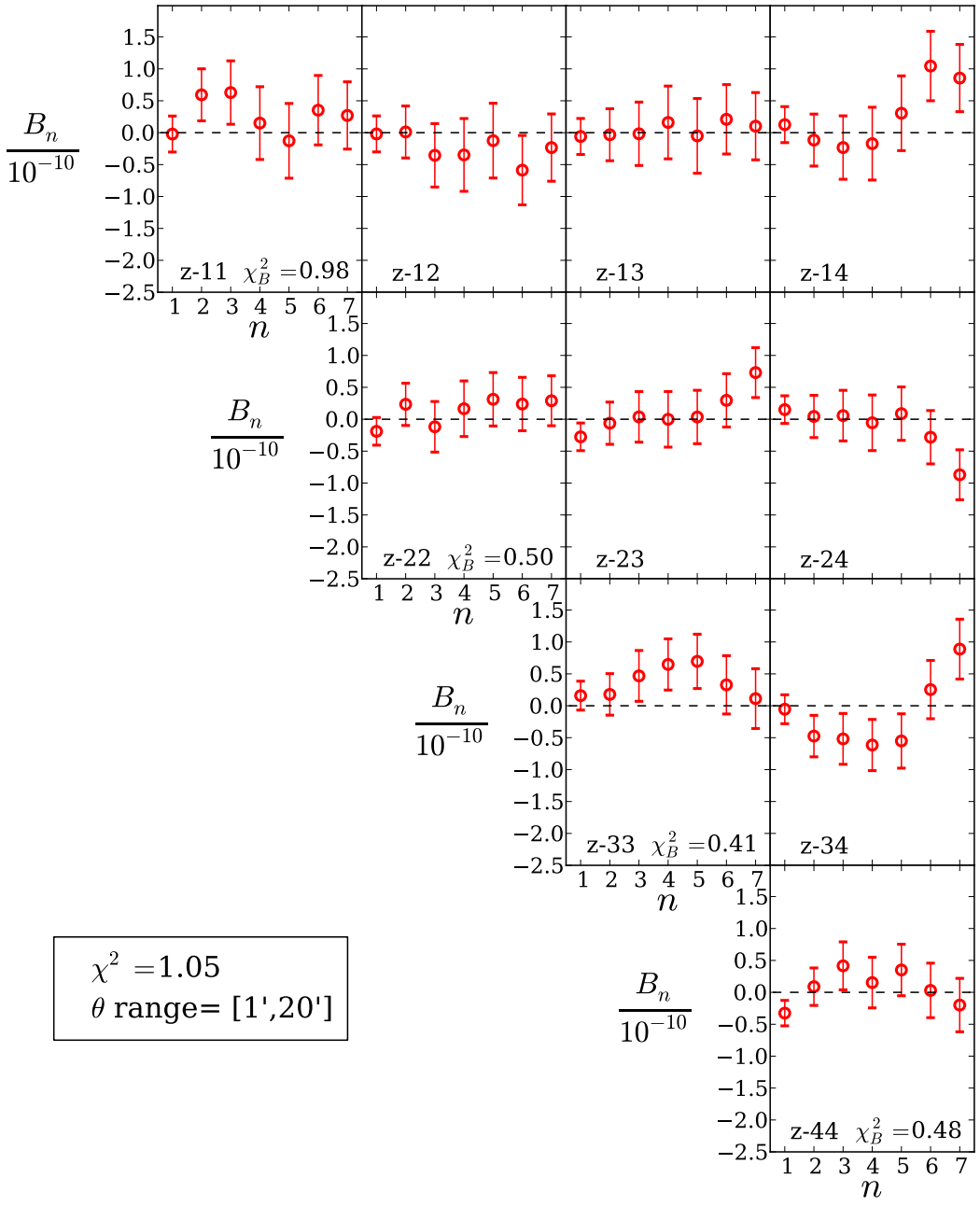

Figure 4.22: Measured B_n from the CFHTLenS data for 4 redshift bins. An angular range of $[1', 20']$ is considered here. See the caption of Fig. 4.21 for more details.

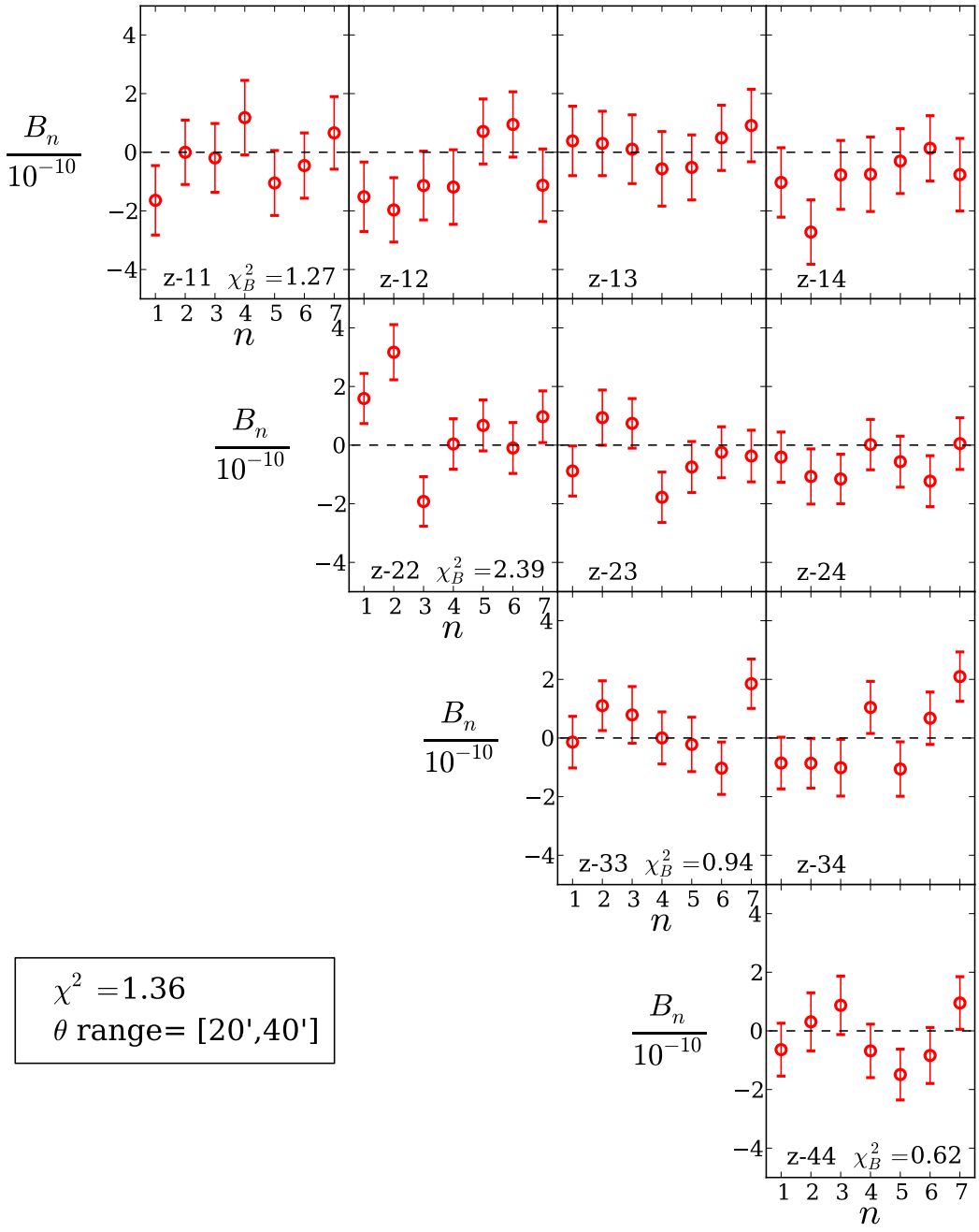

Figure 4.23: Measured B_n from the CFHTLenS data for 4 redshift bins. An angular range of $[20', 40']$ is considered here. See the caption of Fig. 4.21 for more details.

value of all the cosmological parameters, except σ_8 and Ω_m. These are shown in Table. 4.2 along with the priors on the free parameters. Before fitting to these two parameters using the MCMC code, I will show the estimated values of the E_n versus the theory values with the cosmological parameters found in Kilbinger et al. (2013). They find different best fitting values for their parameters when they combine their results with different surveys (for example WMAP7 and BOSS). Here these values,

$$\sigma_8 = 0.813 \qquad \Omega_m = 0.283 \,, \tag{4.12}$$

are chosen for the theory E_n, which are consistent with their results.

Fig. 4.24 shows the E_n modes estimated from the data for the non-tomographic case. The solid curves show the theory values with the parameters above. Recall that the E_n are only defined for natural values of n and the curves are drawn for visual purposes. The goodness-of-fit of the estimated E_n to the theory values is shown using a reduced χ^2. Note that the theory values are not a fit to the data. Therefore, there are 7 degrees of freedom for the χ^2. We can see that the estimated E_n and the theory are in good agreement. The E_n modes are correlated as can be seen through their covariance matrix in Fig. 4.5. We can see a good agreement between the data and the theory from the χ^2 values.

Figs 4.25, 4.26 and Fig. 4.27 show the measured E_n from the CFHTLenS data for the 4 redshift bins case. The solid curves represent the theory values of the observables for the cosmological parameters explained above. The redshift bin pair of each is labelled. The reduced χ^2 values are shown for all redshift bins and for the autocorrelations between the bins. The overall values of the reduced χ^2 indicate a good agreement between the theory and the data. The covariance matrices used in these plots are explained in Sect. 4.3.1 and shown in Fig. 4.6.

The CCOSEBIs are also estimated from the CFHTLenS data as explained in Sect. 4.3.2. Fig. 4.28 and Fig. 4.29 show the observed values of the CCOSEBIs E-/B-modes for a single and 4 redshift bins, respectively. Note that in both cases there are only 5 compressed observables, which depend on the derivatives of the cosmological parameters written below them on the x-axis of the plots. The results are shown for the three angular ranges, $[1', 40']$, $[1', 20']$ and $[20', 40']$. The covariance matrices for the CCOSEBIs are measured using the Clone simulations and are shown in Fig. 4.14 for the E-modes. The modes are correlated as can be seen from the covariance matrix. The solid curves correspond to the theory values for the E^c_μ, with the parameters explained at the start of this section. The reduced χ^2 values are shown for both modes. Note that this is not a fit, hence there are 5 degrees of freedom for the χ^2.

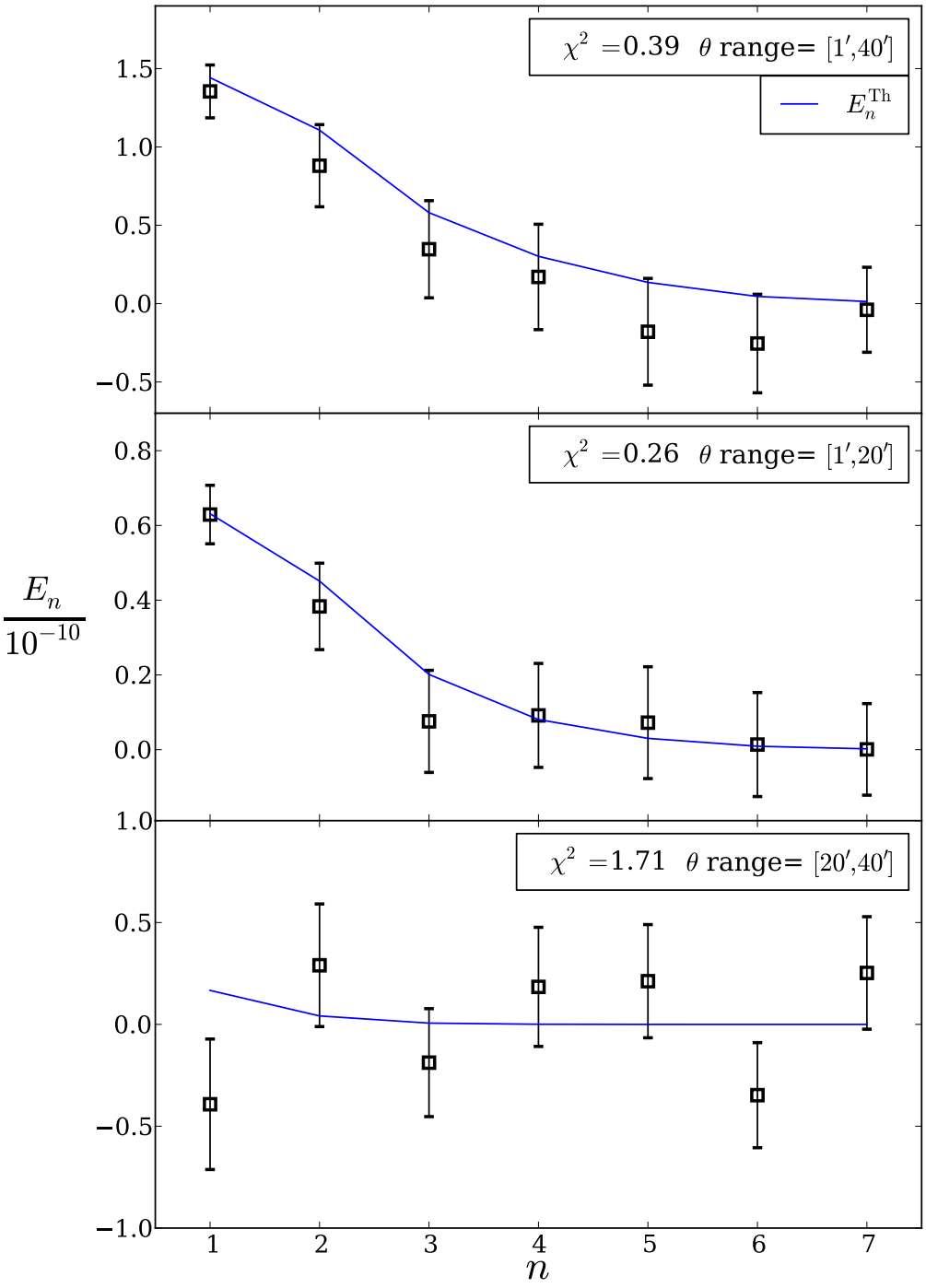

Figure 4.24: The E_n measured from the CFHTLenS data for a single redshift bin. The solid curves show the theory values of E_n with the best fit cosmological parameters of Kilbinger et al. (2013). The angular range of each panel is noted along with the value of the reduced χ^2 for the fit between the theory and the data. The errors are estimated from the Clone simulations (see Fig. 4.5).

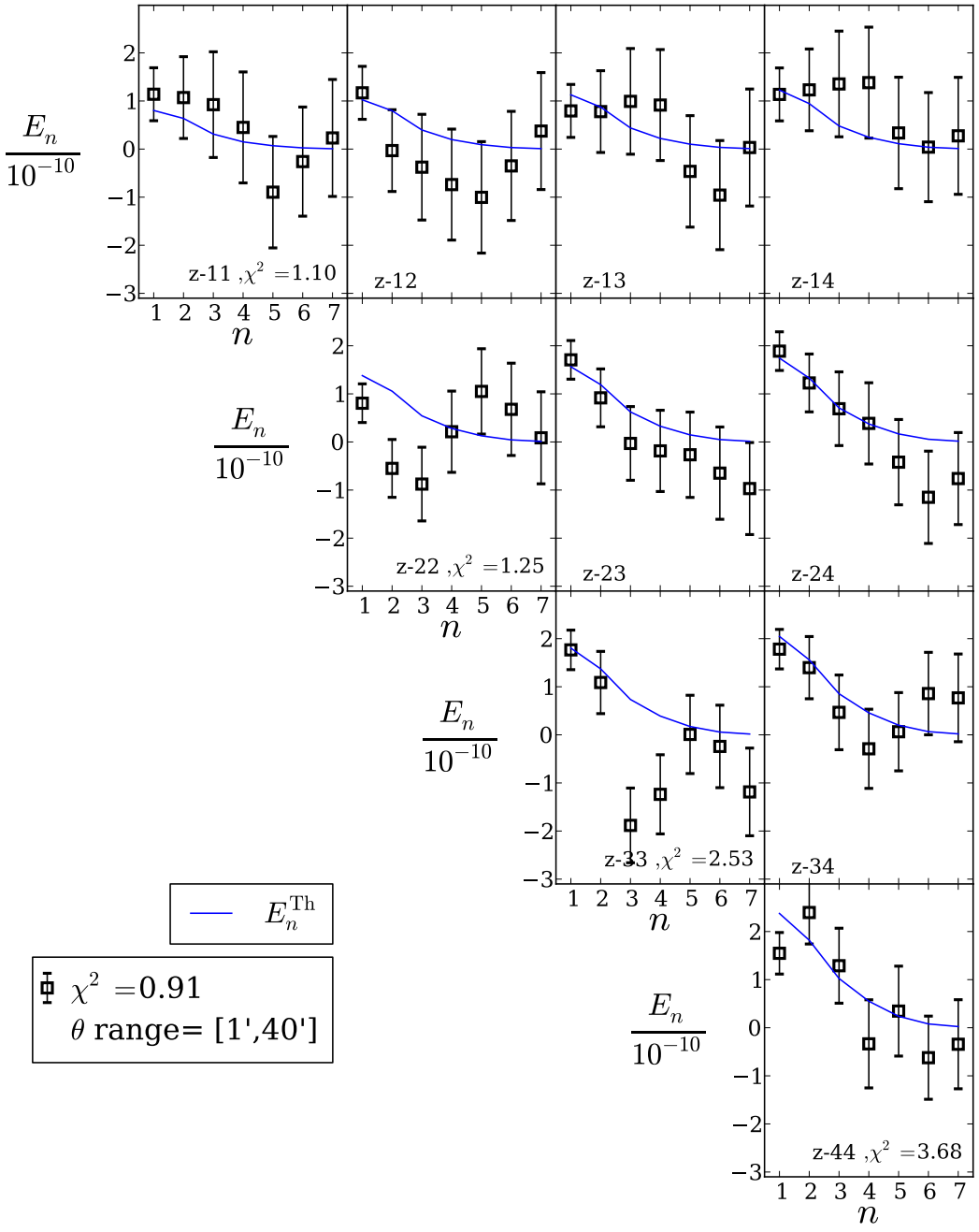

Figure 4.25: Estimated E_n from the CFHTLenS data for 4 redshift bins. An angular range of $[1', 40']$ is considered here. The solid curves show the value of the theory E_n with parameters deduced from previous CFHTLenS data analysis. The reduced χ^2 values show the agreement between the theory and the estimated values of E_n. Each plot shows the E_n for the combination of redshift bins written on its corner. The reduced χ^2 is also written for the autocorrelation cases.

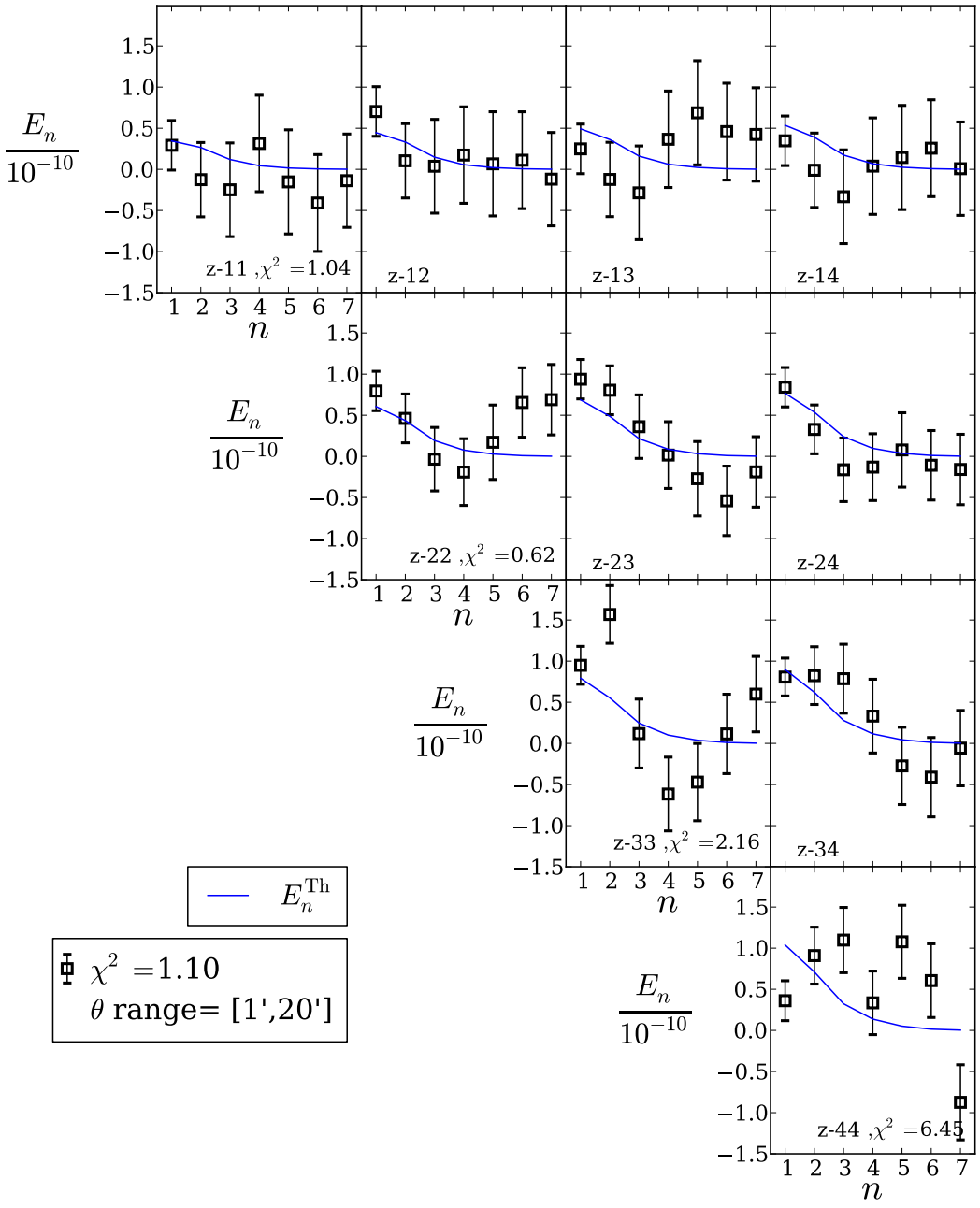

Figure 4.26: Estimated E_n from the CFHTLenS data for 4 redshift bins. An angular range of $[1', 20']$ is considered here. See the caption of Fig. 4.25 for more details.

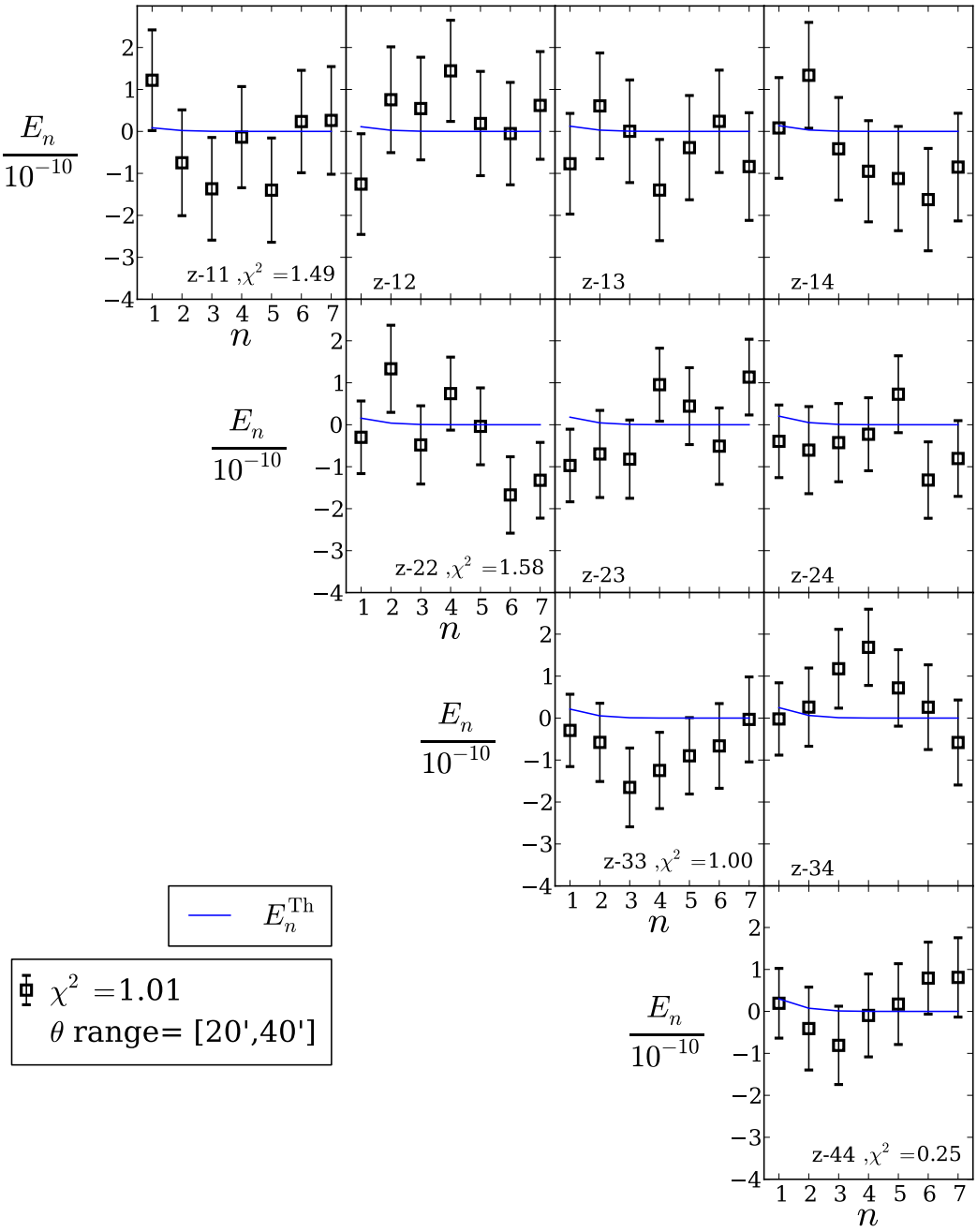

Figure 4.27: Estimated E_n from the CFHTLenS data for 4 redshift bins. An angular range of $[20', 40']$ is considered here. See the caption of Fig. 4.25 for more details.

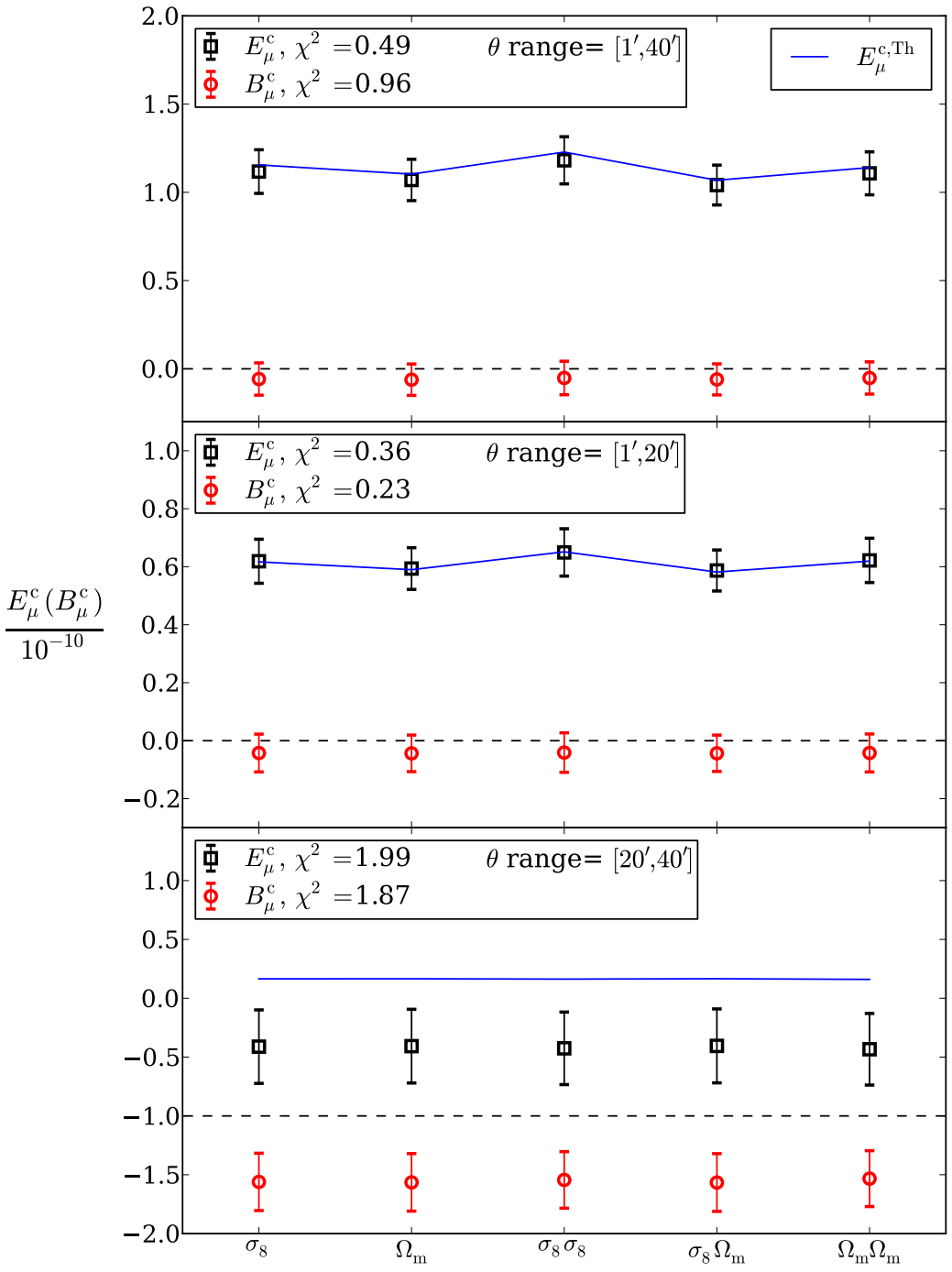

Figure 4.28: Estimated CCOSEBIs from the CFHTLenS data. The CCOSEBIs are linear combinations of the COSEBIs. A single redshift distribution is assumed here. The errors on the measurements come from the Clone simulations. The solid curves show the theory values of the E_μ^c for the cosmological parameters in Kilbinger et al. (2013) (see the start of this section for more details). Note that the CCOSEBIs modes are discrete and the theory values are connected for visual purposes. The reduced χ^2 values are shown for the theory values for each angular range. The B_μ^c and their goodness-of-fit to zero are shown as well. The x-axis shows which derivatives are used for making the CCOSEBIs mode.

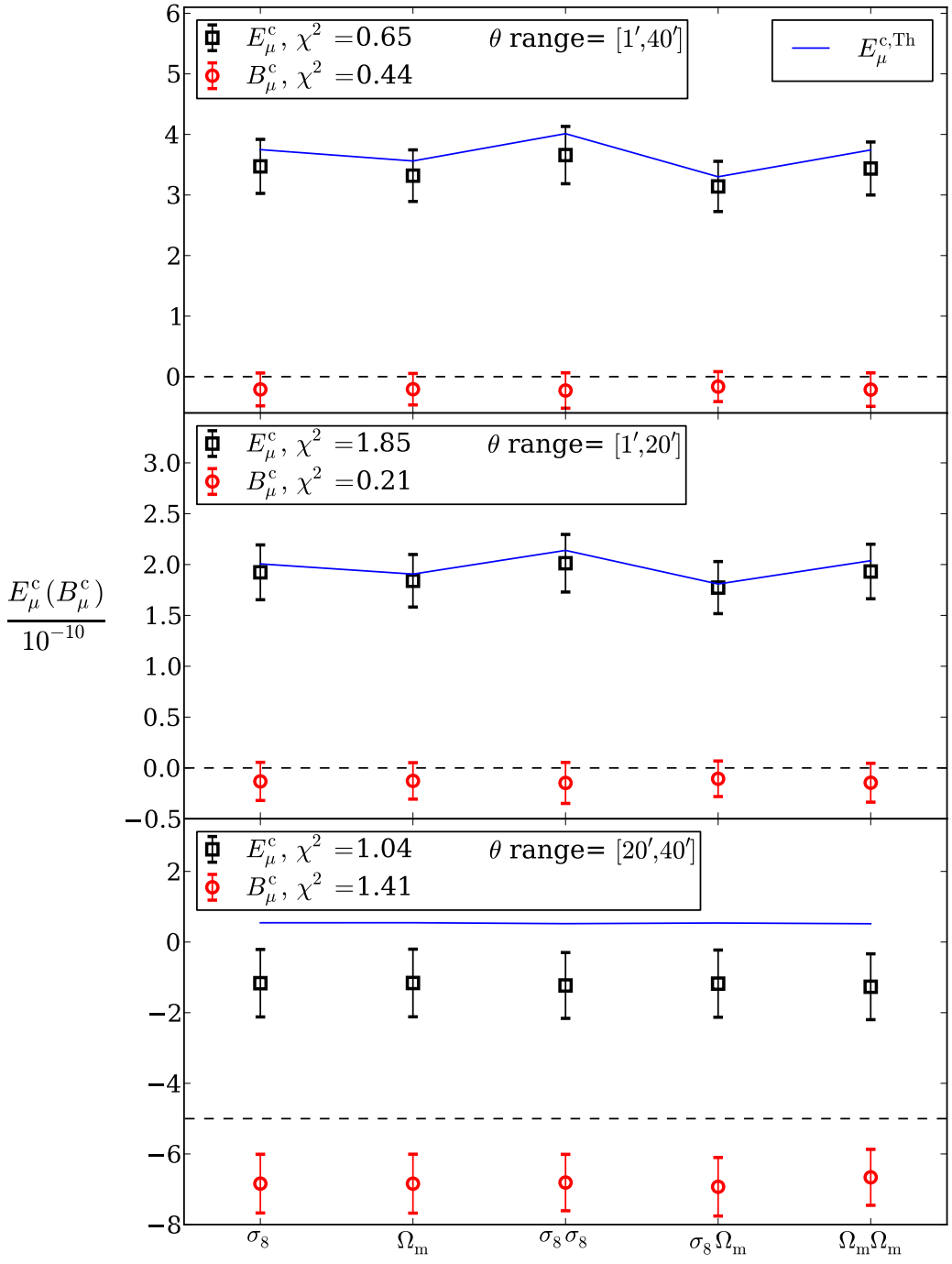

Figure 4.29: Measured CCOSEBIs from CFHTLenS using 4 tomographic redshift bins. See the caption of Fig. 4.28 for more information.

4.4.2 MCMC Results

In Sect. 4.3.3 we saw the measured likelihood confidence regions for the two free parameters σ_8 and Ω_m of the Clone simulations. A comparison between the different cases were shown in the figures of that section. Here I will show similar result using the CFHTLenS data.

Fig. 4.30 shows a comparison between three angular ranges. All of the cases are consistent with the Kilbinger et al. (2013) result, shown with the star.

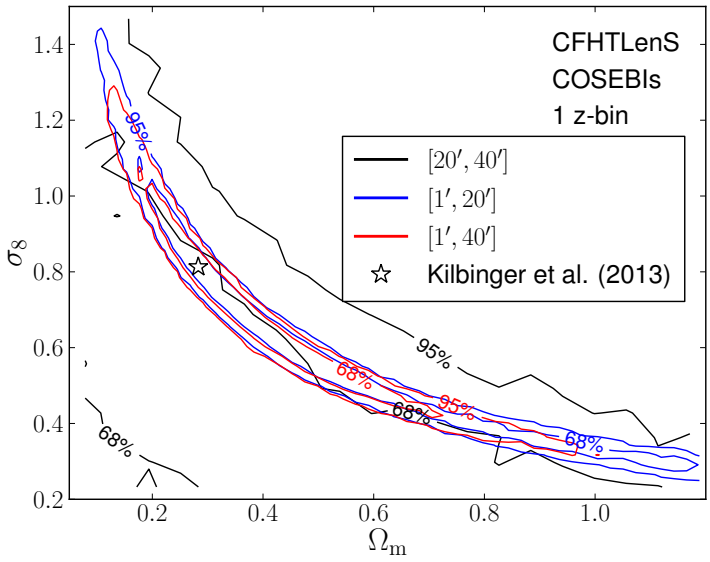

Figure 4.30: Likelihood contours for σ_8 and Ω_m measured from the data. A single redshift distribution and 7 E-mode COSEBIs are assumed here. An MCMC algorithm is used to sample the likelihood. The 68% and 95% confidence regions are shown here. The best fit values in Kilbinger et al. (2013) are shown with the star.

Fig. 4.31 shows the results for $[1', 40']$ with and without tomography. The left panel shows the confidence regions using COSEBIs, whereas the right panels belong to CCOSEBIs. Comparing the two panels we see that the likelihood regions are of similar size for the non-tomographic case. Similar to what we saw for the Clone simulations here we do not see a big information gain with tomography.

Fig. 4.32 shows the likelihood regions using 4 redshift bins for COSEBIs and CCOSEBIs for a better comparison. Similar to the case for the simulations (Fig. 4.19), we see that the contours are smaller for CCOSEBIs due to the Hartlap et al. (2007) correction for the inverse covariance.

Weak lensing data is mostly sensitive to this combination of σ_8 and Ω_m,

$$\Sigma_8 = \sigma_8(\Omega_m/c)^\alpha \ , \tag{4.13}$$

where c is a constant which is typically equal to the best fit value of Ω_m. In this work I use $c = 0.27$ similar to the previous CFHTLenS papers, Kilbinger et al. (2013) and Heymans et al. (2013). We can find Σ_8 and α for all the cases that has been shown in this section. To find α I use a curve fitting routine and fix it to the best fit value. Then a distribution for Σ_8 is estimated using Eq. (4.13). The best fit value of Σ_8 is measured from the maximum value

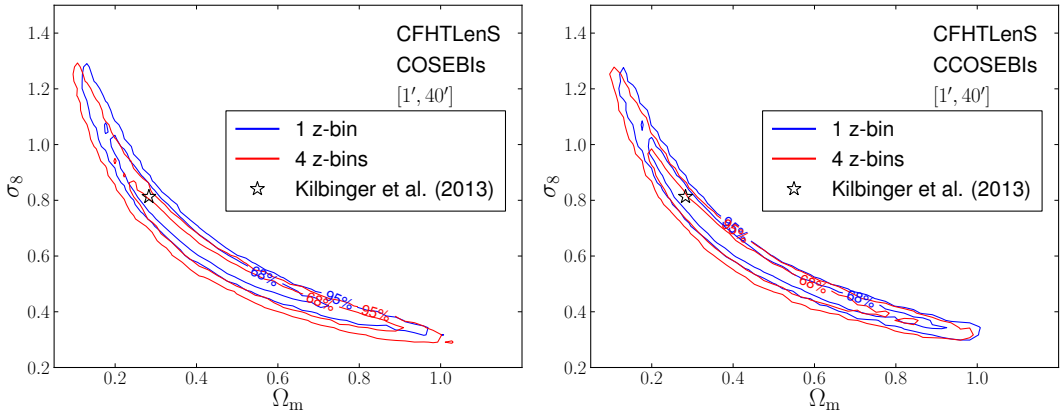

Figure 4.31: Likelihood contours for σ_8 and Ω_m measured from the data for $\theta \in [1', 40']$. The left panel shows the likelihood when COSEBIs is the measured quantity, while the right one shows the same for CCOSEBIs. An MCMC algorithm is used to sample the likelihood. The 68% and 95% confidence regions are shown here. The best fit values in Kilbinger et al. (2013) are shown with the star.

Data	α	Σ_8
E_n, 1-bin, $[20', 40']$	0.34	$0.372^{+0.352}_{-0.097}$
E_n, 1-bin, $[1', 20']$	0.64	$0.822^{+0.046}_{-0.052}$
E_n, 1-bin, $[1', 40']$	0.62	$0.825^{+0.033}_{-0.044}$
E_n, 4-bins, $[1', 40']$	0.62	$0.764^{+0.046}_{-0.048}$
E^c_μ, 1-bin, $[1', 40']$	0.64	$0.818^{+0.035}_{-0.044}$
E^c_μ, 4-bins, $[1', 40']$	0.59	$0.782^{+0.045}_{-0.051}$

Table 4.5: The constraints on Σ_8 from CFHTLenS data. Column one shows the data that is used to find the constraints. α is calculated using by fitting to $\Sigma_8 = \sigma_8(\Omega_m/0.27)^\alpha$. The value of the best fit Σ_8 and its associated errors are estimated from the distribution of Σ_8 by fixing α, and finding the values for Σ_8 around its maximum that correspond to 68% of the full area of the distribution. E_n and E^c_μ represent the COSEBIs and CCOSEBIs, respectively. The results are shown for one and four redshift bins with three angular ranges.

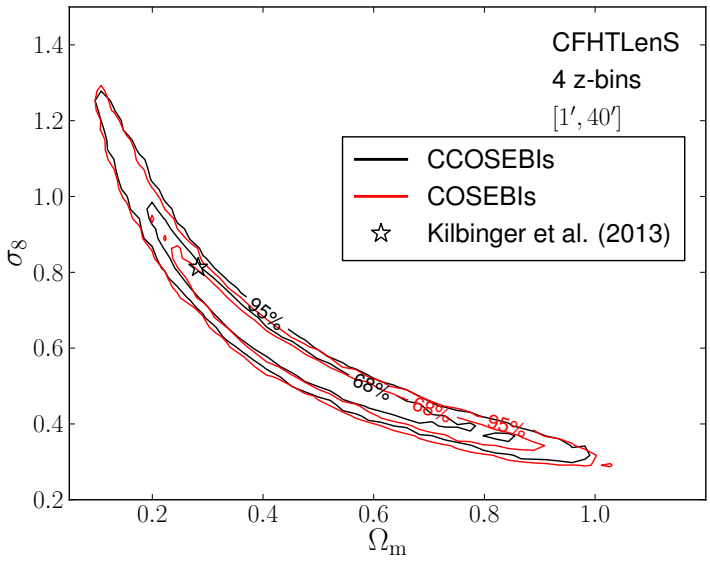

Figure 4.32: Likelihood contours for σ_8 and Ω_m measured from the CFHTLenS data with 4 redshift bins. The black contours show 68% and 95% confidence regions for CCOSEBIs versus the red contours for COSEBIs. An MCMC algorithm is used to sample the likelihood. The best fit values in Kilbinger et al. (2013) are shown with the star.

of the distribution. The error on this measurement is calculated by finding the values of Σ_8 between which 68% of the distribution lies. Table. 4.5 summarizes the results. The errors on Σ_8 indicate the width of the σ_8-Ω_m banana, (see Fig. 4.30, Fig. 4.31 and Fig. 4.32). The largest angular range, $[20', 40']$ as we have seen before, contains little information about these parameters. However, going from $[1', 20']$ to $[1', 40']$ slightly improves the constraints. The tomography cases although consistent with each other, show slightly larger errors compared to to the single redshift distribution cases. The constraints on σ_8-Ω_m, should not vary significantly with tomography, as was seen in Fig. 2.12. Although, we do not expect them to deteriorate. The apparent increase of the constraint areas, is not significant. However, there are effects such as the error in estimating the covariance from a limited number of simulations in the case of COSEBIs, and inaccuracies in the estimated likelihood distributions from the MCMC chains, can result in such differences. Moreover, applying the parameter covariance correction in Eq. (4.11) which has been neglected in this work will decrease the errors on the estimated parameters, especially for the COSEBIs with tomography. For my future work with more free cosmological parameters, the redshift binning will play a more central role. As the other cosmological parameters, specially the dark energy parameters, are more sensitive to the redshift distribution.

4.5 Conclusion

CFHTLenS is the largest completed weak gravitational lensing survey, to date. In this Chapter I showed a blind cosmic shear analysis of this dataset using two methods. To set up for the blind analysis, I first applied and tested the methods numerically and then on the Clone simulations, which have the specifics of CFHTLenS integrated within them. Once the pipeline was

debugged and tested, I applied it to the data without modifications. This is important because changing the pipelines in accordance to the data analysis results, produces confirmation biases.

In this thesis I chose a redshift where it was sufficient to ignore the contribution to the measured 2PCFs signal from galaxy alignments which can contaminate the weak lensing signal, by choosing the redshift range $[0.58, 1.3]$, for which, the contribution from intrinsic alignments was shown to be small (Heymans et al. 2013). Since the area that CFHTLenS covers is small for cosmic shear analysis to be competitive with other probes of cosmology, the expected constraints from its data are loose, as found. However, this is the first ever measurement of CCOSEBIs on data and COSEBIs on tomographic data. The likelihood distribution for these parameters were sampled using an MCMC algorithm explained in Sect. 2.1.3. The Clone simulations were used to find the covariance of the data. These simulations pose limitations on the scales with reliable measurements. Consequently, an angular range of $[1', 40']$ was chosen for this analysis. Future work will use the current state of the art simulations to allow for even larger scales to be probed. In addition, this range was divided into two equal ranges, $[1', 20']$ and $[20', 40']$, to do consistency checks and find out where most of the information comes from. Moreover, two redshift binning setups were considered: a tomographic case with 4 redshift bins and a single redshift distribution.

The first method used in this section was COSEBIs (Complete Orthogonal Sets of E-/B-Integrals), which separates E-modes from B-modes in the data. We expect weak lensing to predominantly produce E-modes, making B-mode undesirable. Most of the B-mode contribution typically comes from remaining systematics from the data reduction procedure. Although, the absence of B-modes does not guarantee a perfect data reduction, it is a necessary condition for a survey like CFHTLenS. For future large scale and space based surveys, where the measurement errors are significantly smaller the B-modes could also indicate other physical phenomena. For example, some of the galaxy alignment models predict these modes (see Blazek et al. 2011, and references therein). The B-modes measured from CFHTLenS were in most cases consistent with zero. The $[20', 40']$ range showed a small B-mode contribution for both the single bin and tomographic analysis. However, the information level in this range is lower compared to the smaller scales and the B-modes that exist here should not raise a flag.

Kilbinger et al. (2013) performed a non-tomographic analysis of CFHTLenS using COSEBIs. They reported some difficulties with the measurements. Since the higher COSEBIs modes have more oscillations in their filter functions, it is necessary to use a finely binned shear two point correlation functions to measure them. Here I did not encounter similar problems, due to binning. The tomographic analysis of cosmic shear data with COSEBIs was done for the first time in this work. Comparing the tomographic case with the single redshift distribution, we saw that the information gain is very small as expected from the Fisher forecast in Fig. 2.12. In fact the tails of the likelihood contours for the tomographic case are longer.

I used the compression method discussed in Chapter 3 to make compressed COSEBIs,

CCOSEBIs. I applied this method for the very first time on data. Previously, I tested this compression method on COSEBIs using a Fisher analysis, to determine if any information is lost in this process. Here I used a full likelihood analysis of σ_8 and Ω_m and showed that no information is lost due to this compression. The conclusion is that this method is robust and is recommended for any cosmological analysis, where data compression is beneficial. The higher the number of observables, the higher the number of simulations needed to estimate their covariance matrix accurately. For a relatively small number of simulations the measured inverse covariance matrix needs to be corrected. This correction can stretch the likelihood regions for the estimated parameters. Using this compression method the number of observables shrinks which makes the estimated covariances more reliable. We see this effect by comparing the likelihood contours of COSEBIs and CCOSEBIs for 4 redshift bins for the simulations and the data alike (Figs 4.19 and 4.32).

The methods and pipelines used in this analysis can be used on any other reduced cosmic shear dataset, such as KiDS[3], DES which has started collecting data (although it has not reached a higher area than CFHTLenS yet), and Euclid. In future I will add the lower redshift galaxies and intrinsic alignment models to the analysis. This will allow more free parameters to be included and constrained. However, CFHTLenS is a rather small survey and any cosmic shear analysis done with its data is more for demonstration and preparation purposes. Here I showed that COSEBIs and CCOSEBIs and in general this compression method is applicable to real data and produces robust results.

After the submission of this thesis I redid the analysis with the full CFHTLenS redshift range and updated simulations (which will be published as Asgari et al. 2015). I found significant B-modes at large scales, especially when tomographic bins are involved. With the updated simulations I no longer encounter very small χ^2 values. Hence, I conclude that the noise in the simulations used in this thesis was responsible for this problem.

[3]Kilo Degree Survey: http://kids.strw.leidenuniv.nl/

Chapter 5

Flat Sky Pseudo-Cl Analysis

Pseudo-Cl (PCl) analysis, which is concerned with Fourier mode-mixing effects of masked data, has been used to analyse cosmic microwave background (CMB) temperature (see Planck Collaboration et al. 2014a, for a full sky analysis of the CMB temperature) and polarization data (see Brown et al. 2005a, 2009, and references therein, for a full sky analysis of simulated and QUaD survey data). Shear and electromagnetic polarization are mathematically identical quantities, therefore, any method that is applicable to one should in principle work for the other as well. CMB data is also analysed using likelihood methods (Planck Collaboration et al. 2014a, see for example), which can also be applied to shear fields (see Seljak 1998; Hu & White 2001) as was done by Brown et al. (2003) for the COMBO-17 survey. However, these methods are too slow for current and future surveys where the number of pixels and the general resolution is high.

Most cosmic shear analysis measure the shear two point correlation functions (2PCFs), as we saw in the earlier Chapters. This is due to the fact that the measured 2PCFs are not biased by the presence of a mask on the images[1], which is always present (see Sect. 2.4). Although, the measurement of the power spectrum is complicated by the presence of a mask, there are advantages in estimating them. The covariance of the power spectrum is diagonal for the linear perturbation regime. Hence, using PCl analysis is specially advantageous for CMB data, where most Fourier modes are still in the linear regime. However, for more recent epochs, where most of the weak gravitational lensing information comes from, the scales with the highest cosmological information are well into their non-linear evolution. Nevertheless, we can still use the larger scales to look for systematic errors, since the theory and the covariance are better known for these scales. As we saw in Sect. 2.6 by using COSEBIs we can choose the angular range used in the analysis. Similarly, a PCl analysis can be restricted to certain Fourier modes to restrict the baryon and clustering effects which are hard to model. In more detail it is worth noting that in a power spectra analysis removing modes contaminated by for example non-

[1]This is true assuming that the mask is not correlated with the shear field. In the presence of these kind of correlations, both 2PCFs and PCls will be biased.

linearity and baryon effects is simpler compared to 2PCFs, from a theoretical point of view. In particular, we can tune n-body simulations to match PCls, so for instance we only need to model to $k = 5h/\mathrm{Mpc}$ for a Euclid analysis with PCls, but would need to reach much larger k-modes for accurate estimation of 2PCFs. The modelling of matter power spectrum is much more challenging for larger Fourier modes, due to non-linear effects. In addition, in contrast to 2PCFs a PCl analysis should result in a positive definite signal, which is a good consistency check.

PCl is divided into two categories, flat and full sky. For a full sky analysis one needs to use spherical harmonics to transform the real space quantities (see Hivon et al. 2002 for a spin-0 and Brown et al. 2005b for a spin-2 field). The advantage of a full sky analysis is that the Fourier modes are well defined. Although, going to small scale (high Fourier modes), requires an impractically high number of pixels. On the other hand, in the flat sky approximation (see Memari 2009, for a flat sky analysis of CMB QUaD data), the Fourier modes are continuous, hence the discretization of the data will be arbitrary, which can bias the results. Nevertheless, flat sky approximation allows one to choose the range of the Fourier modes in the analysis, and as a result reaching higher modes is achievable.

PCl analysis on a curved (full) sky has not yet been applied to weak lensing data. Nevertheless, Hikage et al. (2011) applied this method to simulated data for small masks. They also investigate flat sky PCl estimation using similar masks. Hikage et al. (2011) show that these methods results in sub-percent accuracies in the estimated power spectra. However, in their study they ignore the effects of large masks and assume period boundary conditions for their images in the flat sky case. In addition, Hand et al. (2013) have measured the flat sky PCls for cross-correlations of Atacama[2] CMB maps and CFHTLenS[3] galaxy lensing convergence maps.

In this Chapter I will study the effects of masking for both small and large masks, using Gaussian and lognormal simulated shear fields, for the flat sky limit. I will first go through the formalism of PCls in Sect. 5.1, where I explain how the mode mixing can be modelled via a mixing matrix. In Sect. 5.2 the resulting pseudo power spectra and the recovered power spectra will be shown and compared with their expected values from theory. Finally in Sect. 5.3 I will propagate the random and mask modelling errors to the cosmological parameters using a Fisher analysis.

5.1 Formalism

Cosmic shear analysis tools and observables have been introduced in Chapter 2. In Sect. 2.4.1 we saw how the presence of B-modes affect the convergence, κ, and the shear, γ. Eq. (2.57)

[2]http://www.princeton.edu/act/
[3]http://www.cfhtlens.org/

shows that the convergence can be separated into two real parts $\kappa_{\rm E}$ and $\kappa_{\rm B}$ in real space. In Fourier space, we can write κ in terms of the Fourier transforms of $\kappa_{\rm E,B}(\boldsymbol{\vartheta})$,

$$\hat{\kappa}_{\pm}(\boldsymbol{\ell}) = \hat{\kappa}_{\rm E}(\boldsymbol{\ell}) \pm i\hat{\kappa}_{\rm B}(\boldsymbol{\ell}) \,, \tag{5.1}$$

with

$$\hat{\kappa}_{\rm E,B}(\boldsymbol{\ell}) = \int {\rm d}^2\vartheta \, \kappa_{\rm E,B}(\boldsymbol{\vartheta}) e^{-i\boldsymbol{\ell}\cdot\boldsymbol{\vartheta}} \,. \tag{5.2}$$

Note that, $\hat{\kappa}_{\rm E,B}(\boldsymbol{\ell})$ are complex quantities. We can also write $\gamma_{\pm}$ as,

$$\gamma_{\pm}(\boldsymbol{\vartheta}) = \gamma_1(\boldsymbol{\vartheta}) \pm i\gamma_2(\boldsymbol{\vartheta}) \,, \tag{5.3}$$

and in Fourier space,

$$\hat{\gamma}_{\pm}(\boldsymbol{\ell}) = \hat{\gamma}_1(\boldsymbol{\ell}) \pm i\hat{\gamma}_2(\boldsymbol{\ell}) \,, \tag{5.4}$$

where $\hat{\gamma}_{1,2}(\boldsymbol{\ell})$ are the Fourier transforms of $\gamma_{1,2}(\boldsymbol{\vartheta})$, respectively. The relation between $\hat{\kappa}_{+}(\boldsymbol{\ell})$ and $\hat{\gamma}_{+}(\boldsymbol{\ell})$ is given in Eq. (2.49). To find the relation between $\hat{\kappa}_{-}(\boldsymbol{\ell})$ and $\hat{\gamma}_{-}(\boldsymbol{\ell})$ we use a similar formalism to the one used in Sect. 2.4. $\kappa_{\pm}(\boldsymbol{\vartheta})$ and $\gamma_{\pm}(\boldsymbol{\vartheta})$ are functions of the lensing potential, $\psi_{\pm}$, via,

$$\gamma_{+} = \frac{1}{2}\partial\partial\psi_{+} \,, \qquad \gamma_{-} = \frac{1}{2}\partial^*\partial^*\psi_{-} \,, \qquad \kappa_{\pm} = \frac{1}{2}\partial^2\psi_{\pm} \,, \tag{5.5}$$

where

$$\psi_{\pm} = \psi_{\rm E}(\boldsymbol{\vartheta}) \pm i\psi_{\rm B}(\boldsymbol{\vartheta}) \,, \qquad \psi_{-} = \psi_{+}^* \,, \tag{5.6}$$

∂ was defined in Eq. (2.44) and the last equation shows that ψ_{-} is the complex conjugate of ψ_{+}. The Fourier transform of ∂ is given in Eq. (2.47), as $i\hat{\ell}$, which was used to connect the convergence and shear in Fourier space in Eq. (2.49). Similarly, $\hat{\kappa}_{\pm}$ and $\hat{\gamma}_{\pm}$ are related via,

$$\begin{aligned} \hat{\kappa}_{+}(\boldsymbol{\ell}) &= \hat{\ell}^*\hat{\ell}^*|\hat{\ell}|^{-2}(\hat{\gamma}_1 + i\hat{\gamma}_2)(\boldsymbol{\ell}) \,, \\ \hat{\kappa}_{-}(\boldsymbol{\ell}) &= \hat{\ell}\hat{\ell}|\hat{\ell}|^{-2}(\hat{\gamma}_1 - i\hat{\gamma}_2)(\boldsymbol{\ell}) \,. \end{aligned} \tag{5.7}$$

Simplifying the above equations by substituting for $\hat{\ell}$ from,

$$\hat{\ell} = \ell e^{i\varphi_\ell} \,, \quad \text{with} \quad \ell = |\hat{\ell}| \,, \tag{5.8}$$

results in

$$\hat{\kappa}_{\pm}(\boldsymbol{\ell}) = e^{\mp 2i\varphi_\ell}\hat{\gamma}_{\pm}(\boldsymbol{\ell}) \,, \tag{5.9}$$

where φ_ℓ is the polar angle of both $\boldsymbol{\ell}$ and $\hat{\ell}$.

5.1.1 Applying Masks to the Shear Fields

In any realistic scenario parts of the images are masked. Formally we only need to know the position of source galaxies. However, analysing a gridded image is significantly faster, since fast Fourier transforms can be utilized in this case. A PCl analysis relies on such gridded fields, where any region with no signal resulting from observers choices or faulty and empty pixels produce the mask. For practical reasons which will become clearer in the results section (Sect. 5.2), we can choose to apodise the masked shear field with a smoothing kernel, S. The masks that are going to be used in this thesis, consist of ones and zeros, exclusively. However, in practice the masks present on the measured shear fields are usually smoother, due to dithering of the observed images. If the mask provided by the observer is smooth enough then apodisation will not be necessary. A mask, W, has a multiplicative effect on the shear field,

$$\gamma_{\pm}^{\rm ps}(\boldsymbol{\vartheta}) = W(\boldsymbol{\vartheta})\gamma_{\pm}(\boldsymbol{\vartheta}) \,. \tag{5.10}$$

Any quantity with a ps denotes a masked or pseudo quantity from hereon. There are two ways to apodise a mask, one is to convolve the masked shear field, $\gamma^{\rm ps}(\boldsymbol{\vartheta})$, with S,

$$\gamma_{\pm{\rm ap}}^{\rm ps}(\boldsymbol{\vartheta}) = [S * (W\gamma_{\pm})](\boldsymbol{\vartheta}) \equiv \int {\rm d}^2\boldsymbol{\vartheta}' S(\boldsymbol{\vartheta} - \boldsymbol{\vartheta}')W(\boldsymbol{\vartheta}')\gamma_{\pm}(\boldsymbol{\vartheta}') \,. \tag{5.11}$$

The other method is to take the mask and smoothen its edges with a kernel. Note that when this apodisation method is used, the mask will maintain its original zeros while smoothly transitioning to the unmasked parts, where $W(\boldsymbol{\vartheta}) = 1$. Therefore, using this method enlarges the mask. The original mask is then replaced by the new apodised mask.

In Fourier space the shear field is first convolved with the mask,

$$\hat{\gamma}_{\pm}^{\rm ps}(\boldsymbol{\ell}) = \int \frac{{\rm d}^2\ell'}{(2\pi)^2} \widehat{W}(\boldsymbol{\ell} - \boldsymbol{\ell}')\hat{\gamma}_{\pm}(\boldsymbol{\ell}') \,, \tag{5.12}$$

and then multiplied by the smoothing kernel if the first apodisation method is used,

$$\hat{\gamma}_{\pm{\rm ap}}^{\rm ps}(\boldsymbol{\ell}) = \hat{S}(\boldsymbol{\ell}) \int \frac{{\rm d}^2\ell'}{(2\pi)^2} \widehat{W}(\boldsymbol{\ell} - \boldsymbol{\ell}')\hat{\gamma}_{\pm}(\boldsymbol{\ell}') \,, \tag{5.13}$$

Any quantity that has a subscript "ap" has been apodised with the first smoothing method. Substituting from Eq. (5.9) into the above equation we can find a relation for the masked κ,

$$\hat{\kappa}_{\pm}^{\rm ps}(\boldsymbol{\ell}) = \int \frac{{\rm d}^2\ell'}{(2\pi)^2} \widehat{W}(\boldsymbol{\ell} - \boldsymbol{\ell}')\hat{\kappa}_{\pm}(\boldsymbol{\ell}')e^{\mp 2i\varphi_{\ell\ell'}} \qquad \text{and} \qquad \hat{\kappa}_{\pm{\rm ap}}^{\rm ps}(\boldsymbol{\ell}) = \hat{S}(\boldsymbol{\ell})\hat{\kappa}_{\pm}^{\rm ps}(\boldsymbol{\ell}) \,, \tag{5.14}$$

where $\varphi_{\ell\ell'} = \varphi_{\ell} - \varphi_{\ell'}$. By adding and subtracting the equations above, we can find a relation

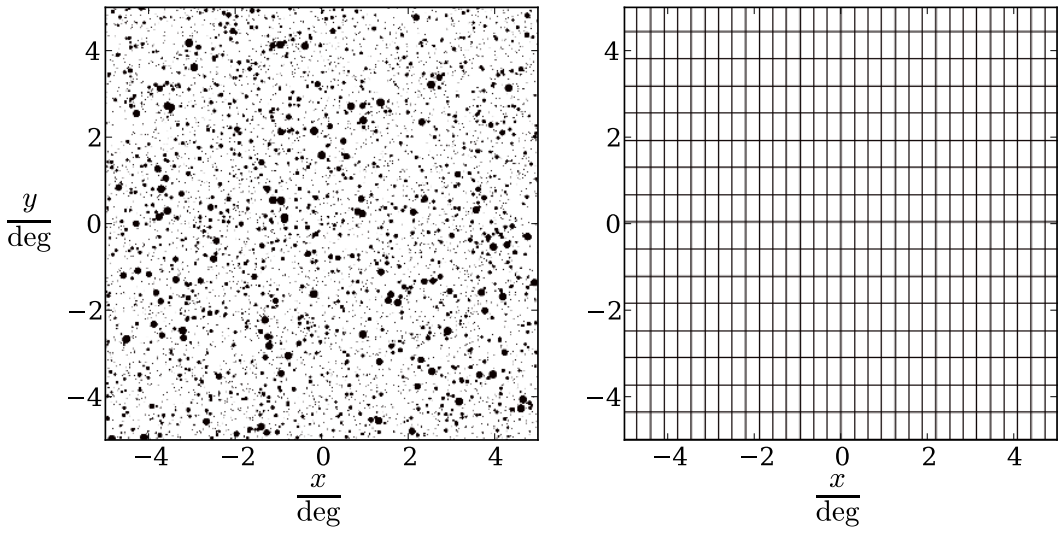

Figure 5.1: Star and Checkerboard masks. The star mask contains randomly positioned circles with random areas picked from three ranges, 2% from $[0.1', 0.5']$, 5% from $[1', 25']$ and 3% from $[15', 100']$ square arcminutes. The checkerboard mask mimics a CCD gap pattern. Three pixels are masked in the gaps. These two masks are also combined to simulate a more realistic senario.

between the masked and unmasked $\kappa_{\rm E,B}$,

$$\hat{\kappa}^{\rm ps}_{\rm E}(\boldsymbol{\ell}) = \int \frac{{\rm d}^2\ell'}{(2\pi)^2} \widehat{W}(\boldsymbol{\ell}-\boldsymbol{\ell}')[\hat{\kappa}_{\rm E}(\boldsymbol{\ell}')\cos 2\varphi_{\ell\ell'} + \hat{\kappa}_{\rm B}(\boldsymbol{\ell}')\sin 2\varphi_{\ell\ell'}] \,, \quad \hat{\kappa}^{\rm ps}_{\rm E(ap)}(\boldsymbol{\ell}) = \hat{S}(\boldsymbol{\ell})\hat{\kappa}^{\rm ps}_{\rm E}(\boldsymbol{\ell}) \,,$$
$$\hat{\kappa}^{\rm ps}_{\rm B}(\boldsymbol{\ell}) = \int \frac{{\rm d}^2\ell'}{(2\pi)^2} \widehat{W}(\boldsymbol{\ell}-\boldsymbol{\ell}')[\hat{\kappa}_{\rm B}(\boldsymbol{\ell}')\cos 2\varphi_{\ell\ell'} - \hat{\kappa}_{\rm E}(\boldsymbol{\ell}')\sin 2\varphi_{\ell\ell'}] \,, \quad \hat{\kappa}^{\rm ps}_{\rm B(ap)}(\boldsymbol{\ell}) = \hat{S}(\boldsymbol{\ell})\hat{\kappa}^{\rm ps}_{\rm B}(\boldsymbol{\ell}) \,, \tag{5.15}$$

where $\hat{\kappa}^{\rm ps}_{\rm E,B(ap)}(\boldsymbol{\ell})$ are the apodised and masked E-/B-mode κ. The above relations show that the mask affects the convergence in Fourier space by mixing some of the E-mode components into the B-modes and vice versa. Consequently, in order to utilize Fourier space information in cosmic shear analysis, the effects of the mask must be modelled.

The masks considered in this work are categorized into two groups: small and large scale masks. We also combine these masks to make complex mask set-ups. Fig. 5.1 shows the star and CCD pattern (checker-board) masks used throughout this work. The masks are zero padded to avoid periodic boundary conditions, to maintain realism. The left panel shows the star mask which contains randomly positioned circles with random areas picked from three ranges; 2% of the field is covered with starts from $[0.1', 0.5']$, 5% from $[1', 25']$ and 3% from $[15', 100']$ square arcminutes. The Checkerboard mask contains 3 dark pixels in the boundary between two chips. These masks contain only ones and zeros. The masked regions are shown in black. As can be seen in Fig. 5.1 the masks have sharp features which motivates smoothing

techniques to be applied on them.

5.1.2 Mask Modelling: The Mixing Matrix

In Fourier space the observable used for cosmic shear analysis is the convergence power spectrum. Measuring the power spectra from κ fields in Fourier space is a form of data compression, which assumes that the Universe is isotropic. As a result the C_ℓ only depend on the absolute value of the Fourier modes and not their angle. However, the mask in general is not isotropic. In order to apply it to a theory power spectrum, averages over its angular dependencies are taken and a mixing matrix is calculated.

The power spectrum is defined as

$$\left\langle |\hat{\kappa}_{\rm x}(\boldsymbol{\ell})\hat{\kappa}^*_{\rm y}(\boldsymbol{\ell}')| \right\rangle = (2\pi)^2 \delta_{\rm D}(\boldsymbol{\ell}-\boldsymbol{\ell}') C^{\rm xy}_\ell \ , \tag{5.16}$$

where x, y represent E, B, and the angle brackets denote an ensemble average. Therefore, C_ℓ is

$$\left\langle C^{\rm xy}_\ell \right\rangle = \frac{1}{A}\int \frac{{\rm d}\varphi_\ell}{2\pi} \left\langle |\hat{\kappa}_{\rm x}(\boldsymbol{\ell})\hat{\kappa}^*_{\rm y}(\boldsymbol{\ell})| \right\rangle \ , \tag{5.17}$$

where the integral is a simple angle averaging and we have replaced the Dirac delta function with its approximate value for a finite field,

$$\delta_{\rm D}(0) \approx \frac{A}{(2\pi)^2} \ , \tag{5.18}$$

where A is the area of the field. In practice we can only measure the pseudo-power spectrum, $\widetilde{C}_\ell$, because of the existing masks on the images. However, the cosmological models provide us with the power spectrum, C_ℓ. In the absence of noise the $\widetilde{C}_\ell$ is defined in the same way as C_ℓ in Eq. (5.17), by replacing $\hat{\kappa}^{\rm ps}(\boldsymbol{\ell})$ with the unmasked convergence, $\hat{\kappa}(\boldsymbol{\ell})$. We can find a relation between Cls and PCls by inserting for $\hat{\kappa}^{\rm ps}(\boldsymbol{\ell})$ from Eq. (5.15) into Eq. (5.17),

$$\begin{aligned}\left\langle \widetilde{C}^{\rm EE}_\ell \right\rangle = \frac{1}{A}\int \frac{{\rm d}\varphi_\ell}{2\pi} |\hat{S}(\boldsymbol{\ell})|^2 \int \frac{{\rm d}^2\ell'}{(2\pi)^2} \widetilde{W}(\boldsymbol{\ell}-\boldsymbol{\ell}') \int \frac{{\rm d}^2\ell''}{(2\pi)^2} \widetilde{W}^*(\boldsymbol{\ell}-\boldsymbol{\ell}'') \\ \times \left\langle |[\hat{\kappa}_{\rm E}(\boldsymbol{\ell}')\cos 2\varphi_{\ell\ell'} + \hat{\kappa}_{\rm B}(\boldsymbol{\ell}')\sin 2\varphi_{\ell\ell'}][\hat{\kappa}^*_{\rm E}(\boldsymbol{\ell}'')\cos 2\varphi_{\ell\ell''} + \hat{\kappa}^*_{\rm B}(\boldsymbol{\ell}'')\sin 2\varphi_{\ell\ell''}]| \right\rangle \ .\end{aligned} \tag{5.19}$$

The mask is not a variable between the realizations (that is assuming that there is no correlation between the mask and the underlying shear field), therefore, we can take $\widetilde{W}$ out of the ensemble averages. Moreover, choosing a symmetric smoothing kernel allows us to take $|\hat{S}(\boldsymbol{\ell})|^2$ out of the integral over φ_ℓ. Using Eq. (5.16) we link the PCls to the Cls,

$$\begin{aligned}\left\langle \widetilde{C}^{\rm EE}_\ell \right\rangle = \frac{|\hat{S}(\ell)|^2}{A}\int \frac{{\rm d}\varphi_\ell}{2\pi} \int \frac{{\rm d}^2\ell'}{(2\pi)^2} |\widetilde{W}(\boldsymbol{\ell}-\boldsymbol{\ell}')|^2 \\ \times [C^{\rm EE}_{\ell'}\cos^2 2\varphi_{\ell\ell'} + (C^{\rm EB}_{\ell'} + C^{\rm BE}_{\ell'})\sin 2\varphi_{\ell\ell'}\cos 2\varphi_{\ell\ell'} + C^{\rm BB}_{\ell'}\sin^2 2\varphi_{\ell\ell'}] \ .\end{aligned} \tag{5.20}$$

The above equation is written for the E-mode power spectrum although it can be extended to the other cases without much effort. $C^{\rm xy}_{\ell'}$ only depend on the length of $\boldsymbol{\ell}'$ and not its orientation. However, $|\widetilde{W}(\boldsymbol{\ell}-\boldsymbol{\ell}')|$ and the trigonometric functions in Eq. (5.20) depend on $\varphi_{\ell'}$, the polar angle of $\boldsymbol{\ell}'$. Therefore, the angle averaging part of the integrals in Eq. (5.20) can be taken independent of the cosmological model (see Memari 2009, for details of the calculations). The masking effect is hence modelled in the form of a mode mixing matrix, $\mathscr{M}$,

$$\mathscr{M}_{\ell\ell'} \equiv \frac{|S(\ell)|^2}{(2\pi)^2 A}\int_0^\pi \mathrm{d}\eta\; W_{\gamma\gamma}(L) M_\eta \;, \tag{5.21}$$

where η is the angle between $\boldsymbol{\ell}$ and $\boldsymbol{\ell}'$, and $L = |\boldsymbol{\ell}-\boldsymbol{\ell}'|$. $W_{\gamma\gamma}(L)$ is the angle average of the square of the mask in Fourier space,

$$W_{\gamma\gamma}(L) \equiv \int \frac{\mathrm{d}\varphi_L}{2\pi}|\widetilde{W}_\gamma(\boldsymbol{L})|^2, \tag{5.22}$$

and

$$M_\eta \equiv \begin{pmatrix} 1+\cos 4\eta & 1-\cos 4\eta & 0 \\ 1-\cos 4\eta & 1+\cos 4\eta & 0 \\ 0 & 0 & 2\cos 4\eta \end{pmatrix}. \tag{5.23}$$

As a result we can write

$$\widetilde{\boldsymbol{C}}_\ell = \int_0^\infty \mathrm{d}\ell'\,\ell'\,\mathscr{M}_{\ell\ell'}\boldsymbol{C}_{\ell'} \;, \tag{5.24}$$

where

$$\boldsymbol{C} \equiv (C_\ell^{\rm EE}, C_\ell^{\rm BB}, C_\ell^{\rm EB})^t \;. \tag{5.25}$$

In practice we need to change all the integrals in the above equations into discrete finite sums. Doing so can cause discrepancies between the measured PCls and the observed ones, which should be corrected for. We define a new mixing matrix $\mathbf{M}$ which satisfies this matrix relation,

$$\widetilde{\boldsymbol{C}}_\ell = \mathbf{M}_{\ell\ell'}\boldsymbol{C}_{\ell'} \;, \tag{5.26}$$

where Einstein summation rules apply. $\mathbf{M}$ is the discrete form of $\mathscr{M}$ with $\mathrm{d}\ell'\ell'$ absorbed in its definition. Consequently, we can write the inverse of Eq. (5.26) and recover the power spectrum,

$$\boldsymbol{C}_\ell^{\rm rec} = (\mathbf{M}^{-1})_{\ell\ell'}\widetilde{\boldsymbol{C}}_{\ell'} \;. \tag{5.27}$$

A prominent source of noise in weak lensing analysis is the galaxy shape noise which is modelled as a Gaussian random noise with zero mean and σ_ϵ dispersion. The mask affects the noise in the same way as the shear field. We can write the noise as a separate source of power with no ℓ dependence,

$$N_\epsilon = \frac{\sigma_\epsilon^2}{2n_{\rm gal}} \;, \tag{5.28}$$

where $n_{\rm gal} = 30$ per square arcminutes is the mean number density of galaxies and $\sigma_\epsilon = 0.3$ is the intrinsic dispersion of galaxy ellipticities. As a result the measured PCl is,

$$\widetilde{\boldsymbol{C}}_\ell = \mathbf{M}_{\ell\ell'}[\boldsymbol{C}_{\ell'} + N_\epsilon] \,, \tag{5.29}$$

and the recovered Cl is,

$$\boldsymbol{C}_\ell^{\rm rec} = (\mathbf{M}^{-1})_{\ell\ell'}\widetilde{\boldsymbol{C}}_{\ell'} - N_\epsilon \,. \tag{5.30}$$

Estimating the Mixing Matrix

The angle averaging of a pixelated field is not exact. To take the angular average over such a field an annulus is chosen and the average is taken over the value of all of the pixels with centres lying inside it. The Fourier mode, ℓ, that corresponds to this estimated value is also calculated by averaging over the value of the $|\boldsymbol{\ell}|$ modes that lie in this annulus. To estimate $W_{\gamma\gamma}$ from Eq. (5.22), I use annuli of width equal to the smallest Fourier mode, $\ell_{\rm min}$, available in field,

$$\ell_{\rm min} = \frac{2\pi}{D}, \tag{5.31}$$

where D (in radian) is the side length of the square field after zero padding. Then the estimated $W_{\gamma\gamma}$ values are fed into Eq. (5.21). Since $W_{\gamma\gamma}$ is only estimated for discrete values, the integral in Eq. (5.21) needs to be transformed into a sum over these values of L. This integral is taken over η which is calculated for each L, ℓ and ℓ' from,

$$\cos\eta = \frac{\ell^2 + \ell'^2 - L^2}{2\ell\ell'} \,. \tag{5.32}$$

The values of ℓ and ℓ' depend on the binning scheme used. Note that in Eq. (5.21), the available η will not form a regular grid, therefore, dη will not be a constant. To find the PCls from theory, first $\mathbf{M}_{\ell\ell'}$ is estimated for the smallest binning (also used for estimating $W_{\gamma\gamma}$), which is then applied to the theory power spectrum and noise. To find a better agreement between the theory and the estimations, the theory C_ℓ is first laid on a similarly pixelated field and angle averaged for the smallest bins in ℓ, as explained above for $W_{\gamma\gamma}$. The theory $\widetilde{C}_\ell = \mathbf{M}_{\ell\ell'}C_{\ell'}$, can then be re-binned as desired.

To recover the C_ℓ using the estimated PCls from the masked fields, a different approach has to be taken. In this case the ℓ binning should be done prior to applying the mixing matrix. The matrix is then inverted and applied to a binned $\widetilde{C}_\ell$, using Eq. (5.30). In this thesis I will use linear binning everywhere with different widths for the annuli, from $\ell_{\rm min}$ to $20 \times \ell_{\rm min}$. This allows an investigation into the effects of binning. The binning is quantized by an integer parameter $n_\ell \in [1, 20]$ which corresponds to the number of merged original annuli to make the wider annulus.

5.1.3 Mask Smoothing: Apodisation

In this thesis, I will use three Gaussian smoothing kernels to apodise the masks using the second method (smoothing the edges of $W(\boldsymbol{\vartheta})$ before applying it to the shear fields). The advantage of this method is that it allows slower variations for the integrands in Eq. (5.21) and Eq. (5.22) which makes their discrete approximation more accurate. As can be seen in Eq. (5.21) in the case of the first apodisation method, the smoothing kernel only comes into play after the angular averages in Eq. (5.21) and Eq. (5.22) have been taken.

Fig. 5.2 shows the general form of the smoothing kernels which act on the edges of the masked regions. Note that using this method increases the effective masked area, since the fully masked regions will remain the same, while their edges will have a smooth transition from zero to one, which is determined by the size of the kernel. In Fig. 5.2, N is an odd number which determines the size of the kernel. The dispersion of this Gaussian is equal to $(N-1)/1.5$. In general, any smoothing function can be chosen as the kernel. In this thesis I will use three sizes for the Gaussian kernels,

$$N = 5, 11, 23 \text{ pixels} . \tag{5.33}$$

The main results will be shown and compared for the original masks and the three apodisation schemes.

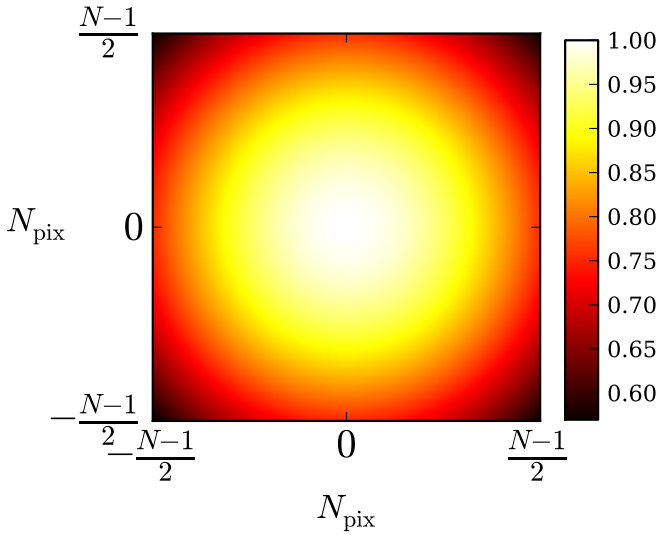

Figure 5.2: The general form of the smoothing kernel used for aposiding the masked shear fields. The total number of pixels used, N, determines the support size of the smoothing Kernel. The kernel is symmetric, hence N is an odd number. The functional form of the smoothing kernel is a Gaussian with zero mean and $\sigma = (N-1)/1.5$.

Fig. 5.3 shows the mixing matrix for the composite mask (star and checkerboard). The left panel shows the matrix for the original ones and zeros mask, while the right panel shows the same for an apodised mask. The apodisation used here is with $N = 11$. Since the EB-EB part of the mixing matrix is independent of the EE and BB parts, it is not shown here and will not be used in any of the analysis. The mixing matrices are shown for the largest binning used

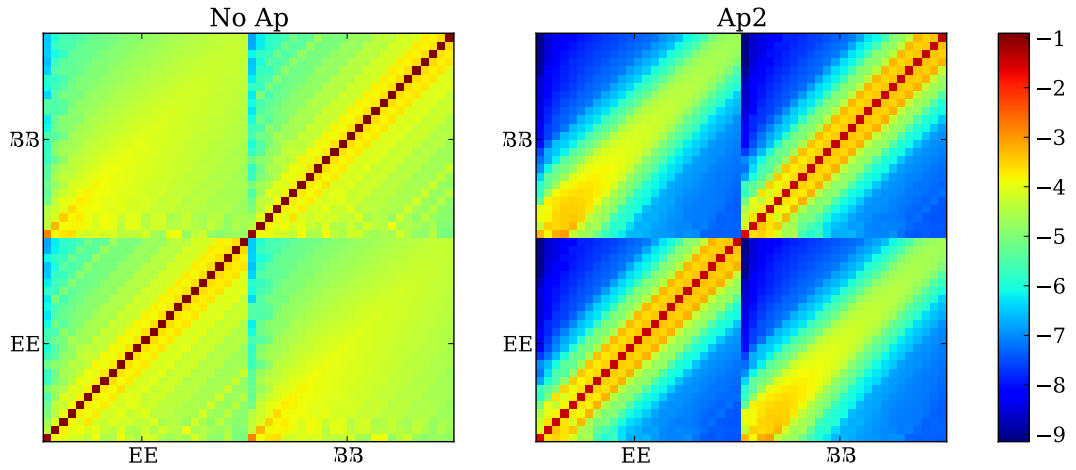

Figure 5.3: The logarithm of the absolute value of the mixing matrices for star and checkerboard mask. The left panel shows the mixing matrix for the original mask with no apodisation, whereas the right panels shows the same for an apodised mask with $N = 11$ pixels smoothing scale. 25 linear ℓ bins in $[245, 8830]$ are considered here, which corresponds to the widest binning that will be used in this thesis.

in this thesis ($n_\ell = 20$, $\Delta\ell = n_\ell \times \ell_{\rm min}$). The matrices are plotted in terms of the logarithm of the absolute value of their elements. As can be seen in this figure a smooth mask has a more diagonal mixing matrix and a smoother off diagonal behaviour. The importance of this property of the mixing matrix will become clear in the next sections.

5.2 Theory vs Simulations

To test the mask modelling I use two sets of simulations: random realisations of Gaussian and lognormal shear fields (provided by Benjamin Joachimi). The input power spectrum is identical for both cases and is based on a cold dark matter Universe with a dominant dark energy component, with cosmological parameters given in Table. 5.1.

The linear power spectrum is determined assuming a primordial power law power spectrum with Bond & Efstathiou (1984) transfer function. Additionally, the halo fit formula of Smith et al. (2003) is used for calculating the non-linear scales.

All the simulations are originally made for a larger field ($20° \times 20°$, 2048×2048 pixels), then a $10° \times 10°$ field (1024×1024 pixels) is cut out of the middle, to simulate the non-periodic nature of the Universe. 100 random simulations are generated for each case in the analysis. After adding a Gaussian random shape noise with $\sigma_\epsilon = 0.3$ to the shear fields, they are masked, and then zero-padded before the Fourier transform. The zero padding scheme used here, changes the size of the fields to their original size, which means, doubling the size of the field on each side by adding zeros. The zero padding ensures that a periodic boundary

σ_8	Ω_m	Ω_Λ	w_0	n_s	h	Ω_b
0.8	0.27	0.73	−1.0	0.96	0.72	0.045

Table 5.1: The fiducial cosmological parameters consistent with Planck results. The normalization of the power spectrum, σ_8, is the standard deviation of perturbations in a sphere of radius $8h^{-1}$Mpc today. Ω_m, Ω_Λ, and Ω_b are the matter, the dark energy and the baryonic matter density parameters, respectively. w_0 is the dark energy equation of state parameter, which is equal to the ratio of dark energy pressure to its density. The spectral index, n_s, is the power of the initial power spectrum. The dimensionless Hubble parameter, h, characterizes the rate of expansion today.

condition is not assumed for the field when a fast Fourier transform (FFT) is applied to it. Zero-padding the field more than this results in a computationally more expensive analysis while the result remains similar. The FFT of a zero padded field has a higher resolution, hence zero-padding is also a non-unique form of interpolation between the Fourier modes. As a result the resolution of the Fourier transformed fields, $\ell_{\rm min}$ (defined in Eq. 5.31), is equal to 18.

The convergence, $\hat{\kappa}^{\rm ps}(\boldsymbol{\ell})$, is estimated from the Fourier transformed shear field, from which $\widetilde{C}_\ell$ is calculated using Eq. (5.17). The angle averaging involved here is not exact, since the fields are pixelized. The pixelization effect on angle averaging is more prominent for the low ℓ values, where the number of pixels in an annulus is small. This effect is one of the main sources of discrepancies between the modelled and the measured PCls. Having a smooth mask results in a slowly varying convergence in Fourier space which in turn reduces the differences between the model and the measured quantities.

Fig. 5.4 shows the average estimated and theory PCls and Cls of the lognormal fields for the composite mask, with $n_\ell = 10$ (see Sect. 5.1.2 for the details of ℓ binning and the definition of n_ℓ). The top plot shows the results for the original star and checkerboard mask, while the bottom plot shows the same for an apodised version of this mask with the medium sized kernel. The curves show the expected values of the PCls and Cls from theory, whereas the markers show their estimated values. The dashed line shows the value of noise power spectrum given in Eq. (5.28). The theory B-mode C_ℓ is zero, however, due to incompleteness the recovered power spectrum has a non-zero B-mode contribution. The recovered C_ℓ underestimates the power on large scales while overestimating it on small scales. However, the mid-range scales show the best agreement between the theory and recovered C_ℓ. Apodisation pushes the inconsistency between the theory and $C_\ell^{\rm rec}$ to larger ℓ values and also decreases the B-modes for $3000 \lesssim \ell \lesssim 7000$. The PCls are dominated by noise on small scales, hence, $\widetilde{C}_\ell^{\rm E} \simeq \widetilde{C}_\ell^{\rm B}$ for large ℓ. The error-bars shown here are estimated from the field-to-field variance of the simulated fields, for the mean value.

As we have seen in Sect. 5.1.2 the mixing matrices calculated from a pixelated mask are not accurate. The inaccuracy in mask modelling is more severe for small angular scales which will propagate to all scales (see Eq. 5.21). As a result, a constant multiplicative bias will always be present in the mixing matrix. This bias is shown in the panels of Fig. 5.4. $B_{\rm CL}$ and $B_{\rm PCL}$ are the

respective multiplicative biases for C_ℓ and $\widetilde{C}_\ell$. These biases are not equal because the angular averaging which results in the binned C_ℓ and $\widetilde{C}_\ell$ are carried out at different stages of their calculation (as was explained in Sect. 5.1.2), hence the bias at low ℓ is propagated differently.

I have tested several different methods to estimate the mixing matrix and concluded that the only way to systematically tackle this challenge is to find this constant bias by employing simulations, as was done here. Note that while changing the integration scheme or the number of zero pad pixels, may alleviate this problem for certain masks it will not be applicable to the others. For example, for a spherically symmetric mask, the best method to estimate the angular averages is to take averages over all the pixels with the exact distance to the middle of the field (in Fourier space). Using this method results in a smaller number of points for each ℓ mode, which will in turn result in a very inaccurate mask modelling, for asymmetric masks. In conclusion, the method used in this thesis is the most robust approach to mask modelling for a flat sky analysis.

Since the variance on the mean of 100 fields is very small, it is difficult to compare the theory to estimated $\widetilde{C}_\ell$ or C_ℓ values in Fig. 5.4. Therefore, the relative power spectra are plotted in Fig. 5.5 and Fig. 5.6. The top plot in these figures show the ratio of the power spectra for the original star and checkerboard mask and the rest of the plots show the same for the apodised masks with the kernels discussed in Sect. 5.1.3. The gray areas show the cosmic variance for the simulated fields. The cosmic variance for C_ℓ assuming a Gaussian distribution is estimated using,

$$\sigma^2_{\text{cosm, Cl}} = \frac{2}{2\ell+1}\frac{1}{f_{\text{sky}}\Delta\ell}(C_\ell + N_\epsilon)^2 \,, \tag{5.34}$$

where f_{sky} is the fraction of the sky that is not covered by any masks. $\Delta\ell$ is the ℓ-bin width and N_ϵ is the noise power. For the 100 simulated fields of 100 square degrees each,

$$f_{\text{sky}} = \frac{100 * 100 * (\pi/180)^2}{4\pi}\, f_{\text{image}} \simeq 0.24\, f_{\text{image}} \,, \tag{5.35}$$

where f_{image} is the fraction of each image covered by the mask. To find the cosmic variance for $\widetilde{C}_\ell$ we need to use the mixing matrix on the $\langle \Delta C^2_\ell \rangle$. Doing so results in some off diagonal terms which we are not interested in. The diagonal terms are

$$\sigma^2_{\text{cosm, PCl}} = \frac{2}{2\ell+1}\frac{1}{f_{\text{sky}}\Delta\ell}\widetilde{C}^2_\ell \,, \tag{5.36}$$

which is the cosmic variance term for the PCls, where $\widetilde{C}_\ell$ is their expected value from theory.

Fig. 5.5 which depicts the ratios for PCls, shows a poor agreement between the theory and estimated values for small ℓ. This is not improved by apodisation. Smoothing the mask changes the value of the constant multiplicative bias, B_{PCL}, at the cost of decreasing the observed effective fraction of the image, f_{image}. On small scales ($\lesssim$ 2500) the fluctuations

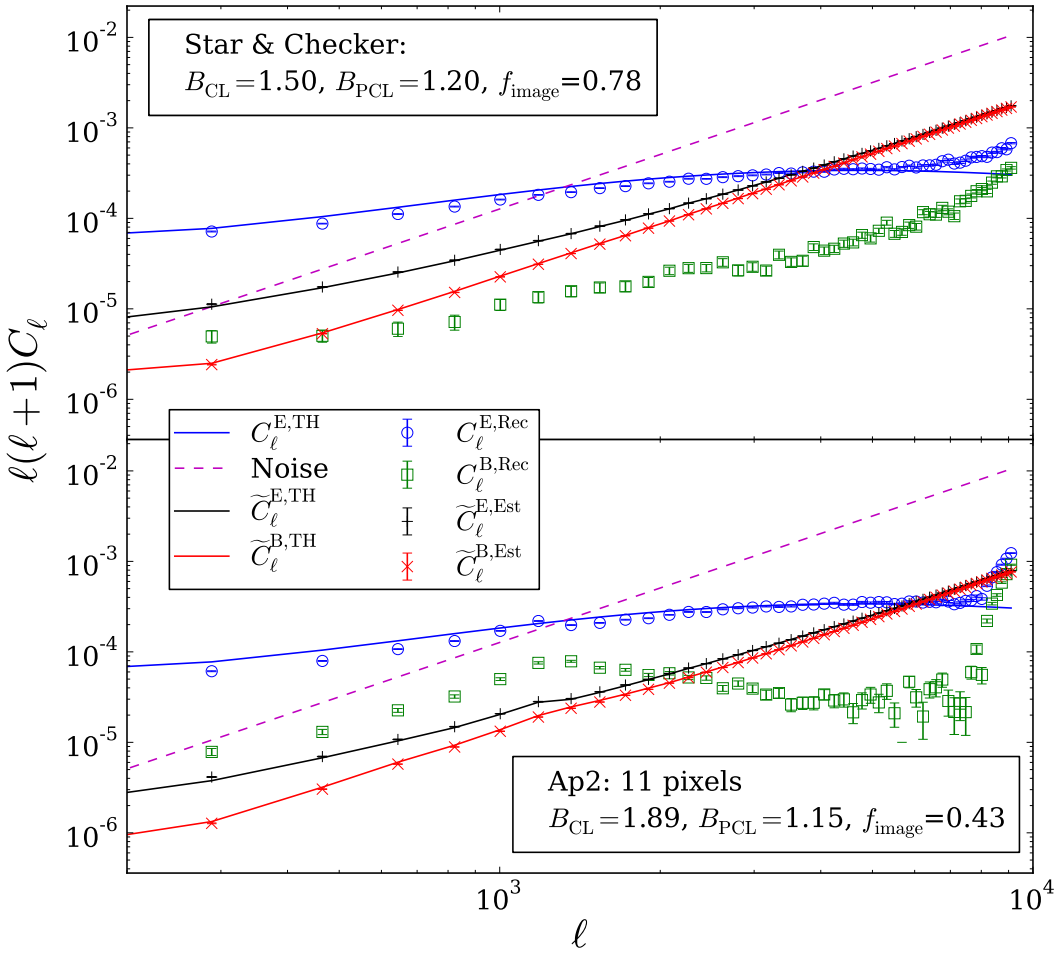

Figure 5.4: The estimated power spectra from the lognormal simulations versus their expected values from theory for the composite mask. Ten of the smallest ℓ bins are merged to make these power spectra ($n_\ell = 10$, $\Delta\ell \approx 180$). The solid curves show the expected theory values and the symbols show the estimated values from the simulations. The blue circles show the recovered C_ℓ after subtracting the noise contribution. The green squares show the remaining B-mode contribution in the recovered C_ℓ. The theory value of the noise power spectrum is shown by the dashed magenta line. The black and red curves show the theory E-/B-mode pseudo power spectra, respectively. The × and + symbols show their estimated values. The error bars correspond to field-to-field variations between the realisations of the shear fields. The top panel shows the results for the original mask, while for the bottom panel the mask is apodised with a kernel of size 11 pixels. $B_{\rm CL}$ and $B_{\rm PCL}$ show the value of the constant multiplicative bias of the mixing matrix, for the C_ℓ and $\widetilde{C}_\ell$, respectively. This bias is corrected for using the simulations. The effective relative uncovered area of the simulated images is shown via $f_{\rm image}$. The apodisation covers a larger area of the field.

of the PCls are within the cosmic variance. Nevertheless, the large scale dependencies are significantly beyond the amount allowed by cosmic variance. On small scales the theory underestimates the value of E-modes while overestimates the B-mode contribution. On very small scales as we saw for Fig. 5.4 the PCls are dominated by noise, which means that in Fig. 5.5 we are comparing noise with noise for these scales and unsurprisingly get a good agreement.

These discrepancies are even more pronounced in Fig. 5.6. The recovered Cl, $C_\ell^{\rm rec}$, underestimates the large scale values. The black points in this plot show the values of the ratios of $C_\ell^{\rm rec}$ without noise correction to $C_\ell + N_\epsilon$ with the cosmic variance region shown in dark gray, while the red points show the ratio for a noise corrected $C_\ell^{\rm rec}$ to C_ℓ with the light gray area for its corresponding cosmic variance. Consequently, the errors on the red points are larger and they diverge for large ℓ where the noise contribution is significantly larger than the shear power spectrum.

By comparing Fig. 5.5 and Fig. 5.6 we see that in general the forward modelling shows a better agreement between the theory and estimated values. Therefore, we expect a smaller bias in the cosmological models using $\widetilde{C}_\ell$ values as observables.

5.3 Error Propagation

The ultimate goal of a weak lensing analysis is to constrain cosmological models and their parameters, in a typical scenario. In Sect. 2.1.2 we saw how to use Fisher matrices to translate the observed quantities and their covariance to constraints on model parameters. Furthermore, we saw how to estimated the bias in measured parameters, to linear order, using a similar formalism. Here I will use a Fisher analysis to put upper limits on the constraining power of PCls. I will use both $\widetilde{C}_\ell$ (forward modelling) and C_ℓ (backward modelling) to compare the constraints and biases on model parameters. In this thesis I use lognormal shear fields as default and will compare the final results with Gaussian shear fields. The covariance of a Gaussian field has been calculated in the literature. Here I will show a general calculation for finding the moments of a lognormal field, which is then used to find the covariance of the shear power spectra.

5.3.1 Lognormal Moments

We can find the moments of a lognormal field using its relation to a Gaussian field. These moments will be used to calculate the covariance matrix of the power spectrum of the lognormal shear fields, which will then be used to estimate the Fisher matrices. In this section I will show how all the moments of a lognormal field can be written in terms of its power spectrum, by taking the following steps.

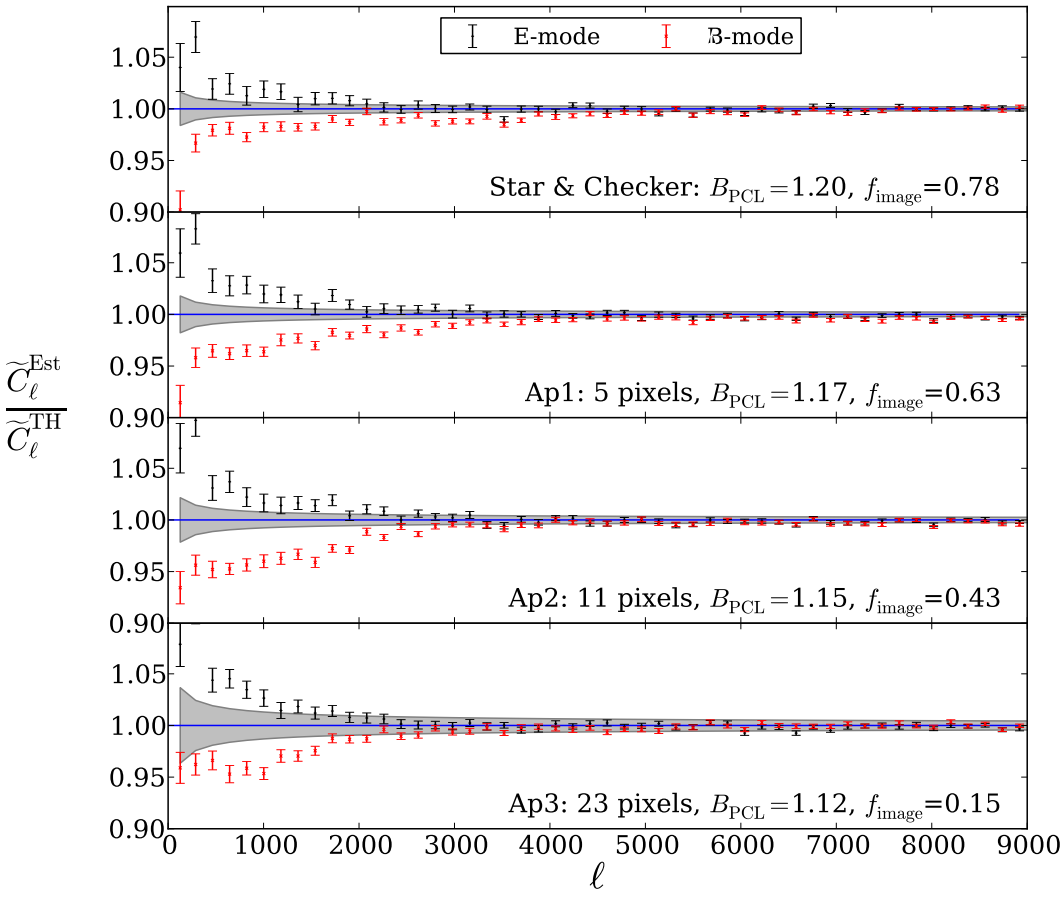

Figure 5.5: The ratio of estimated to theory PCls. The top plot shows this ratio for the original star and checkerboard mask which has sharp edges, while the bottom plots shows the same for smoothed masks, with different kernel sizes. The black point show the ratio for the E-mode $\widetilde{C}_\ell$, whereas the red ones correspond to the B-modes. The blue solid line shows where a perfect correspondence between the theory and the estimated quantities from the simulations would lie. The gray area shows the expected cosmic variance.

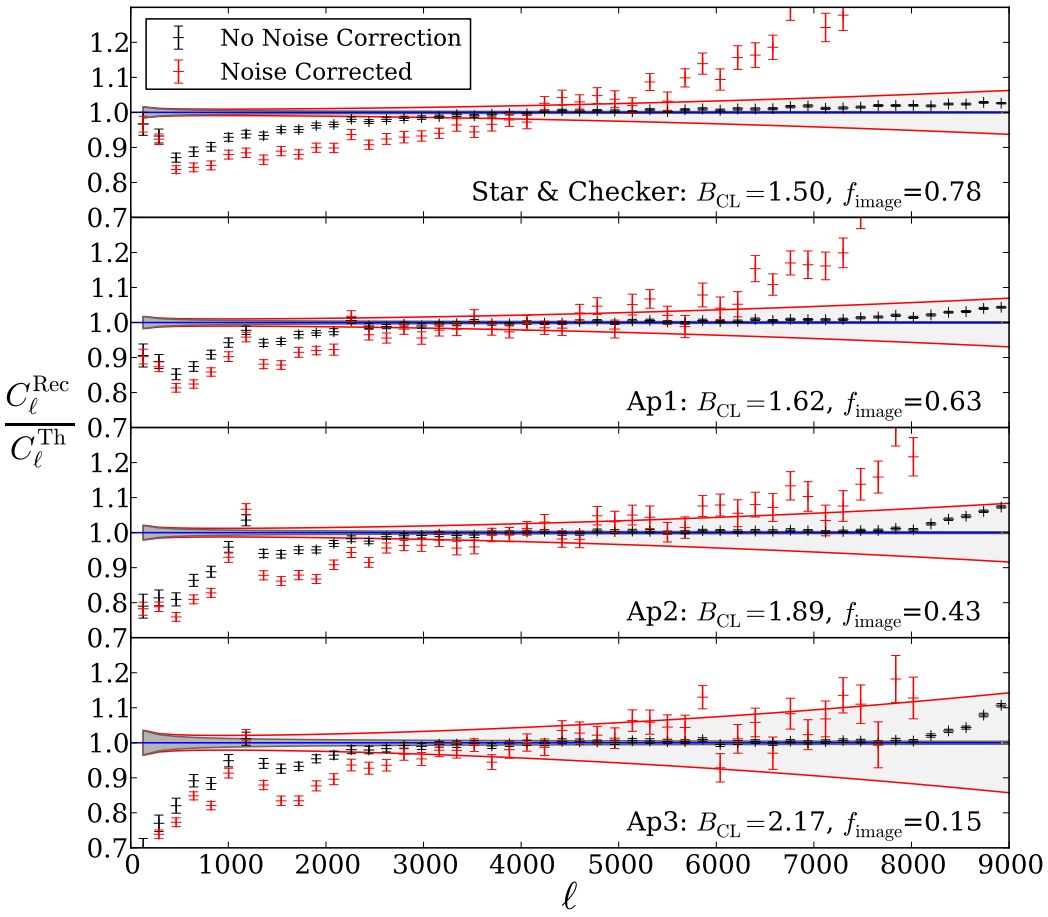

Figure 5.6: The ratio of recovered to theory E-mode C_ℓ. The top plot shows this ratio for the original composite mask, while the bottom plots shows the same for smoothed masks, with different kernel sizes. The black point show the ratio for the C_ℓ before noise correction, and the red points show this ratio after noise correction. The blue solid line shows where a perfect correspondence between the theory and the estimated quantities from the simulations would lie. The light gray area with red edges show the expected cosmic variance for the noise corrected C_ℓ ratio, while the dark gray area shows the cosmic variance before noise correction.

A lognormal field, $\delta^{\rm ln}$, with zero mean is defined with respect to a Gaussian field δ as

$$\delta^{\rm ln}(\mathbf{x}) \equiv e^{\delta(\mathbf{x})-\sigma^2/2} - 1 \ , \tag{5.37}$$

where σ^2 is the variance of the Gaussian field. This implies that the variance of the lognormal field is $\exp\sigma^2$. To find the moments of this lognormal field, we take the following steps. In Fourier space the Nth lognormal moment for a non-zero $\mathbf{k}_i$ can be written as

$$\Big\langle \prod_i^N \hat{\delta}^{\rm ln}(\mathbf{k}_i) \Big\rangle = \Big\langle \prod_i^N [(2\pi)^n \delta_{\rm D}(\mathbf{k}_i) + \hat{\delta}^{\rm ln}(\mathbf{k}_i)] \Big\rangle \ , \tag{5.38}$$

where $\delta_{\rm D}(\mathbf{k}_i)$ is the Dirac delta function. Note that $\mathbf{k}_i$ has n dimensions and $\delta(\mathbf{k}_i)$ is a one dimensional quantity on a multidimensional grid. Next we write the lognormal moments with respect to their real space counterparts,

$$\Big\langle \prod_i^N [(2\pi)^n \delta_{\rm D}(\mathbf{k}_i) + \hat{\delta}^{\rm ln}(\mathbf{k}_i)] \Big\rangle = \Big\langle \prod_i^N \int {\rm d}\mathbf{x}_i e^{-i\mathbf{k}_i.\mathbf{x}_i} [1 + \delta_i^{\rm ln}] \Big\rangle \ , \tag{5.39}$$

where $\delta_i^{\rm ln} \equiv \delta^{\rm ln}(\mathbf{x}_i)$. We can take the ensemble average inside the integral and rewrite the above equation as,

$$\Big\langle \prod_i^N \hat{\delta}^{\rm ln}(\mathbf{k}_i) \Big\rangle = \prod_i^N \Big[\int {\rm d}\mathbf{x}_i e^{-i\mathbf{k}_i.\mathbf{x}_i} \Big] \Big\langle \prod_i^N [1 + \delta_i^{\rm ln}] \Big\rangle \ . \tag{5.40}$$

The ensemble average, $\langle \prod_i^N [1 + \delta_i^{\rm ln}] \rangle$, can be expressed in terms of the two point correlation functions, by writing the lognormal fields in terms of their Gaussian generators from Eq. (5.38),

$$\Big\langle \prod_i^N (1 + \delta_i^{\rm ln}) \Big\rangle = \int_{-\infty}^{+\infty} {\rm d}\boldsymbol{\delta} \ p(\boldsymbol{\delta}) \ \prod_i^N \Big[e^{\delta_i} \ e^{-\sigma_i^2/2} \Big] \ , \tag{5.41}$$

where $\boldsymbol{\delta} \equiv (\delta_1, \delta_2, ..., \delta_N)$ and

$$p(\boldsymbol{\delta}) = \frac{e^{-\boldsymbol{\delta}\ \mathbf{C}^{-1}\boldsymbol{\delta}^t/2}}{\sqrt{(2\pi)^N \det \mathbf{C}}} \ , \tag{5.42}$$

is the multivariate Gaussian distributed probability of $\boldsymbol{\delta}$ with $\mathbf{C}$ as the covariance.

Since $\mathbf{C}$ is a covariance, i.e. symmetric and positive definite, we can write it in terms of its eigenvalues and eigenvectors as

$$\mathbf{C} = \mathbf{O}\mathbf{D}\mathbf{O}^{\rm t} \ , \tag{5.43}$$

where $\mathbf{O}$ is the orthogonal matrix made out of the eigenvectors of $\mathbf{C}$ and $\mathbf{D}$ is a diagonal matrix of the eigenvalues of $\mathbf{C}$. As a result

$$\mathbf{C}^{-1} = \mathbf{O}\mathbf{D}^{-1}\mathbf{O}^{\rm t} \ . \tag{5.44}$$

Substituting for $\mathbf{C}^{-1}$ in Eq. (5.41) and defining $X \equiv \boldsymbol{\delta}\mathbf{O}$, yields,

$$\Big\langle \prod_i^N (1+\delta_i^{\rm ln}) \Big\rangle = \frac{\prod_i^N e^{-\sigma_i^2/2}}{\sqrt{(2\pi)^N \det \mathbf{C}}} \int_{-\infty}^{+\infty} {\rm d}X \, e^{-X\mathbf{D}^{-1}X^t/2} \times \prod_i^N \Big[e^{X_j O^{\rm t}_{ji}} \Big] \,, \tag{5.45}$$

where $|\det \mathbf{O}| = 1$ was used to simplify the result. Rewriting Eq. (5.45) in terms of its components results in,

$$\Big\langle \prod_i^N (1+\delta_i^{\rm ln}) \Big\rangle = \frac{\prod_i^N e^{-\sigma_i^2/2}}{\sqrt{(2\pi)^N \det \mathbf{C}}} \times \prod_j^N \Big[\int_{-\infty}^{+\infty} {\rm d}X_j \, e^{-X_j^2 D_{jj}^{-1}/2 + X_j \sum_i O_{ij}} \Big] \,. \tag{5.46}$$

With the aid of another variable change, $Y_i = X_i \sqrt{D_i^{-1}}$, and completing the square we solve this integral and find the desired relationship,

$$\Big\langle \prod_i^N (1+\delta_i^{\rm ln}) \Big\rangle = \exp\left(\frac{\sum_{ij} C_{ij} - \sigma_i^2}{2} \right) = \exp\left(\sum_{i<j} C_{ij} \right). \tag{5.47}$$

Inserting for $\langle \prod_i^N \hat{\delta}^{\rm ln}(k_i) \rangle$ from the above equation into Eq. (5.40) results in,

$$\langle \prod_i^N \hat{\delta}^{\rm ln}(k_i) \rangle = \prod_i^N \int {\rm d}x_i e^{-ik_i.x_i} \prod_{i<j}^N e^{C_{ij}} \,. \tag{5.48}$$

The covariance of the Gaussian and lognormal fields are related via,

$$e^{C_{ij}} = 1 + \langle \delta_i^{\rm ln} \delta_j^{\rm ln} \rangle = 1 + \xi_{ij}^{\rm ln} \,, \tag{5.49}$$

where $\xi_{ij}^{\rm ln}$ is the correlation between $\delta_i^{\rm ln}$ and $\delta_j^{\rm ln}$. Consequently, we can write the lognormal moments in Eq. (5.48) in terms of their two point correlation functions or alternatively their power spectra,

$$\langle \prod_i^N \hat{\delta}^{\rm ln}(k_i) \rangle = \frac{1}{(2\pi)^{Mn}} \prod_i^N \int {\rm d}x_i e^{-ik_i.x_i} \prod_{j>i;m}^{N;M} \int {\rm d}l_m e^{il_m.(x_i - x_j)} [(2\pi)^n \delta_{\rm D} + P](l_m) \,, \tag{5.50}$$

where $P(l)$ is the power spectrum of the lognormal field, n is the dimension of the field (for gravitational lensing $n = 2$), the subscript m belongs to each pair of $\delta_i^{\rm ln}$ and $\delta_j^{\rm ln}$ which make $\langle \delta_i^{\rm ln} \delta_j^{\rm ln} \rangle = \xi_{ij}^{\rm ln}$ and $M = N(N-1)/2$. The above integrals can be simplified by integrating with respect to x_i, since $[(2\pi)^n \delta_{\rm D} + P](l_m)$ have no dependency on x_i. The x_i integrals will result in N delta functions of dimension n which depend on k_i and l_m. There are M, l_m integrals and 2^M, $[(2\pi)^n \delta_{\rm D}(l_m) + P(l_m)]$ combinations. Writing the Delta functions found from the x_i integrals in

the following form,

$$\begin{aligned}\delta_i^{\bar{\sum} j - \bar{\sum} k} &\equiv \delta_i^{+j+j'+...-k-k'-...} \\ &\equiv \delta_{\rm D}(k_i + l_j + l_{j'} + ... - l_k - l_{k'} - ...) \,,\end{aligned} \tag{5.51}$$

will simplify the notation. Note that $\bar{\sum}$ is not a real sum. We find that the two sums over the positive and negative l_m modes can be formulated as follows,

$$\begin{aligned}\bar{\sum} j &= \bar{\sum}_{r=1}^{i-1}(r-1)N + i - r(r+1)/2 \\ -\bar{\sum} k &= -\bar{\sum}_{r=i}^{N-1}(i-1)N - i(i-1)/2 + r - i + 1 \,,\end{aligned} \tag{5.52}$$

for a given N and i. We are now left with M integrals with 2^M components for each,

$$\begin{aligned}\langle \prod_i^N \hat{\delta}^{\rm ln}(k_i)\rangle = \frac{1}{(2\pi)^{n(M-N)}} \prod_{j>i;m}^{N;M} \int {\rm d}l_m [(2\pi)^n \delta_{\rm D} + P](l_m) \\ \times \prod_i^N \delta_i^{+\bar{\sum}_{r=1}^{i-1}(r-1)N+i-r(r+1)/2-\bar{\sum}_{r=i}^{N-1}(i-1)N-i(i-1)/2+r-i+1} \,.\end{aligned} \tag{5.53}$$

The remaining M integrals over l_m can be simplified using the N delta functions and the delta functions in the 2^M combinations of $\delta_{\rm D}(l_m)$ and $P(l_m)$. Some of these integral vanish after considering the delta functions. In any integral if we come about a $\delta_{\rm D}(k_i)$ then that term is equal to zero since we are not interested in $k_i = 0$ terms and for the rest of the values the delta function vanishes.

We can immediately see that for the third moment $N = M = 3$, i.e. only one integral will remain after the simplifications and the rest of the term will either vanish or are products of power spectra and Delta functions which depend on several k_i modes. I have developed an algorithm which can simplify the moments for any given N.

The fourth moment of the lognormal fields are essential for calculating the covariance of their power spectra. Therefore, here the results for the fourth order moment will be explicitly shown. The fourth moment has many terms. These terms can be divided into four groups, depending on the number of remaining integrals over the power spectra and an extra group which contains the Gaussian only contribution. Hence, in the following each group will be represented separately. The fourth lognormal moment in Fourier space can be written as,

$$\langle \hat{\delta}^{\rm ln}(k_1)\hat{\delta}^{\rm ln}(k_2)\hat{\delta}^{\rm ln}(k_3)\hat{\delta}^{\rm ln}(k_4)\rangle = (2\pi)^n \delta_{\rm D}(k_1+k_2+k_3+k_4)\{I + II + III + IV\} + G \,, \tag{5.54}$$

where G is the pure Gaussian term,

$$\begin{aligned} G =&(2\pi)^{2n}\delta_{\rm D}(k_2+k_3)\delta_{\rm D}(k_1+k_4)P(k_1)P(k_2) \\ &+(2\pi)^{2n}\delta_{\rm D}(k_1+k_3)\delta_{\rm D}(k_2+k_4)P(k_1)P(k_2) \\ &+(2\pi)^{2n}\delta_{\rm D}(k_1+k_2)\delta_{\rm D}(k_3+k_4)P(k_1)P(k_3)\ , \end{aligned} \tag{5.55}$$

and I, II, III and IV are the pure lognormal terms, shown bellow. The highest number of integrals remaining after the simplifications is three. There is only a single term of this form,

$$I = \int {\rm d}l_4{\rm d}l_5{\rm d}l_6 P(l_4)P(l_5)P(l_6)P(l_4+l_5-k_2)P(l_6-l_4-k_3)P(k_4+l_5+l_6)\ . \tag{5.56}$$

There are six terms with two integrals, which can be factorized as,

$$\begin{aligned} II = &\int {\rm d}l_5{\rm d}l_6 P(l_5)P(l_6)P(k_4+l_5+l_6)\big[P(l_5+l_6-k_2-k_3)P(l_5-k_2) \\ &\qquad + P(l_5+l_6-k_2-k_3)P(l_6-k_3) + P(l_5-k_2)P(l_6-k_3)\big] \\ &+\int {\rm d}l_4{\rm d}l_6 P(l_4)P(l_6)P(l_6-l_4-k_3)P(k_4+l_6)\big[P(l_4-l_6+k_1+k_3)+P(l_4-k_2)\big] \\ &+\int {\rm d}l_4{\rm d}l_5 P(l_4)P(l_5)P(l_4+l_5-k_2)P(l_4+k_3)P(k_4+l_5)\ . \end{aligned} \tag{5.57}$$

The fifteen terms that have one remaining integral are,

$$\begin{aligned} III = &\int {\rm d}l_6 P(l_6)\big[P(k_1)P(l_6-k_3)P(l_6-k_2-k_3)+P(k_1)P(k_4+l_6)P(l_6-k_1-k_3) \\ &\quad + P(k_1)P(k_4+l_6)P(l_6-k_3)+P(k_2)P(l_6-k_3)P(l_6-k_1-k_3) \\ &\quad + P(k_2)P(k_4+l_6)P(l_6-k_2-k_3)+P(k_2)P(k_4+l_6)P(l_6-k_3) \\ &\quad + P(k_4+l_6)P(l_6-k_2-k_3)P(l_6-k_3)+P(k_4+l_6)P(l_6-k_1-k_3)P(l_6-k_3)\big] \\ &+\int {\rm d}l_5 P(l_5)\big[P(k_4+l_5)P(l_5-k_2)P(l_5-k_2-k_3)+P(k_3)P(l_5-k_2)P(l_5-k_1-k_2) \\ &\quad + P(k_3)P(k_4+l_5)P(l_5-k_2-k_3)+P(k_3)P(l_5-k_2)P(k_4+l_5)\big] \\ &+\int {\rm d}l_4 P(l_4)P(k_4)\big[P(l_4-k_2)P(l_4-k_1-k_2)+P(l_4+k_3)P(l_4+k_1+k_3) \\ &\quad + P(l_4-k_2)P(l_4+k_3)\big]\ . \end{aligned} \tag{5.58}$$

And finally there are 16 terms which do not have any remaining integrals and only depend on

the power spectra of the lognormal modes,

$$
\begin{aligned}
IV = {} & P(k_1)P(k_2)P(k_3) + P(k_1)P(k_2)P(k_4) + P(k_1)P(k_3)P(k_4) + P(k_2)P(k_3)P(k_4) \\
& + [P(k_1)P(k_2) + P(k_3)P(k_4)][P(k_1 + k_3) + P(k_2 + k_3)] \\
& + [P(k_1)P(k_3) + P(k_2)P(k_4)][P(k_1 + k_2) + P(k_2 + k_3)] \\
& + [P(k_1)P(k_4) + P(k_2)P(k_3)][P(k_1 + k_2) + P(k_1 + k_3)] \, .
\end{aligned} \tag{5.59}
$$

IV has the highest contribution out of all of the pure lognormal terms as was shown by Hilbert et al. (2011b) for the covariance of the two point correlation functions. Ergo, to find the covariance of the power spectra for a lognormal field I will neglect I, II and III.

Weak Lensing Covariance

A lognormal distribution provides a more realistic characterization of the convergence field, $\kappa(\theta)$ (see Hilbert et al. 2011a, and references therein). Therefore, we can use the formalism in Sect. 5.3.1 to estimate the covariance of such a field. However, the minimum value of $\kappa(\theta)$ is not minus one unlike the density contrast $\delta(\mathbf{x})$. We can write the lognormal convergence field in terms of a Gaussian field,

$$
\kappa(\theta) = e^{n(\theta)} - \kappa_0 \, , \tag{5.60}
$$

where $n(\theta)$ is a Gaussian random field and κ_0 is the absolute value of the minimum convergence (see Hilbert et al. 2011a). For this thesis the value assumed for κ_0 is 0.012 which corresponds to the value found by Hilbert et al. (2011a) for the Millennium simulation with source galaxy redshift of 0.76. Using this definition instead of the one for $\delta^{\ln}$ in Eq. (5.37) introduces extra constant coefficients in Eq. (5.47). The final result after applying these changes, are shown here.

The covariance matrix of the power spectrum of a lognormal convergence field can be written in terms of the sum of the covariance of a Gaussian field and a purely lognormal term,

$$
\mathbb{C}^{\mathrm{tot}}(\ell, \ell') \equiv \mathbb{C}^{\ln}(\ell, \ell') + \mathbb{C}^{\mathrm{G}}(\ell, \ell') = \langle \hat{P}(\ell)\hat{P}(\ell') \rangle - P(\ell)P(\ell') \, , \tag{5.61}
$$

where $P(\ell)$ is the expected value of $\hat{P}(\ell)$. $\mathbb{C}^{\mathrm{G}}(\ell, \ell')$ is given in Kaiser (1998) as,

$$
\mathbb{C}^{\mathrm{G}}(\ell, \ell') = \frac{4\pi}{A\ell\Delta\ell}(P(\ell) + N_\epsilon)^2 \delta_{\ell\ell'} \, , \tag{5.62}
$$

where $\delta_{\ell\ell'}$ is the Kronecker delta, which make this covariance diagonal. N_ϵ is the noise power spectrum given in Eq. (5.28). $\Delta\ell$ is the width of the ℓ bin considered and A is the area of the field.

The purely lognormal term, $\mathbb{C}^{\ln}(\ell, \ell')$, can be calculated using the purely lognormal terms

of the lognormal moments. We put $k_1 = \ell$, $k_2 = -\ell$, $k_3 = \ell'$ and $k_4 = -\ell'$ in Eq. (5.54) and simplify IV (the other lognormal terms are neglected) to find the desired relation,

$$\mathbb{C}^{\rm ln}(\ell,\ell') \simeq \frac{1}{A^2\kappa_0^2}\int\frac{{\rm d}\varphi_\ell}{2\pi}\int\frac{{\rm d}\varphi'_\ell}{2\pi}\langle\hat{\kappa}(\ell)\hat{\kappa}(-\ell)\hat{\kappa}(\ell')\hat{\kappa}(-\ell')\rangle_{IV} \tag{5.63}$$
$$= \frac{1}{A\kappa_0^2}\left\{2[P^2(\ell)P(\ell')+P^2(\ell')P(\ell)]+[P(\ell)+P(\ell')]^2\int\frac{{\rm d}\varphi_{\ell\ell'}}{2\pi}[P(|\boldsymbol{\ell}-\boldsymbol{\ell}'|)+P(|\boldsymbol{\ell}+\boldsymbol{\ell}'|)]\right\},$$

where $\varphi_{\ell\ell'} = \varphi_\ell - \varphi_{\ell'}$, is the angle between $\boldsymbol{\ell}$ and $\boldsymbol{\ell}'$. Eq. (5.61) is written for the $C_\ell = P(\ell)$. To covariance matrix for $\widetilde{C}_\ell$ is simply,

$$\widetilde{\mathbb{C}}^{\rm tot}(\ell,\ell') = \mathbf{M}\mathbb{C}^{\rm tot}(\ell,\ell')\mathbf{M}^{\rm t}\,, \tag{5.64}$$

where $\mathbf{M}$ is the mixing matrix and matrix multiplications are implied.

5.3.2 Fisher Analysis: Results

The main purpose of this Chapter is to find out how accurate a PCl analysis is and its limitations. Using a Fisher formalism we can quantify the accuracy of such analysis. The main quantity that is of interest here is the ratio of the bias on a deduced model parameter to the errors associated with it. Ideally, the estimation is unbiased. However, as we have seen throughout this Chapter and the previous ones, there are always bias sources. In the case of PCls the main source of bias is due to inaccuracies in mask modelling. Given these considerations the best case scenario is when the bias is significantly smaller than the estimated errors. This renders the bias statistically insignificant. Hence we are going to look at the ratio of the absolute value of the bias, $|B|$, on a parameter to its standard deviation, σ, while the other parameters are marginalized over. In this Chapter the study is limited to two free parameters, σ_8 and $\Omega_{\rm m}$. Since, this work is dedicated to demonstrate the validity of the PCls method for weak gravitational lensing and the number of free parameters do not change the main results. Moreover, having a single redshift distribution results in loose constraints on more than two parameters. We can choose to use different ℓ ranges in the Fisher analysis study. Here I choose the range $\ell \in [2500, 5000]$ by inspecting Fig. 5.6 and Fig. 5.5. All the results shown in this section are for this range.

We first look at $|B|/\sigma$ with respect to n_ℓ for the backward modelling in Fig. 5.7 and forward modelling in Fig. 5.8. These figures consist of twelve panels for the three masks plus their three kinds of apodisation. The $|B|/\sigma$ for σ_8 and $\Omega_{\rm m}$ closely follow each other for all n_ℓ (see Sect. 5.1.2 for the definition of n_ℓ). In the case of backwards modelling the error is the same for all of the plots, since it is calculated from the theory covariance which is independent of the mask considered. By estimating $C_\ell^{\rm rec}$ we try to recover the information lost to the mask. This recovering is not exact which is the source of the bias. On the contrary, the covariance for

forward modelling depends on the mask. Using a larger n_ℓ (a wider ℓ-bin), results in a larger σ. However, as we have seen in Sect. 3.6, narrower band powers are generally more biased. As a results, $|B|/\sigma$ should decrease with an increasing n_ℓ. Although this trend is seen for most cases in Fig. 5.7 and also Fig. 5.8, some of the cases deviate from it. This is due to lucky cancellation effects using different binning, which decrease the bias for certain combinations. By comparing these figures we can conclude that the forward modelling (using PCls as observables) is more reliable than attempting to recover the input C_ℓ. Since, in Fig. 5.8, $|B|/\sigma < 1$ for large n_ℓ for all the cases and generally decreases with apodisation.

A comparison of different masks in Fig. 5.7 and Fig. 5.8 shows that the large scale mask is generally more difficult to model and results in a larger relative bias. This effect is more pronounced for the composite mask where all the scales are affected by the masks. Note that in all case in this work the zero padding is present, which affects the small ℓ-modes, and propagates through to all modes in the mixing matrix estimation. Apodising the mask increases the area covered by the mask, which in turn results in a more unstable recovery of the C_ℓ which is seen in Fig. 5.7. However, in Fig. 5.8 we see that apodisation benefits the analysis.

To further investigate the apodisation and bin width effects, we can look at Fig. 5.9 and Fig. 5.10. Both figures show the simultaneous constraints on σ_8 and Ω_m and the corresponding bias in their estimated values, using Fisher analysis. The difference between the two figures is the width of the ℓ-bins used. Fig. 5.9 corresponds to $n_\ell = 10$ ($\Delta\ell \approx 180$), whereas, $n_\ell = 20$ ($\Delta\ell \approx 360$) for Fig. 5.10. The contours are in general smaller for the narrower bins as expected. The top plots in both figures show the one and two σ results for the backward modelling and the bottom ones display the one σ contours for forward modelling. These figures belong to the composite mask in its original form and with the medium sized apodisation scheme. Results are shown for both lognormal and Gaussian shear fields. The contours belonging to the Gaussian fields are shifted in the σ_8 direction by 0.1, to assist with inspections. Black and green contours belong to the Gaussian fields, for the original and apodised masks, respectively. Similarly, red and blue contours show the results for the lognormal fields for the original and apodised masks, respectively. The fiducial value of the parameters is shown by +. The symbol $\times$ shows the shifted position of the fiducial value, which should be compared with the Gaussian contours.

As we have seen throughout Sect. 5.2 and this section, the forward modelling provides better and more consistent results, which can also be seen in Fig. 5.9 and Fig. 5.10. Although the apodisation does not affect the backward modelling significantly, it decreases both the bias and estimated errors on the parameters for forward modelling. The results are similar for the lognormal fields and the Gaussian ones. The lognormal contribution to the covariance is small for the ℓ-range considered here. Fig. 5.9 shows that without apodisation for $n_\ell = 10$ the bias is too large. However, when the mask is apodised the bias is reduced and is inside the one σ contour for the forward modelling case (bottom panel). Even with the larger n_ℓ in Fig. 5.10 the

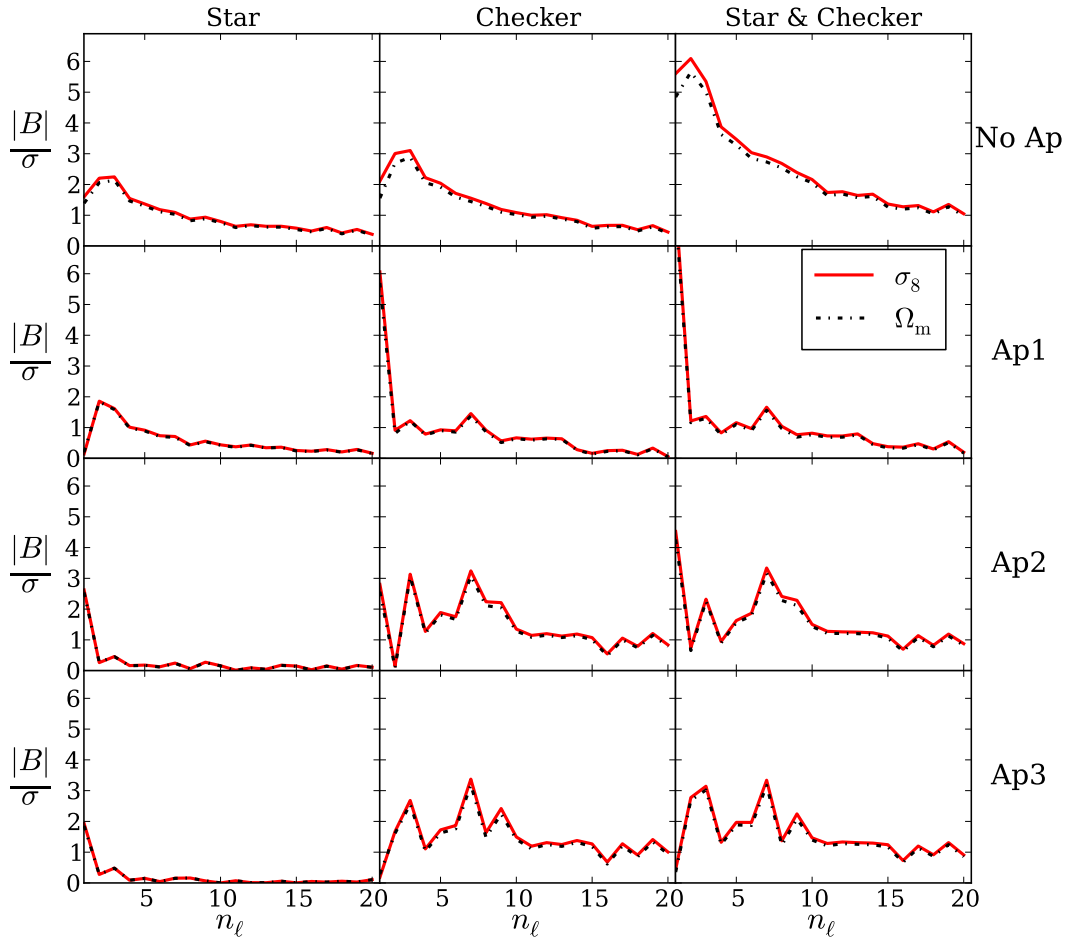

Figure 5.7: The ratio of bias to error on the estimated parameters with respect to the binning, using C_ℓ as the observable. n_ℓ is the number of original ℓ bins merged to make the new binning. A larger n_ℓ corresponds to a wider bin. The estimated parameters are σ_8 (red solid line) and Ω_m (black dashed line). All the cosmological parameters are fixed to their fiducial values in Table. 5.1 expect for these two parameters. A Fisher analysis is used for estimating the bias and error on the estimated parameter. To find the σ for each parameter the other one is marginalized over. The top panels show this ratio for original masks with no apodisation. The lower panels display $|B|/\sigma$ for the apodised versions of the masks, with Ap1 corresponding to the smallest kernel, Ap2 to the medium size and Ap3 to the largest (see Eq. 5.33). The columns correspond to the three masks used in this thesis.

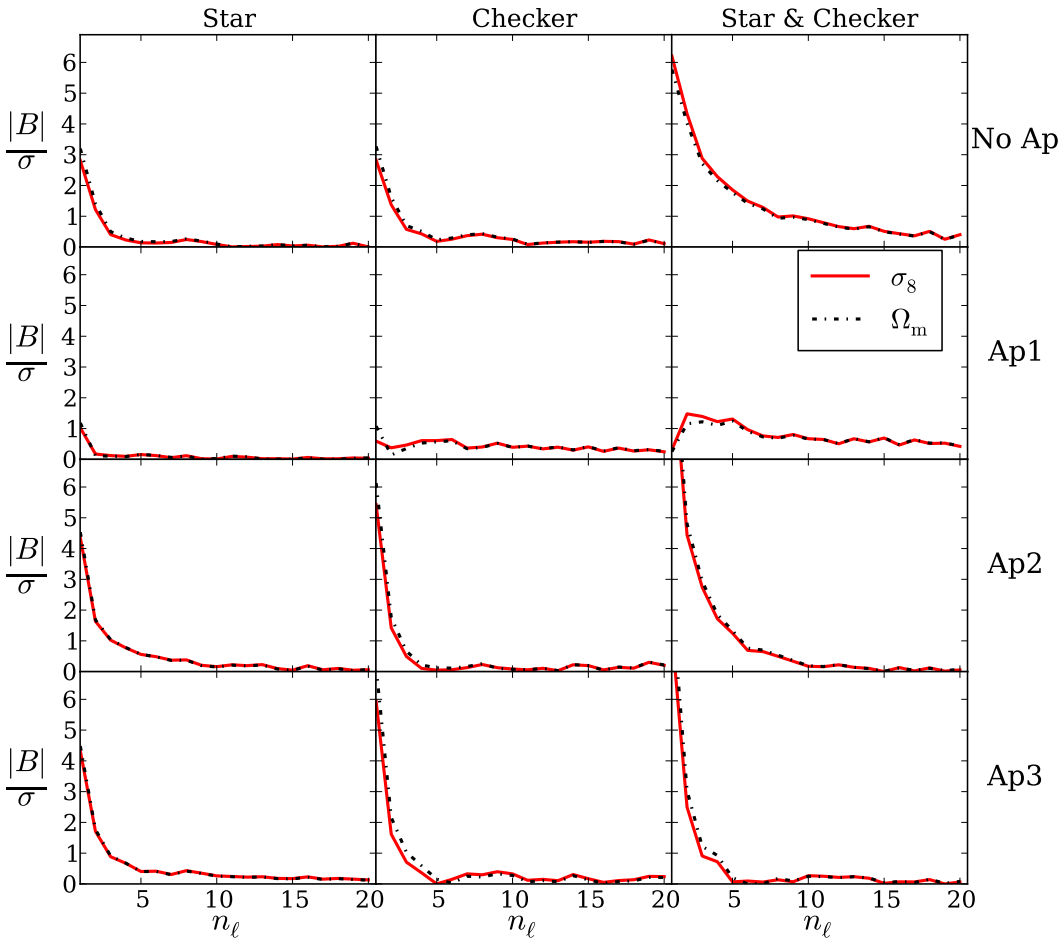

Figure 5.8: The ratio of bias to error on the estimated parameters with respect to the binning, using $\widetilde{C}_\ell$ as the observable. See the caption of Fig. 5.7 for more details.

fiducial point lies outside of the 95% range for forward modelling.

5.4 Conclusion

Pseudo Cl analysis is a method that models the effects of masks on the Fourier transform of a field. In this Chapter I used the flat sky version of this method on simulated shear fields to investigate its limitations and validity for weak lensing analysis. This is particularly interesting for future large scale surveys such as Euclid[4] and LSST[5]. Although, PCl has been used for cosmic microwave background analysis as was mentioned in the introduction of this Chapter, it has never been tested to the extent that was done here. Specifically, the effects of incomplete mask modelling on the estimated cosmological parameters has never been demonstrated in the literature.

In this Chapter I started by showing the general PCl formalism, which shows how the E-/B-modes are mixed in Fourier space as a result of the masks. Then in Sect. 5.1.2 I explained how to model the mask in practice, given a pixelized field. In this section smoothing the mask (apodisation) was introduced as a recipe to avoid some of the mask modelling difficulties. Complicated masks with small and large scale features were used in this work to investigate the mask modelling in more detail. Hikage et al. (2011) who performed an analysis of simulated shear fields with full and flat sky PCl only considered small scale masks and did not propagate the error to the measured parameters.

All the fields used in this Chapter were cutoffs of larger simulated fields, using Gaussian and lognormal fields, with the lognormal field used as the default. As a result, to analyse the fields they were first zero padded to avoid periodic boundary conditions. Therefore, the masks which are also modelled in Fourier were first zero padded. The zero padded mask is effectively the original mask plus a large scale square shaped mask. Hence, a large scale mask was always present in this analysis. The E-/B-mode mixing of the power spectra due to the mask is modelled in a "mixing matrix". The mixing matrix is unable to model the mask effects perfectly due to pixelization, which renders the angular averages on the Fourier transform of the mask inaccurate. This inaccuracy is specially important for large scale masks. Thus, having a zero padded mask, which is a necessity for simulating realistic scenarios, generates challenges in estimating the mixing matrix. The inaccuracies which result from such masks propagate through all the Fourier modes (see Eq. 5.21), and cause a constant multiplicative bias in the measured power spectra. The solution to this problem is to use simulations to correct for this bias, since it is insensitive to the input power spectrum.

To apodise the fields I changed the input mask with a smoothed version of it, using Gaussian kernels with three sizes. This smoothing increases the size of the masked region, however,

[4] http://sci.esa.int/euclid/
[5] http://www.lsst.org/lsst/

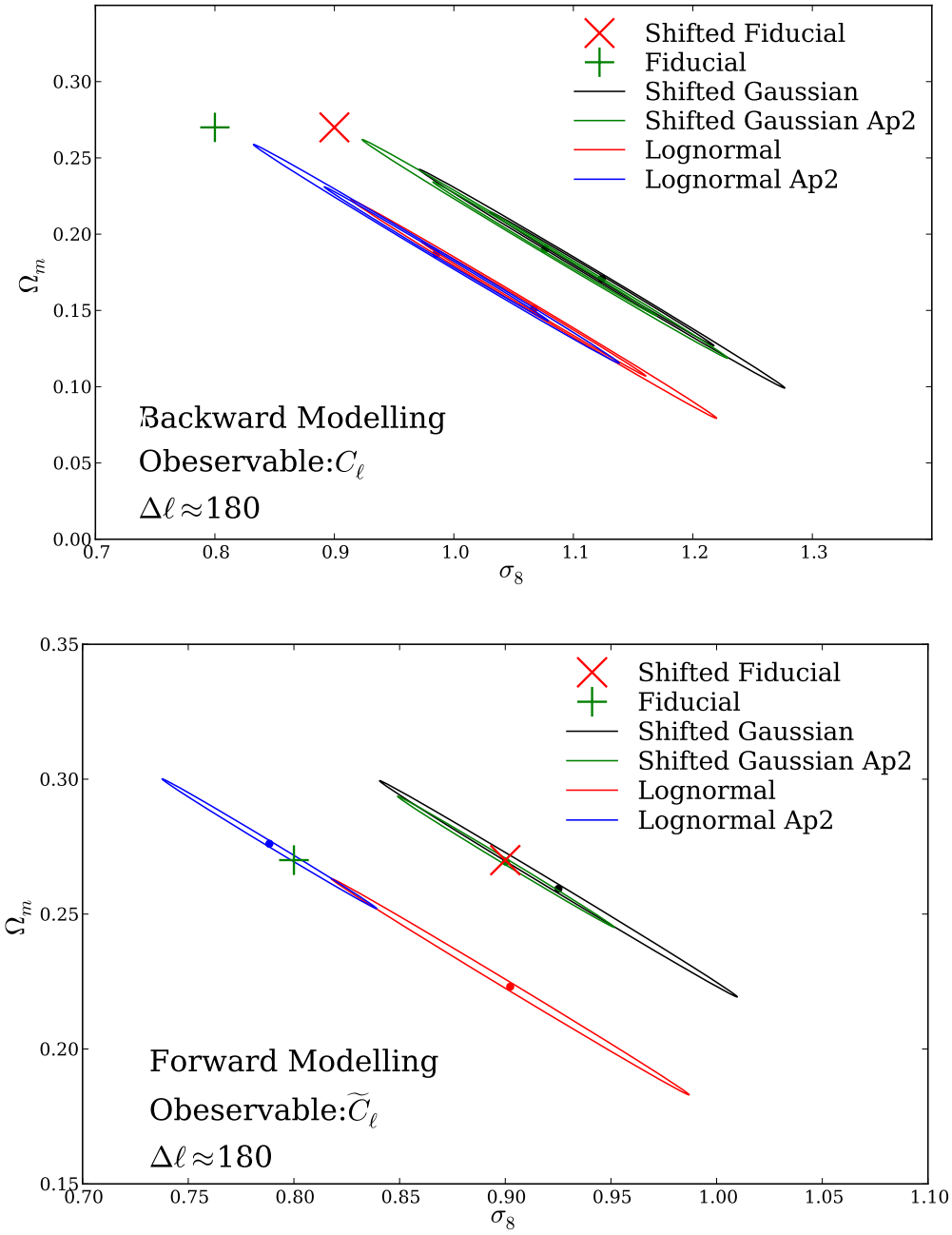

Figure 5.9: Constraints on σ_8 and Ω_m. Here $n_\ell = 10$ which means $\Delta\ell \approx 180$. The one and two σ contours are drawn around the best fit value for the top panel which uses the Cls. The bottom panel corresponds to the forward modelling which uses PCls as observables. The contours for this panel show the one σ constraints. The contours are shown for Gaussian and lognormal fields. The Gaussian contours are shifted in σ_8 by 0.1 for an easier comparison. The fiducial values of the parameters is shown by + and the shifted value by $\times$. The black and red contours belong to the original composite mask, while the green and blue ones show the estimated errors for its apodised version. The medium sized kernel with 11 pixels is used here.

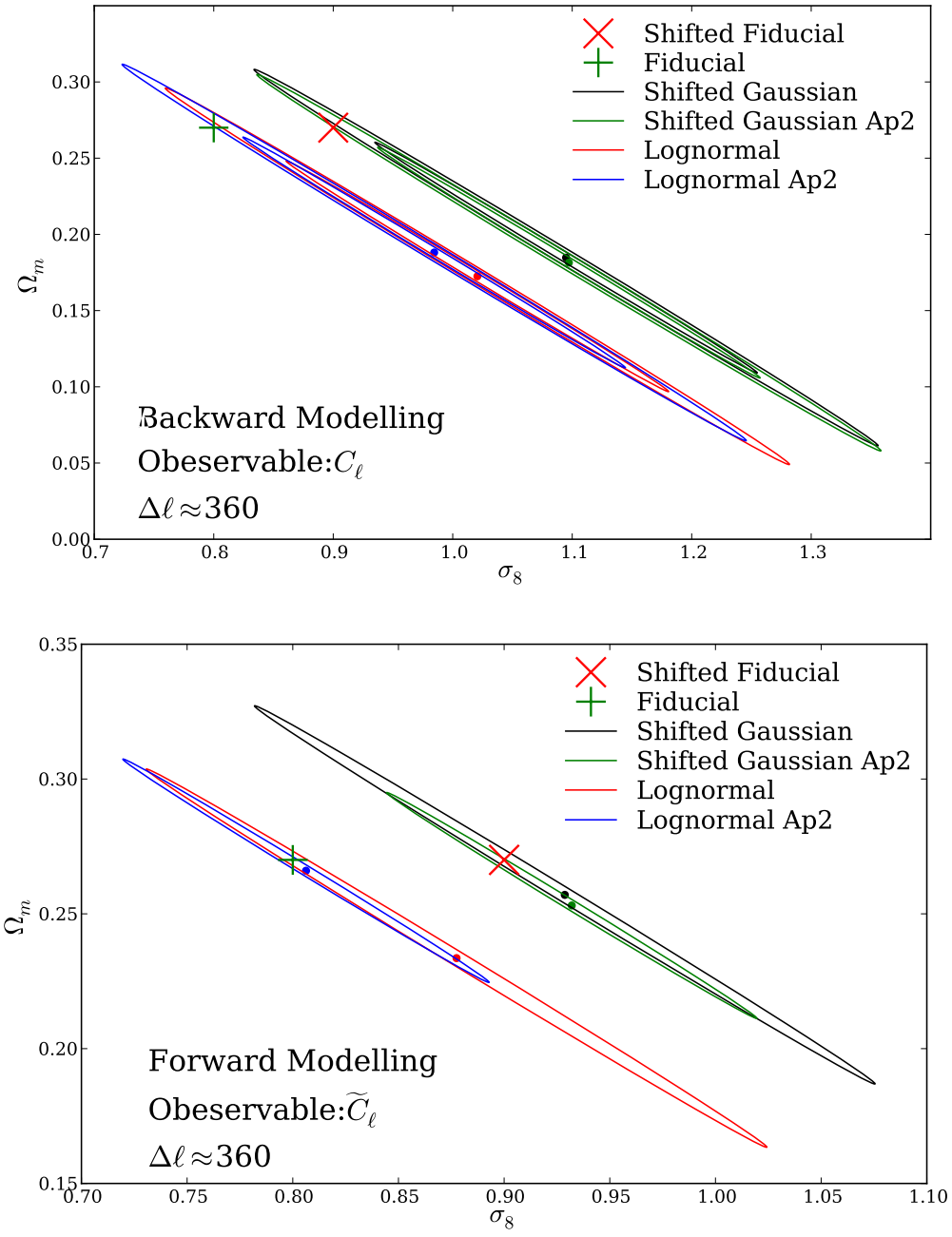

Figure 5.10: Constraints on σ_8 and Ω_m with $n_\ell = 20$ ($\Delta\ell \approx 360$). See the caption of Fig. 5.9 for more details.

since the apodised masked convergence field has no sharp edges, its Fourier transform varies slower, which makes the angular averages more accurate. The mixing matrices corresponding to the apodised masks are in general more diagonal, in other words, different modes have smaller mixing with each other.

In Sect. 5.2 we compared the estimated values of the PCls from the simulations to their expected value from theory. Moreover, using the inverse of the mixing matrix the input power spectrum was recovered. The ratios of theory and estimated values of PCls and Cls show how accurate the mask modelling is. The modelling of very large or small scales are in general difficult and that is where most of the discrepancies lie. The small scale signal is dominated by galaxy shape noise, hence they provide very little information. Most of the gravitational lensing information comes from midrange scales, which can be measured using the PCl analysis. Recovering the input power spectrum is considerably more challenging than the forward modelling. We can see that by comparing Fig. 5.5 and Fig. 5.6.

I used Fisher analysis to propagate the errors and biases in mask modelling to two cosmological parameters, σ_8 and Ω_m. To estimate the Fisher matrix, the covariance matrix of the observables is needed. In Sect. 5.3.1 I provided an algorithm for calculating the moments of a lognormal field. This was then used to estimate the covariance of the power spectra for the simulated lognormal shear fields.

With the Fisher formalism, I compared the linear bias on the two parameters to their estimated error, for different masks and apodisation schemes. These results were shown for the midrange scales, $2500 \lesssim \ell \lesssim 5000$. We saw that the checkerboard mask is in general harder to model compared to the star mask, since the ratio of the biases to errors is higher for this mask. In addition, apodisation in general decreases the bias to error ratio.

In Chapter 3 I showed that narrow band power spectra are generally biased and contaminated by power from outside their range. Also here I compared the relative value of bias to error using different bin sizes in Fourier space, which lead to the same conclusion: increasing the width of the ℓ-bin decreases the relative bias.

Finally, I showed the simultaneous error and bias on the two parameters for lognormal and Gaussian shear fields. The covariance of the lognormal fields is dominated by the Gaussian shape noise. As a result there is little statistically significant difference between the two cases. We also saw that the apodised fields show a smaller bias and tighter constraints on the estimated parameters.

In conclusion for a flat sky PCl analysis the following considerations should be taken into account:

- Forward modelling is superior to backward modelling, and causes smaller biases.
- Using narrow ℓ-bins can cause large biases.
- The inaccuracies in mask modelling will result in a constant bias, which should be

corrected for using simulations.

- Apodising (smoothing) the mask before applying them on the fields will reduce the biases and increase the constraint on deduced parameters.

Chapter 6

Conclusions

Modern cosmology is at a very exciting point in its relatively young life. The advances in technology have provided us -for the first time- the ability to look far into the past. Cosmic microwave background (CMB) surveys have made incredible maps of temperature fluctuations which show (statistically) the seeds of all the large and small structures we see today. We have observed the light of exploding stars from half way through space-time, and concluded that the Universe is accelerating in its expansion. Large scale structures have been observed in the form of clusters, filaments and walls. Aside from the progress in astronomical instruments, we have seen a revolution in processing power, provided by supercomputers. This is the first time in the history of Earth[1], that the opportunity to understand the essence of the Universe has arisen. Although, these observations have provided us with valuable information about our Universe, the picture is not yet complete. In fact it is far from complete. Our current standard cosmological model, although simple, is not particularly insightful. It states that the vast majority of the energy budget of the Universe is "dark"! In other words (we know that) we have no idea what the Universe is mainly made up of, and the small percentage that we do know about (baryons) is too difficult to model in a comprehensive manner.

Other probes of the cosmos have the potential to shed some light onto this problem. Weak gravitational lensing (cosmic shear) is one of the most promising probes of the dark sector. Cosmic shear is a probe of the large scale structures and the expansion history of the Universe in relatively late times. The future large scale cosmic shear surveys such as Euclid[2], LSST[3] and WFIRST[4] have the potential to, for example, reveal a varying dark energy component with the equation-of-state parameter $w_0 \neq -1$ or $w_1 \neq 0$. If they provide evidence against the cosmological constant, it will be the discovery of the century[5]! Another possible discovery that will have a great impact on our understanding of dark energy is finding that it has structures.

[1] as far as we know.
[2] http://sci.esa.int/euclid/
[3] http://www.lsst.org/lsst/
[4] http://wfirst.gsfc.nasa.gov/
[5] The question is then: what happened to the vacuum energy?

This also would mean that it is not the cosmological constant and will favor dark energy over modified gravity models.

As I write this final Chapter the LHC[6] has started looking for the dark matter particle at a record high energy level. If they find traces of this weakly interacting particle, the standard particle theory will have to expand and cosmology will be more intertwined with particle physics. However, this particle may be much heavier than the level of energy LHC can reach. In which case we need to again go back to the Universe to look for clues as to the nature of dark matter.

The origin of the Universe is also an open question. Inflation which covers a vast variety of theories is the most popular model available. Although there are other arguably equally plausible models which can solve the same problems inflation was designed to tackle (for example string gas cosmology, Brandenberger 2011). Recently a press conference by the BICEP[7] group about the detection of primordial B-modes in the CMB data caused a lot of excitement. Even though their signal was later proved to be dominated by dust using data from Planck[8], it demonstrated how much such a discovery can affect our knowledge of the Universe, as it inspired many articles to be written in the short time it was believed to be true.

I have focused on studying cosmology from the gravitational lensing point of view. There are different levels of difficulty in extracting physical understanding from cosmic shear data, from the image processing to model selection. I was mainly interested in developing, implementing, testing and applying statistical methods of cosmic shear analysis. In the next section I will go through a summary of my thesis and then in the last section I will briefly speculate on the applications of my work in the future and how it can be expanded and improved.

6.1 Summary

This thesis started with an introduction to cosmology in Chapter 1. In Chapter 2 we went through the basics of gravitational lensing and cosmic shear analysis. My research for this thesis was presented in Chapters 3, 4 and 5.

In Chapter 3 I explain a compression method that I developed (Asgari & Schneider 2015) which is based on the sensitivity of observables (statistics) to the parameters to be measured. This method relies on our understanding of these parameters, since the compressed observables depend on the covariance and derivatives of the parent observable to the parameters at their fiducial value. We expect to lose information in the compression if these fiducial values are substantially different from the true parameters. The main motivation for data compression is that the number of simulations needed to estimate the data covariance matrix accurately,

[6] http://home.web.cern.ch/topics/large-hadron-collider
[7] https://www.cfa.harvard.edu/CMB/bicep2/science.html
[8] http://sci.esa.int/planck/

depends on the number of observables. Therefore, having a smaller set of observables reduces the number of cosmological simulations needed. This is a great advantage as running simulations is computationally expensive. To define the compressed observables we do however need some knowledge of the parent observable covariance matrix.

I applied this method of compression to COSEBIs (Complete Orthogonal Sets of E-/B-Integrals explained in Chapter 2) to make compressed COSEBIs (CCOSEBIs). I used a Fisher analysis to find out how much information is lost if the data covariance is inaccurate. I found that using the first order CCOSEBIs alone, with an inaccurate covariance, results in a considerable loss of information, however, the second order statistics and the combination of first and second orders, does not suffer significantly. As a result, I concluded that using this method for data analysis is highly desirable.

In the second part of Chapter 3 I use the COSEBIs filter functions to show how well a band power spectrum can be estimated from a finite field data. A top hat (the function used to define the band power spectra) has finite support in Fourier space, which translates to infinite support in real space. When the angular ranges available in real data are limited the estimated band powers will be contaminated by information from outside the band. Using COSEBIs filters I showed that the band power is biased in a cosmology independent way. The contamination from outside of the band is smaller for a large angular range and a wide bin. Hence, if band powers are to be defined for cosmic shear analysis these contaminations should be considered in the analysis.

In Chapter 4 I used CFHTLenS (current state-of-the-art cosmic shear survey) data to estimate cosmological parameters using COSEBIs and CCOSEBIs. I used a redshift range for which intrinsic alignments are believed to be small (Heymans et al. 2013). I also used tomographic bins in the analysis and compared that to a single redshift distribution. COSEBIs separates E-/B-modes on a finite angular range efficiently (lossless). Any method that can do so can be transformed into COSEBIs, or is based on it. I used two free parameters for this thesis, σ_8 and Ω_m, since the redshift range I considered was limited. Before performing the analysis on the data, I setup a blind pipeline for each method, which was first tested on a set of simulations that mimic the data. I estimated the cosmological parameters of the simulations using a likelihood analysis with a flat prior. The likelihood distribution was sampled using an MCMC algorithm. After finding satisfactory results with the simulations by confirming the the codes reproduce the input cosmological parameters, I applied the pipelines without any changes to the CFHTLenS data. The best fit values I found are statistically consistent with the values found by the CFHTLenS team (see for example Kilbinger et al. 2013, , $\sigma_8 = 0.813$ and $\Omega_\mathrm{m} = 0.283$). The results are consistent for COSEBIs and CCOSEBIs. I found a small remaining B-mode contribution on large scales in the data. The significance of this B-mode component depends on the accuracy of the covariance estimated from the simulations. However, in other words if the simulations are not representative of reality the covariances will

not be either. This is most probably the case since the box size of the simulations we used is relatively small.

Chapter 5 views the cosmic shear analysis from a different angle. I implemented a Fourier space method, pseudo Cls (PCls), and tested it on Gaussian and lognormal simulated shear fields. Although, the presence of masks on images complicates the estimation of convergence (shear) power spectra, there are advantages in Fourier space analysis of cosmic shear data. The power spectrum is more readily comparable with theory and for large scales shows a more diagonal covariance. In addition, we can choose the Fourier modes that we want to use in our analysis, according to our understanding of the theory. For a PCl analysis the effects of the mask on the Fourier (harmonic) modes is modeled in a mode-mixing matrix. PCl analysis can be done on the full sky with spherical harmonics or on a patch of sky where flat sky approximation can be applied. Estimating PCls from a curved sky has the advantage of having well defined discrete bins, in contrast to the flat sky version. In other words the full sky estimation in practice is much closer to the theory than the flat sky. For a flat sky estimation the choice of binning is arbitrary which complicates and biases the estimation of the mixing matrix. I used large and small scale masks to show this effect.

I compared flat sky PCls estimated from the simulated shear fields with small and large scale masks to their theory values to test the mask modeling. These fields were zero padded to avoid periodic boundary conditions. The zero padding effectively adds a very large square shaped mask to the field. In Fourier space this square is characterized approximately by a sinc function which peaks at low Fourier modes. The mixing matrix is estimated using angle averages of the mask multiplied to trigonometric functions. The angle averaging of a pixelized field is specially inaccurate for small Fourier values, where the number of pixels that lie in an annulus is small. This inaccuracy will bias the mixing matrix on all scales. Using these simulations I found that there is a constant multiplicative bias in the mixing matrix which can be corrected for using the simulations, since it is insensitive to the input power spectrum.

The product of the mixing matrix and the input power spectrum provides the theory PCls. The mixing matrix can also be inverted and applied to the estimated PCls to recover the power spectrum. I compared a forward and backward PCl modeling and found that in general forward modeling (using the pseudo power as the observable) is more reliable and less biased. However, even in this case , large scales are unreliable because of the mask modeling inaccuracies explained before. As a result, I chose the range 2500-5000 for the Fisher analysis section. I propagated the errors on the PCls to cosmological parameters and also estimated the bias on these parameters using Fisher analysis. I found consistent results with Chapter 3 that binning band power spectra results in biased values. The wider the band power the smaller the biases on the estimated parameters. I also explored the effects of apodisation on sharp masks and found that apodisation provides a more diagonal mixing matrix which results in smaller biases and tighter constraints on cosmological parameters, using forward modeling.

6.2 Prospects

The largest completed cosmic shear data set is CFHTLenS, which is a rather small survey. The KiDS[9] collaboration has just released its first science papers. It will ultimately reach an area about 10 times as large as the CFHTLenS area. Although it is not as deep as CFHTLenS it can provide tighter constraints on model parameters. DES[10] is two years into its five year mission of mapping one-eighth of the sky. The HSC[11] installed on the Subaru telescope has recently started imaging the sky with great resolution for a five year weak lensing survey. Other surveys such as Euclid and LSST will start imaging in the future. The data analysis with CFHTLenS provides a testbed for all of these surveys. The CFHTLenS area was too small to compete with other cosmological probes such as CMB. However, these new and ongoing weak lensing surveys will be able to provide complimentary and competitive information to our cosmological knowledge. That is assuming we understand how to analyze them.

Currently, the most popular real space method used in cosmic shear analysis is not COSEBIs. I hope to change that by publishing the analysis in Chapter 4 on CFHTLenS to show that COSEBIs should be the preferred method since it is efficient and accurate. The improvements that I plan to apply to that Chapter before publication involve using a new suite of N-body simulations from Harnois-Déraps & van Waerbeke (2015). I will also include intrinsic ellipticity alignments so that I can use the lower redshift bins, explore the B-modes I detected on large scales and hopefully bypass the low value of some of the χ^2.

Studying cosmic shear data in Fourier space provides an independent way to test for systematics. Although PCls have been used on CMB data, they have never been applied to weak lensing. I explored the effects of masking in detail (for a flat sky PCl analysis) for the first time. The results I found will be particularly interesting for Euclid which has requirements on the estimated power spectra. Although flat sky PCls has its limitations, it can be used with certain precautions. It will at least provide sanity and consistency checks for future analysis.

One of the biggest challenges for the future of weak lensing analysis is making accurate and large enough cosmological simulations. In particular, simulating the effects of baryons is not only time consuming but also uncertain as different feedback models generate very different results. Accurate modeling of redshift space distortions requires very large simulations, while galaxy formation models need high resolutions. The use of hybrid simulations have been suggested to solve some of these problems. Regardless, by using PCls and COSEBIs we can control the Fourier modes and angular ranges used in the analysis, respectively and avoid scales for which the simulations are not to be trusted. Running high resolution large simulations is very time consuming. Realistically, it is currently impossible to run enough simulations for most statistical methods, since the number of simulations needed to estimate an accurate

[9] Kilo Degree Survey: http://kids.strw.leidenuniv.nl/
[10] Dark Energy Survey: http://www.darkenergysurvey.org/index.shtml
[11] Hyper Suprime Cam: http://www.naoj.org/Projects/HSC/

enough covariance matrix depends on the number of observables. The accuracy needed for the covariance matrix also depends on the survey. For large future surveys with small noise the covariances need to be estimated with a much higher accuracy compared to the requirements for CFHTLenS. Although COSEBIs and PCls both compress the data substantially, with the presence of tomography their data vectors become too large. Tomographic redshift bins are necessary to see the evolution of the Universe. It provides a great increase in the information level, especially regarding the dark energy parameters. Moreover, in order to incorporate intrinsic alignments in the analysis about 10 or more tomographic bins are required. This will quickly increase the total number of observables. Hence compression methods are highly desirable to tackle this issue. Consequently, the compression method I showed in Chapter 3 will provide a very strong tool for the future data analysis and reduce the number of simulations needed substantially.

Two point statistics have been the focus of cosmic shear analysis so far, with a few exceptions (see Fu et al. 2014, for example). This is mainly due to the difficulty of higher order analysis and the fact that the number of observables quickly increases. For a non-Gaussian field higher order statistics provide extra information. Therefore, a weak lensing analysis can benefit from higher order information. Currently, E-/B-mode decomposition for 3-point statistics has not been done in a COSEBI-like manner, which means that there is room for significant improvements. PCl analysis can also be carried out for higher order statistics. It would be highly interesting to develop efficient methods for 3-point statistics. Also in this case data compression will become even more crucial. The information from the higher orders can be complimentary to the 2-point statistics and also provide sanity checks.

We will see a boom in the cosmic shear field in the next decade. There are still challenges to face. However, as the statistical methods improve and other analysis methods are introduced to this field (for example machine learning), the chances of triumph increases. The future looks bright!

Appendix A

An Unsuccessful Compression Method

In this appendix we see the steps taken to develop a compression method which was tried before the method in Chapter 3 was developed. This compression method is also based on Taylor expansions, however, it lacks the flexibility of the method in Chapter 3 as it is sensitive to the choice of the covariance.

Let us write the observables in terms of their zeroth and first order Taylor series components with respect to the parameters to be estimated,

$$X_n(\phi_\mu) = X_n^{\rm f} + \mathbf{D}_{n\mu} p_\mu;, \tag{A.1}$$

where $p_\mu = \phi_\mu - \phi_\mu^{\rm f}$, and

$$\mathbf{D}_{n\mu} = \left(\frac{\partial X_n}{\partial \phi_\mu}\right)_{|\phi_\kappa^{\rm f}} , \tag{A.2}$$

are the first derivatives of the expectation values with respect to the model parameters, taken at the fiducial point in parameter space. The Taylor expansion in Eq. (A.1) can be written in matrix form,

$$\Delta \boldsymbol{X}(\Phi) = \mathbf{D}\boldsymbol{p} \, . \tag{A.3}$$

We can write a χ^2 for Eq. (A.3),

$$\chi^2 = (\Delta \boldsymbol{X} - \mathbf{D}\boldsymbol{p})^{\rm t} \, \mathbf{C}^{-1} \, (\Delta \boldsymbol{X} - \mathbf{D}\boldsymbol{p}) \ , \tag{A.4}$$

where $\mathbf{C}$ is the covariance of the $\boldsymbol{X}$. We minimize the χ^2 with respect to the cosmological parameters by taking its partial derivatives and equating them to zero,

$$\frac{\partial \chi^2}{\partial p_\kappa} = -2 D_{m\kappa} C^{-1}_{mn} \, (\Delta X_n - D_{n\nu} p_\nu) = 0 \, , \tag{A.5}$$

rearranging Eq. (A.5) leads to

$$\mathbf{D}^{\rm t} \mathbb{C}^{-1} \Delta \boldsymbol{X} = \mathbf{D}^{\rm t} \mathbb{C}^{-1} \mathbf{D} \boldsymbol{p} \, . \tag{A.6}$$

We define $\mathbf{V} \equiv \mathbf{D}^{\mathrm{t}}\mathbb{C}^{-1}$ and $\mathbf{U} \equiv \mathbf{D}^{\mathrm{t}}\mathbb{C}^{-1}\mathbf{D}$ and hence conclude,

$$\boldsymbol{p}^{(1)} = \mathbf{U}^{-1}\mathbf{V}\Delta\boldsymbol{X} \equiv \mathbf{R}\Delta\boldsymbol{X} \,, \tag{A.7}$$

here the superscript (1) indicates that this is a first order approximation for $\boldsymbol{p}$. We can further improve the approximation for $\boldsymbol{p}$ by going to higher orders in the Taylor expansion.

$$X_n(\Phi) = X_n^{\mathrm{f}} + X_{n,\mu}^{\mathrm{f}} p_\mu + \frac{1}{2} X_{n,\mu\nu}^{\mathrm{f}} p_\mu p_\nu \,, \tag{A.8}$$

where

$$X_{n,\mu\nu}^{\mathrm{f}} = \left.\frac{\partial^2 X_n}{\partial p_\mu \partial p_\nu}\right|_{\Phi^{\mathrm{f}}} \equiv Z_{n\mu\nu} \,. \tag{A.9}$$

We can write $\boldsymbol{p} = \boldsymbol{p}^{(1)} + \boldsymbol{p}^{(2)}$, where $\boldsymbol{p}^{(1)}$ and $\boldsymbol{p}^{(2)}$ are the first and second order approximations for $\boldsymbol{p}$. Keeping the terms up to second order in Eq. (A.8) results in

$$X_n(\Phi) = X_n^{\mathrm{f}} + D_{n\mu}(p_\mu^{(1)} + p_\mu^{(2)}) + \frac{1}{2} Z_{n\mu\nu} p_\mu^{(1)} p_\nu^{(1)} \,, \tag{A.10}$$

The χ^2 is

$$\chi^2 = \left(\Delta X_m - D_{m\mu}(p_\mu^{(1)} + p_\mu^{(2)}) - \frac{1}{2} Z_{m\mu\nu} p_\mu^{(1)} p_\nu^{(1)}\right) C_{mn}^{-1} \left(\Delta X_n - D_{n\mu}(p_\mu^{(1)} + p_\mu^{(2)}) - \frac{1}{2} Z_{n\mu\nu} p_\mu^{(1)} p_\nu^{(1)}\right) , \tag{A.11}$$

Minimizing χ^2 with respect to $p_\kappa^{(2)}$ as before gives,

$$\frac{\partial \chi^2}{\partial p_\kappa^{(2)}} = -2 D_{m\kappa} C_{mn}^{-1} \left(\Delta X_n - D_{n\mu}(p_\mu^{(1)} + p_\mu^{(2)}) - \frac{1}{2} Z_{n\mu\nu} p_\mu^{(1)} p_\nu^{(1)}\right) = 0 \,. \tag{A.12}$$

Rearranging the above results in,

$$V_{\kappa n}\Delta X_n - U_{\kappa\mu} p_\mu^{(1)} - \frac{1}{2} V_{\kappa n} Z_{n\mu\nu} p_\mu^{(1)} p_\nu^{(1)} = U_{\kappa\mu} p_\mu^{(2)} \,. \tag{A.13}$$

The first two components in the above equation cancel each other (see Eq. A.6) hence we find the following relation for $p_\mu^{(2)}$ in terms of $\boldsymbol{p}^{(1)}$

$$\begin{aligned} p_\delta^{(2)} &= -\frac{1}{2} (U^{-1})_{\delta\kappa} V_{\kappa n} Z_{n\mu\nu} p_\mu^{(1)} p_\nu^{(1)} \\ &= -\frac{1}{2} R_{\delta n} Z_{n\mu\nu} R_{\mu m} \Delta X_m R_{\nu l} \Delta X_l \\ &\equiv Q_{\delta ml} \Delta X_m \Delta X_l \,. \end{aligned} \tag{A.14}$$

From the above equation we can see that $p^{(1)}$ depends on the product of two ΔE's. In

conclusion the estimated $\boldsymbol{p}$ is

$$p_\mu = p_\mu^{(1)} + p_\mu^{(2)} = (R_{\mu m} + Q_{\mu nm}\Delta X_n)\Delta X_m \tag{A.15}$$

which gives us the same number of linear combinations of COSEBIs as the free parameters. Using this method for the data compression does not allow for a larger number of compressed observables. In fact, it requires the same number of observables as the number of parameters regardless of the expansion order considered. Therefore, given a wrong covariance matrix the compressed quantities will not span the same parameter space as the true underlying parameters. Which is the reason behind developing the compression method in Chapter 3. The Taylor expansion in that case is used as a guideline for making the compressed observables, unlike here where I have strictly used χ^2 to define the compression matrices.

Bibliography

Anderson, L., Aubourg, E., Bailey, S., et al. 2012, MNRAS, 427, 3435

Asgari, M. & Schneider, P. 2015, A&A, 578, A50

Asgari, M., Schneider, P., & Simon, P. 2012, A&A, 542, A122

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

Bartelmann, M., Fabis, F., Berg, D., et al. 2014, ArXiv e-prints

Bartelmann, M. & Schneider, P. 2001, Phys. Rep., 340, 291

Bassett, B. A., Fantaye, Y., Hlozek, R., & Kotze, J. 2009, ArXiv e-prints

Becker, M. R. & Rozo, E. 2014, ArXiv e-prints

Blazek, J., McQuinn, M., & Seljak, U. 2011, J. Cosmology Astropart. Phys., 5, 10

Bond, J. R. & Efstathiou, G. 1984, ApJ, 285, L45

Brainerd, T. G., Blandford, R. D., & Smail, I. 1996, ApJ, 466, 623

Brandenberger, R. H. 2011, Classical and Quantum Gravity, 28, 204005

Brown, M. L., Ade, P., Bock, J., et al. 2009, ApJ, 705, 978

Brown, M. L., Castro, P. G., & Taylor, A. N. 2005a, MNRAS, 360, 1262

Brown, M. L., Castro, P. G., & Taylor, A. N. 2005b, MNRAS, 360, 1262

Brown, M. L., Taylor, A. N., Bacon, D. J., et al. 2003, MNRAS, 341, 100

Crittenden, R. G., Natarajan, P., Pen, U.-L., & Theuns, T. 2002, ApJ, 568, 20

Delubac, T., Bautista, J. E., Busca, N. G., et al. 2015, A&A, 574, A59

Dodelson, S. 2003, Modern Cosmology (Academic Press)

Eifler, T. 2011, MNRAS, 418, 536

Eifler, T., Krause, E., Schneider, P., & Honscheid, K. 2014, MNRAS, 440, 1379

Erben, T., Hildebrandt, H., Miller, L., et al. 2013, MNRAS, 433, 2545

Fu, L. & Kilbinger, M. 2010, MNRAS, 401, 1264

Fu, L., Kilbinger, M., Erben, T., et al. 2014, MNRAS, 441, 2725

Gelmann, A. & Rubin, D. 1992, Statistical Science, 7, 457

Hand, N., Leauthaud, A., Das, S., et al. 2013, ArXiv e-prints

Harnois-Déraps, J., Pen, U.-L., Iliev, I. T., et al. 2013, MNRAS, 436, 540

Harnois-Déraps, J., Vafaei, S., & Van Waerbeke, L. 2012, MNRAS, 426, 1262

Harnois-Déraps, J. & van Waerbeke, L. 2015, MNRAS, 450, 2857

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

Heavens, A. 2009, ArXiv e-prints

Hetterscheidt, M., Simon, P., Schirmer, M., et al. 2007, A&A, 468, 859

Heymans, C., Brown, M., Heavens, A., et al. 2004, MNRAS, 347, 895

Heymans, C., Grocutt, E., Heavens, A., et al. 2013, MNRAS, 432, 2433

Heymans, C., Van Waerbeke, L., Miller, L., et al. 2012, MNRAS, 427, 146

Hikage, C., Takada, M., Hamana, T., & Spergel, D. 2011, MNRAS, 412, 65

Hilbert, S., Hartlap, J., & Schneider, P. 2011a, A&A, 536, A85

Hilbert, S., Hartlap, J., & Schneider, P. 2011b, A&A, 536, A85

Hildebrandt, H., Erben, T., Kuijken, K., et al. 2012, MNRAS, 421, 2355

Hivon, E., Górski, K. M., Netterfield, C. B., et al. 2002, ApJ, 567, 2

Hoekstra, H. & Jain, B. 2008, Annual Review of Nuclear and Particle Science, 58, 99

Hoekstra, H., Yee, H. K. C., & Gladders, M. D. 2002, ApJ, 577, 595

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

Huff, E. M., Eifler, T., Hirata, C. M., et al. 2014a, MNRAS, 440, 1322

Huff, E. M., Eifler, T., Hirata, C. M., et al. 2014b, MNRAS, 440, 1322

Jarvis, M., Bernstein, G. M., Fischer, P., et al. 2003, AJ, 125, 1014

Joachimi, B. & Bridle, S. L. 2010, A&A, 523, A1

Joachimi, B. & Schneider, P. 2008, A&A, 488, 829

Joachimi, B. & Schneider, P. 2010, ArXiv e-prints: 1009.2024

Joachimi, B., Schneider, P., & Eifler, T. 2008, A&A, 477, 43

Kaiser, N. 1995, ApJ, 439, L1

Kaiser, N. 1998, ApJ, 498, 26

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

Kendall & Stuart. 1960, The Advanced Theory of Statistics Vol 2 (Charles Griffin and Company)

Kenney, J. F. & Keeping, E. S. 1951, Mathematics of Statistics, Part II (Van Nostrand)

Kilbinger, M., Fu, L., Heymans, C., et al. 2013, MNRAS, 430, 2200

Kilbinger, M., Schneider, P., & Eifler, T. 2006, A&A, 457, 15

Kim, A. G., Linder, E. V., Miquel, R., & Mostek, N. 2004, MNRAS, 347, 909

King, L. J. & Schneider, P. 2003, A&A, 398, 23

Knox, L., Scoccimarro, R., & Dodelson, S. 1998, Physical Review Letters, 81, 2004

Komatsu, E., Smith, K. M., Dunkley, J., et al. 2011, ApJS, 192, 18

Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, ArXiv:1110.3193

Marian, L., Smith, R. E., Hilbert, S., & Schneider, P. 2013, MNRAS, 432, 1338

Memari, Y. 2009, Cosmological parameter estimation with QUaD CMB polarization and temperature experiment (Edinburgh Research Archive)

Miller, L., Heymans, C., Kitching, T. D., et al. 2013, MNRAS, 429, 2858

Peacock, J. 1999, Cosmological physics (Cambridge University Press)

Peacock, J. A., Cole, S., Norberg, P., et al. 2001, Nature, 410, 169

Peacock, J. A. & Smith, R. E. 2000, MNRAS, 318, 1144

Peebles, P. J. & Ratra, B. 2003, Reviews of Modern Physics, 75, 559

Pielorz, J., Rödiger, J., Tereno, I., & Schneider, P. 2010, A&A, 514, A79

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2014a, A&A, 571, A16

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2014b, A&A, 571, A16

Plebanski, J. & Krasinski, A. 2006, An Introduction to General Relativity and Cosmology (Cambridge University Press)

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 2002, Numerical Recipes in C++ (Cambridge University Press)

Riess, A. G., Macri, L., Casertano, S., et al. 2011, ApJ, 730, 119

Riess, A. G., Strolger, L.-G., Casertano, S., et al. 2007, ApJ, 659, 98

Sato, M., Hamana, T., Takahashi, R., et al. 2009, ApJ, 701, 945

Sato, M., Takada, M., Hamana, T., & Matsubara, T. 2010, ArXiv e-prints

Schneider, P. 1996, MNRAS, 283, 837

Schneider, P. 2009a, Cosmology Lecture Notes (Argelander-Institute for Astronomy)

Schneider, P. 2009b, Gravitational Lensing Lecture Notes (Argelander-Institute for Astronomy)

Schneider, P., Ehlers, J., & E.E., F. 1999, Gravitational Lenses (Springer)

Schneider, P., Eifler, T., & Krause, E. 2010, A&A, 520, A116

Schneider, P. & Kilbinger, M. 2007, A&A, 462, 841

Schneider, P., Kochanek, C., & Wambsganss, J. 2006, Gravitational Lensing: Strong, Weak and Micro: Saas-Fee Advanced Course 33 (Springer)

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

Schneider, P., van Waerbeke, L., Kilbinger, M., & Mellier, Y. 2002a, A&A, 396, 1

Schneider, P., van Waerbeke, L., Kilbinger, M., & Mellier, Y. 2002b, A&A, 396, 1

Schneider, P., van Waerbeke, L., & Mellier, Y. 2002c, A&A, 389, 729

Schneider, P., van Waerbeke, L., & Mellier, Y. 2002d, A&A, 389, 729

Schramm, T. & Kayser, R. 1995, A&A, 299, 1

Schutz, B. 2003, Gravity from ground up (Cambridge University Press)

Seitz, C. & Schneider, P. 1997, A&A, 318, 687

Seljak, U. 1998, ApJ, 506, 64

Seljak, U. 2000, MNRAS, 318, 203

Semboloni, E., Hoekstra, H., Schaye, J., van Daalen, M. P., & McCarthy, I. G. 2011, MNRAS, 417, 2020

Smith, R. E., Peacock, J. A., Jenkins, A., et al. 2003, MNRAS, 341, 1311

Takada, M. & Jain, B. 2009, MNRAS, 395, 2065

Takahashi, R., Soma, S., Takada, M., & Kayo, I. 2014, ArXiv:1405.2666

Takahashi, R., Yoshida, N., Takada, M., et al. 2011, ApJ, 726, 7

Taylor, A. & Joachimi, B. 2014, MNRAS, 442, 2728

Taylor, A. N., Kitching, T. D., Bacon, D. J., & Heavens, A. F. 2007, MNRAS, 374, 1377

Tegmark, M., Blanton, M. R., Strauss, M. A., et al. 2004, ApJ, 606, 702

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

Troxel, M. A. & Ishak, M. 2014, ArXiv e-prints

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

Van Waerbeke, L., Mellier, Y., Erben, T., et al. 2000, A&A, 358, 30

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

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

Zhang, F. 2005, The Schur Complement and Its Applications (Springer)

Publications

Asgari, Marika & Schneider, Peter. A new data compression method and its application to cosmic shear analysis. In *Astronomy & Astrophysics, Volume 578, id.A50, 10 pp.*, 2015.