Third-order Cosmic Shear Statistics:

Covariance, Nulling, and E/B-mode decomposition

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Xun Shi

aus

Nanjing, China

Bonn, 2011

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der
Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Referent: Prof. Dr. Peter Schneider
2. Referent: Prof. Dr. Cristiano Porciani

Tag der Promotion: 03.08.2011
Erscheinungsjahr: 2012

We shall not cease from exploration
And the end of all our exploring
Will be to arrive where we started
And know the place for the first time.

T.S. Eliot, *Little Gidding*

Contents

Chapter 1

Motivation & Overview

The physical understanding of the Universe has greatly advanced during the last century as a result of joint development of theory and observation. Landmark achievements include the first self-consistent model of the Universe as a whole (Einstein 1917), the standard world models of General Relativity (Friedman 1922; Lemaître 1927), the discovery of the recession of the nearby galaxies (Hubble 1929), the derivation of the Robertson-Walker metric (Robertson 1935; Walker 1937), the measurement of the rotation curves of spiral galaxies (see Rubin et al. 1980) and the mass-to-light ratio of the Coma cluster (Zwicky 1937), the development of the Big Bang theory and nucleosynthesis (Gamow 1946; Alpher et al. 1948; Alpher & Herman 1948, 1950), the discovery of the Cosmic Microwave Background (Penzias & Wilson 1965), the discovery of the Large-Scale Structure (e.g. Davis et al. 1982), the discovery of the accelerating expansion of the Universe (Riess et al. 1998), and many others. All these discoveries point to a consistent paradigm: an expanding Universe described in the framework of General Relativity which is dominated by cold dark matter and dark energy in terms of energy. This paradigm, summarized as the Lambda-Cold Dark Matter (ΛCDM) model, has a set of free parameters which can be determined from cosmological observations. So far the different observations show a remarkable consistency in their determined parameter values.

Despite its empirical success, the ΛCDM model is still not complete in the theoretical sense. The two exotic components it assumes, namely dark matter and dark energy, are not yet explained by our current knowledge of fundamental physics. The dark matter behaves like common baryonic matter in gravitational interaction, but cannot emit any electromagnetic radiation. The observational fact that galaxies form before structures of larger physical scales restricts the dark matter to be 'cold', i.e. with a thermal velocity much less than the speed of light. Theoretically there are massive elementary particles which are considered candidates of the cold dark matter, but they are not yet detected experimentally. The dark energy, which is required to explain the late-time acceleration of the expansion of the Universe, is even less understood. In particular it has a negative pressure, which poses a conceptual problem for our understanding.

One possible explanation for dark energy is Einstein's cosmological constant. Many exotic forms of matter, e.g. the quintessence, have also been proposed as dark energy candidates. In either case the picture is far from complete. Moreover, the theory of General Relativity, which most of the current explanations are based on, may fail at cosmic scales. This situation makes the study of the nature of dark energy one of the most important problems in cosmology, and even in all physical science.

Due to a lack of compelling theoretical clues, observational studies of dark energy are especially valuable. It is expected that the question if and how dark energy evolves with time will play a decisive role in judging the possible explanations. To reach this answer, one needs to determine the

cosmological parameters and their time evolution to percent level accuracy or even higher precision.

The current values of the cosmological parameters are mainly constrained by the observation of the Cosmic Microwave Background (CMB), aided by the Hubble parameter measured in the local Universe [1]. Other observations, such as that of the Large-Scale Structure (LSS), Type Ia supernovae, the Lyman-alpha forest, and weak gravitational lensing effects, are needed to improve the precision of the parameter determination. Particularly, each of these observations reflects different physical processes and thus is sensitive to different combinations of cosmological parameters. It is therefore essential to combine them to check the consistency as well as to break degeneracies among different parameters.

Concerning the constraint of the time evolution of dark energy, four observational techniques are considered the most promising. They are the Baryon Acoustic Oscillations which are observed in surveys of the spatial distribution of galaxies, galaxy cluster surveys, surveys of Type Ia supernovae, and weak lensing surveys (Albrecht et al. 2006). These observational techniques can all probe the time evolution of dark energy through the expansion history of the Universe. Additionally, galaxy cluster surveys and weak lensing surveys also provide information through the history of structure formation.

This thesis is concerned with the weak lensing effect, which is considered to be potentially the most powerful one among all dark energy probes. In a weak lensing survey one uses the gravitational shear, which is the coherent shape distortion of galaxies, as the observable effect to probe the statistical properties of the underlying matter density field. The forthcoming large-field multicolor imaging surveys (e.g. DES[2], KIDS[3], EUCLID[4], etc) will obtain photometric redshift and shape information of a huge number of galaxies. This will render weak lensing a higher statistical power compared to other probes. Such constraining power can be further enhanced by the use of higher-order statistics. The higher-order statistics, contrary to the second-order (two-point) ones, can probe non-Gaussian signatures in the matter density field, and thus are necessary tools to fully exploit the wealth of information on small, non-linear scales. Our work focuses on the lowest order of them – the third-order (three-point) statistics.

The performance of weak lensing surveys depends critically on the control of systematic errors. The major sources of systematics lie in the measurement process, specifically, in galaxy shape measurement and the determination of the galaxy redshifts. In addition to them, there are systematics originating from astrophysical processes, the most worrisome one of them being the intrinsic-shear alignment effect. How much weak lensing surveys are affected by these systematics, and how well the systematics can be controlled, is still uncertain to a large extent. This situation emphasizes the importance of studying the systematic errors on one hand, and on the other hand makes it a necessity to perform systematics checks on the lensing signal. The latter can be realized by doing an E/B-mode decomposition, namely separating the lensing signal into an electric field-like E-mode component, and a magnetic field-like B-mode component. Since the physical lensing signal has only E-mode components to the first order while most of the systematic effects do not make this distinction, possible B-mode components in the data provide a valuable check on the level of systematics.

To compare the three-point (3-pt) statistics estimated from a weak lensing survey to that predicted by theory, an unavoidable step is to relate the shear 3-pt statistics to that of the underlying matter density field. The currently available approach uses the relation between the Fourier space 3-pt statistics for the matter density field and the configuration space shear 3-pt statistics which can

[1] The Hubble Space Telescope Key Project, http://www.ipac.caltech.edu/H0kp

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

[3] http://www.astro-wise.org/projects/KIDS/

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

be directly measured from a survey catalog. This theoretical relation, however, contains very oscillatory functions and is thus hard to handle numerically. Therefore, it is helpful to relate the two directly in configuration space, where the corresponding functions are expected to be smooth.

The work in this thesis is committed to these four questions concerning weak lensing 3-pt statistics:

- How much information is contained in weak lensing 3-pt statistics?
- How to relate the 3-pt shear observables to the configuration space statistics of the underlying matter density field?
- How to perform an E/B-mode decomposition for weak lensing 3-pt statistics?
- How to deal with the intrinsic-shear alignments for weak lensing 3-pt statistics?

After introducing the theoretical background in Chap. 2 and Chap. 3, we derive some fundamental relations between weak lensing 3-pt statistics, including those relating the shear observables to the configuration space statistics of the underlying matter density field in Chap. 4. These relations also allow us to formulate the condition of E/B-mode decomposition at the 3-pt level. In Chap. 5 we use a more rigorous approach than Hu (2000) to derive an expression for the covariance matrix of the bispectrum, the Fourier counterpart of the 3-pt correlation function. This presents a theoretical way of quantifying the information content in lensing 3-pt statistics. In Chap. 6 we generalize the nulling technique, a method to control the intrinsic-shear alignment, to the 3-pt level, and thereby provide the first method to control the corresponding 3-pt systematics. A summary of the work presented in this thesis together with suggestions for possible future research is given in Chap.7.

Chapter 2

The cosmological standard model and the large-scale structure

During the past decades, a physical view of the Cosmos summarized as the ΛCDM model, often called the "standard model" as well, has been developed and widely accepted as the modern picture of the Cosmos. This model consistently explains the expansion of the Universe, the formation and growth of structure in the Universe, the existence and anisotropies of the Cosmic Microwave Background, as well as the abundances of chemical elements. In this chapter we will explain the relevant theoretical knowledge in the standard model, on which this thesis is based. For a more detailed view of the model we refer to Peacock (1999), Bernardeau et al. (2002b), and Dodelson (2003).

2.1 The homogeneous and isotropic Universe

2.1.1 Friedmann world model

Considering how the Universe evolves dynamically as a whole, one can, to first order, simplify it as a homogeneous and isotropic medium of matter and energy. On the other hand, the dominating source of interaction in the Universe at large scales is gravitation, due to its long-range and non-cancelling property. These together explain the two theoretical pillars of the cosmological standard model: the cosmological principle which assumes that the Universe is homogeneous and isotropic on large scales, and Einstein's General Relativity (GR) as the theory of gravitation.

The cosmological principle corresponds to a point of view that our observational location in the Universe is in no way unique or special. During the 1920s and 1930s when the theoretical foundations of the standard cosmological model were developed, observational knowledge of the Universe was also exploding: the extragalactic nature of spiral nebulae had been established (1920s), basic types of galaxies were classified (Hubble, 1926), and the velocity-distance relation for nearby galaxies was determined for the first time (Hubble, 1929). These observations supported the philosophy of the cosmological principle, however they did not provide direct evidence for large-scale homogeneity and isotropy. Nowadays large galaxy redshift surveys (e.g. the Sloan Digital Sky Survey[1] and the 2dF Galaxy Redshift Survey[2]) and the mapping of the Cosmic Microwave Background (e.g.

[1] http://www.sdss.org
[2] http://www2.aao.gov.au/2dfgrs

the Wilkinson Microwave Anisotropy Probe[3]) have probed unprecedentedly large volumes of the Universe. Their results suggest that the Universe is indeed homogeneous and isotropic on scales larger than ∼200 Mpc, and thus strongly support the idea of the cosmological principle.

In General Relativity, gravity is regarded as a geometric property of space and time, or *spacetime*, which is mathematically represented by a (3+1)-dimensional Riemannian manifold. The spacetime gets distorted in the presence of matter according to the Einstein field equation of General Relativity, which reads

$$G_{\mu\nu} + \Lambda\, g_{\mu\nu} = -\frac{8\pi G}{c^4} T_{\mu\nu}\,. \tag{2.1}$$

The left hand side of the equation, with $G_{\mu\nu}$ being the Einstein Tensor and Λ the cosmological constant, describes the geometry of the spacetime. It is a non-linear function of the metric $g_{\mu\nu}$ and its first and second derivatives. The right-hand side of the equation describes the matter distribution, with $T_{\mu\nu}$ being the energy-momentum tensor. The high degree of non-linearity is the major challenge in solving the Einstein field equation. Simplifying assumptions are often required, e.g. the cosmological principle in the case of cosmology.

The cosmological principle, i.e. the assumption of homogeneity and isotropy on large scales, simplifies both sides of the Einstein field equation. Under this assumption, the metric can be written in the form of the Robertson-Walker metric,

$$\begin{aligned} \mathrm{d}s^2 &:= g_{\mu\nu}\,\mathrm{d}x^\mu \mathrm{d}x^\nu \\ &= c^2\mathrm{d}t^2 - a^2(t)\left[\mathrm{d}\chi^2 + f_\mathrm{K}^2(\chi)\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2\right)\right], \end{aligned} \tag{2.2}$$

where the expansion of the Universe has been accounted for by the cosmic scale factor $a(t)$. The comoving distance χ between comoving objects, i.e. objects whose movements are caused only by the expansion of the Universe, does not change over time. The comoving angular diameter distance $f_K(\chi)$, which is the radial comoving distance corresponding to a solid angle $\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2\right)$, takes the form

$$f_K(\chi) = \begin{cases} K^{-1/2}\sin(K^{1/2}\chi) & (K>0) \\ \chi & (K=0) \\ (-K)^{-1/2}\sinh[(-K)^{1/2}\chi] & (K<0)\,. \end{cases} \tag{2.3}$$

One can see that when the curvature signature K equals zero, the comoving angular diameter distance $f_K(\chi)$ is additive, suggesting that the Universe is Euclidean. When $K > 0$ ($K < 0$), $f_K(\chi)$ takes a trigonometric (hyperbolic) form, corresponding to a closed (open) Universe.

On the other hand, the cosmological principle implies that the matter content in the Universe can be described, to first order, by a uniform ideal fluid whose density ρ and pressure P depends only on time. Thus the energy-momentum tensor can be reduced to

$$T_{\mu\nu} = \left(\rho + \frac{P}{c^2}\right) U_\mu U_\nu - g_{\mu\nu} P\,, \tag{2.4}$$

where U_μ is the four-velocity.

Inserting (2.2) and (2.4) into the Einstein field equation (2.1), one can obtain the Friedmann equations

$$\begin{aligned} \left(\frac{\dot{a}}{a}\right)^2 &= \frac{8\pi G}{3}\rho(t) - \frac{Kc^2}{a^2(t)} + \frac{\Lambda}{3}\,, \\ \frac{\ddot{a}}{a} &= -\frac{4\pi G}{3}\left(\rho(t) + \frac{3p(t)}{c^2}\right) + \frac{\Lambda}{3}\,, \end{aligned} \tag{2.5}$$

[3] http://map.gsfc.nasa.gov

which gives a description of the average dynamical behavior of the Universe. The next step is to specify the r.h.s. of (2.5) with the knowledge of the energy/matter content of the Universe.

The cosmological constant Λ was originally added by Einstein to the field equation (2.1) as a geometric term in order to allow a static universe as well as eliminating the problem of boundary conditions at infinity. As Einstein later put it himself, it is 'not justified by our actual knowledge of gravitation' but merely 'logically consistent', and 'detracts from the formal beauty of the theory' (Longair 2006).

In the current framework of cosmology, it is more common to consider Λ as an energy component with density $\rho_\Lambda = \Lambda c^2/8\pi G$ and pressure $P_\Lambda = -\rho_\Lambda c^2$. This so-called *dark energy* is one of the most important energy contributions to our Universe, along with a matter component ρ_m and a radiation component ρ_r. The matter component includes the familiar baryonic matter as well as a dark matter component whose existence is inferred through its gravitational interaction with the visible matter. Sometimes these two components are listed separately as ρ_b and ρ_DM. The radiation component is dominated by photons, but generally speaking it contains all relativistic particles.

The evolution of the density of these components as the Universe expands can be studied by the adiabatic equation

$$\frac{\mathrm{d}\left(a^3\rho c^2\right)}{\mathrm{d}t} + P\frac{\mathrm{d}a^3}{\mathrm{d}t} = 0\,, \tag{2.6}$$

which can be derived from (2.5). Characterizing each component by its *equation of state*

$$P = w_\mathrm{eos}\,\rho c^2\,, \tag{2.7}$$

and making the ansatz $\rho \propto a^n$, one obtains from (2.6)

$$\rho(a) = \rho_0\; a^{-3(w_\mathrm{eos}+1)} \tag{2.8}$$

where '0' denotes the value in the present time, and the scale factor for the present Universe has been specified to be 1.

The matter component has zero pressure, i.e. $w_\mathrm{eos} = 0$. According to (2.8), its density evolves as $\rho_\mathrm{m} \propto a^{-3}$. The radiation component has $w_\mathrm{eos} = 1/3$ and thus evolves as $\rho_\mathrm{r} \propto a^{-4}$. The dark energy, if it is taken to be the cosmological constant with $w_\mathrm{eos} = -1$, then its density ρ_de stays constant as the Universe expands. In a wider sense, the name 'dark energy' is used to denote the dominant repulsive component at the present time which is required by the observation of the recent accelerated expansion of the spacetime. In this sense the w_eos of dark energy is only required to be smaller than $-1/3$ at the present time according to (2.5), and can in principle vary with time. A frequently adopted parametrization of the dark energy w_eos is $w_\mathrm{eos} = w_0 + (1-a)w_a$ (e.g. Esposito-Farèse & Polarski 2001; Linder 2003; Albrecht et al. 2006). Obviously, the cosmological constant corresponds to $w_0 = -1$ and $w_a = 0$.

With these we can write the Friedmann equation as

$$\left(\frac{\dot{a}}{a}\right)^2 =: H^2(a) = \frac{8\pi G}{3}\left[\rho_{\mathrm{m},0}\; a^{-3} + \rho_{\mathrm{r},0}\; a^{-4} + \rho_{\mathrm{de},0}\; a^{-3[w_0+(1-a)w_a+1]} - Kc^2a^{-2}\right], \tag{2.9}$$

in which we have defined the Hubble parameter $H = \dot{a}/a$, which describes the expansion rate of the Universe.

In order to non-dimensionalize this equation, we define a critical density

$$\rho_\mathrm{cr}(a) = \frac{3H^2(a)}{8\pi G} \tag{2.10}$$

which is the total density the Universe has in order to be spatially flat, i.e. with $K = 0$. Then we can define dimensionless density parameters

$$\Omega_x := \frac{\rho_{x,0}}{\rho_{\rm cr}(a_0)} = \frac{8\pi G\rho_{x,0}}{3H_0^2}\,, \tag{2.11}$$

where x could be m (matter), b (baryon), DM (dark matter), r (radiation), de (dark energy), or total (matter plus dark energy plus radiation). The curvature term could also be written in terms of these density parameters by letting $a = 1$ in (2.9), as

$$Kc^2 = H_0^2(\Omega_{\rm total} - 1)\,. \tag{2.12}$$

Substituting (2.10), (2.11), and (2.12) into (2.9), we obtain

$$H^2(a) = H_0^2\left[\Omega_{\rm m}a^{-3} + \Omega_{\rm r}a^{-4} + (1-\Omega_{\rm total})a^{-2} + \Omega_{\rm de}a^{-3[w_0+(1-a)w_a+1]}\right]. \tag{2.13}$$

With this equation we have described the expansion of the Universe as a function of a few quantities (the Ω's and H_0), the values of which can be considered as free parameters in a cosmological model.

We can also see from (2.13) that the early expansion history of the Universe was dominated by radiation. The transition from radiation to matter dominance happens at

$$a_{\rm eq} = \frac{\Omega_{\rm r}}{\Omega_{\rm m}} \simeq 3\times10^{-4}\,. \tag{2.14}$$

After that the expansion was matter-dominated. Only very recently ($a \simeq 1$) dark energy became the dominating enery component of the Universe and began to play a major role in the expansion of the Universe.

2.1.2 Cosmological redshift and distances

In a Euclidean space there exists a unique way to specify a 'distance' between two objects, but this is not the case for an expanding spacetime as our Universe. In such an expanding universe, the distance between objects with a fixed comoving separation is constantly changing. When we observe a distant object, we look both out in distance and back in time. The two most common ways of measuring the distance of a light source, namely comparing the measured angle it subtends to its intrinsic size and comparing the measured flux to its intrinsic luminosity, give different results in an expanding universe. Nevertheless they remain important distance measures. Before explaining them in detail we need to introduce the concept of redshift.

Cosmological redshift

The redshift z of a photon describes the change in its wavelength λ with respect to the rest frame wavelength λ_0 when it was emitted,

$$z := \frac{\lambda - \lambda_0}{\lambda_0}\,. \tag{2.15}$$

In our Universe, all photons are subjected to a redshift caused by the expansion of the spacetime, called the cosmological redshift. Quantatively, as light travels along null geodesics, i.e. $\mathrm{d}s^2 = 0$, for them $c^2\mathrm{d}t^2 = a^2(t)\mathrm{d}\chi^2$, which gives

$$\frac{\mathrm{d}t}{a(t)} = -\frac{\mathrm{d}\chi}{c}\,, \tag{2.16}$$

where the minus sign on the r.h.s. is taken to indicate that the light is traveling towards $\chi = 0$.

Now consider a comoving light source emitting a photon with rest-frame frequency ν_0 at time t_0 and comoving distance χ_0, and the photon is observed at t_1 and χ_1 with frequency ν_1 by a comoving observer. Since the comoving distance between the light source and the observer

$$\chi_0 - \chi_1 = c \int_{t_0}^{t_1} \frac{\mathrm{d}t}{a(t)} \tag{2.17}$$

is a constant regardless of when the photon was emitted, one reaches the result

$$\frac{\mathrm{d}t_0}{a(t_0)} = \frac{\mathrm{d}t_1}{a(t_1)} \tag{2.18}$$

by taking the derivative of (2.17) with respect to t_0 while keeping $t_1 - t_0$ fixed.

Equation (2.18) can be seen as an expression of the time dilation $\mathrm{d}t_1/\mathrm{d}t_0$ between two comoving observers at t_0 and t_1, respectively. The time-dilation affects the wavelength of the photon as well,

$$\frac{\mathrm{d}t_1}{\mathrm{d}t_0} = \frac{a(t_1)}{a(t_0)} = \frac{\lambda_1}{\lambda_0} = 1 + z_1 \,. \tag{2.19}$$

Thus we obtain the relation between the redshift of photons caused by the expansion of the Universe and the scale factor of the Universe. In the standard model, the Universe has always been expanding, therefore the higher the redshift, the earlier the photon was emitted. This allows the use of redshift itself as an indicator of time and distance of the light source.

Horizon

The horizon size r_h is defined to be the comoving size of the observable Universe. Letting t_0 be the age of the Universe, and using (2.16), r_h can be expressed as

$$r_\mathrm{h}(t_0) = \int_0^{t_0} \frac{c\,\mathrm{d}t}{a(t)} = \int_0^{a(t_0)} \frac{c\,\mathrm{d}a}{a^2 H(a)} \,. \tag{2.20}$$

Expectedly, r_h depends on the expansion history of the Universe. Since the speed of light is also the upper limit of signal transmission speed, the horizon size is also the size of the region with causal connections.

Angular diameter distance

The angular diameter distance D_ang is of great importance for this thesis. Suppose a source with a redshift of z has an intrinsic transverse size R and is observed to have an angular diameter of δ (in radians), its angular diameter distance is defined to be

$$D_\mathrm{ang}(z) := \frac{R}{\delta} = a(z) f_K(\chi(z)) \,, \tag{2.21}$$

where $\chi(z)$ is the comoving distance of the source, f_K is the comoving angular diameter distance whose form is given in (2.3). The second equation in (2.21) follows from the metric (2.2). With a description of the expansion of the Universe (2.13), we can specify the functional form of $\chi(z)$, or more generally, the comoving distance between sources at two different redshifts

$$\begin{aligned} \chi(z_1, z_2) &= \frac{c}{H_0} \int_{a(z_2)}^{a(z_1)} \left\{ \Omega_\mathrm{r} + a\Omega_\mathrm{m} + a^2 \left(1 - \Omega_\mathrm{total}\right) + a^{-3[w_0 + (1-a)w_a + 1] + 4} \Omega_\mathrm{de} \right\}^{-1/2} \mathrm{d}a \\ &= \chi(z_2) - \chi(z_1) \,. \end{aligned} \tag{2.22}$$

For a spatially flat universe ($K = 0$), the comoving angular diameter distance $f_K(\chi)$ equals the comoving distance χ and is additive. However, this is not true for the angular diameter distance $D_{\rm ang}$ itself. According to the original definition, the angular diameter distance of a source at redshift z_2 as seen by an observer at $z_1 < z_2$ is

$$D_{\rm ang}(z_1, z_2) = a(z_2) f_K \left[\chi(z_1, z_2)\right] . \tag{2.23}$$

Another special property of angular diameter distance is that it does not increase monotonously with increasing redshift. In a standard ΛCDM cosmology it turns over at a redshift around unity, which means more distant objects actually appear larger in angular size.

Luminosity distance

The luminosity distance $D_{\rm lum}$ is obtained by relating the bolometric flux $F_{\rm bol}$ of a source at z_2 observed at $z_1 < z_2$ to its intrinsic bolometric luminosity L,

$$D_{\rm lum}(z_1, z_2) := \sqrt{\frac{L}{4\pi F_{\rm bol}(z_1, z_2)}} . \tag{2.24}$$

The surface brightness S of the source at z_2 is related to its intrinsic bolometric luminosity L as

$$S = \frac{L}{4\pi (R/2)^2} , \tag{2.25}$$

where R is the transverse size of the source, same as in (2.21). The bolometric flux $F_{\rm bol}$ of the source observed at z_1 is related to the surface brightness of the source at z_2, as

$$F_{\rm bol}(z_1, z_2) = S \frac{(1+z_1)^4}{(1+z_2)^4} \left(\frac{\delta(z_1, z_2)}{2}\right)^2 , \tag{2.26}$$

where δ is the angular diameter (in radians) of the source at z_2 observed at z_1, and we have used the fact that the surface brightness S of a receding light source is reduced by a factor $(1+z)^{-4}$. Inserting (2.25) and (2.26) into (2.24), we obtain the relation between $D_{\rm lum}$ and $D_{\rm ang}$ (see e.g. Hogg 1999)

$$D_{\rm lum}(z_1, z_2) = \frac{(1+z_2)^2}{(1+z_1)^2} \frac{R}{\delta} = \frac{(1+z_2)^2}{(1+z_1)^2} D_{\rm ang}(z_1, z_2) = \frac{1+z_2}{(1+z_1)^2} f_K \left[\chi(z_1, z_2)\right] . \tag{2.27}$$

2.2 Formation of the large scale structure

In contrast to the homogeneity on large scales, the Universe we observe today has an abundance of structures on smaller scales. Fig. 2.1 presents the distribution of about 10^5 galaxies observed in the local Universe. One can clearly see in it clusters of galaxies, filaments and voids, which are the principal elements of the LSS. According to the standard cosmological scenario, the Universe began in a much more homogeneous state, and these structures we observe today are formed via the amplification of primordial quantum fluctuations by gravitational instability (see e.g. Peebles 1980). Observational support for this include the temperature fluctuations of the CMB which are found to be five orders of magnitude smaller than the mean temperature, suggesting a high degree of homogeneity at the epoch CMB photons were emitted ($z \simeq 1100$). Additionally, from various observations of the Universe at redshift of order unity, e.g. those of galaxies, galaxy clusters and the Lyman-alpha forest, one can see a distinct growth of structures towards lower redshift.

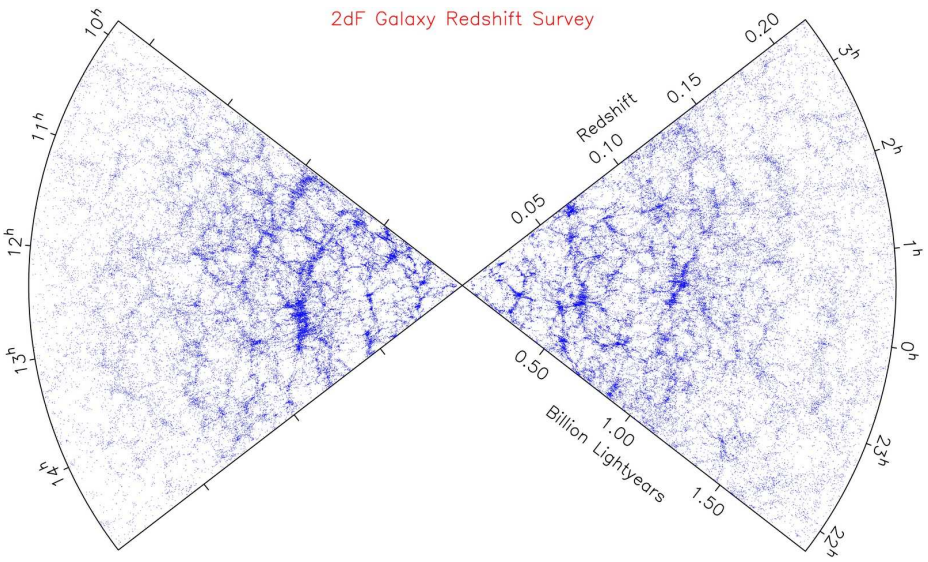


Figure 2.1: Projected distribution of galaxies within two surveyed patches of the observable Universe. The distribution of galaxies reflect indirectly the distribution of the dark matter. The Earth is at the center of the image. Credit: 2dF team

The process of structure formation is sensitive to the evolution of spacetime, but is not considered to have a substantial back-reaction on it. The latter is instead driven by the uniform mean field, as described in the previous chapter. Consequently, the structure formation process can act as a probe of the uniform mean field which can be parametrized by a couple of cosmological parameters.

In this section we consider the formation and evolution of structures under gravity. The dark matter plays a dominant role in this process and is therefore our major concern here.

2.2.1 Vlasov equation

Two physical scales are of great importance in the theory of cosmic structure formation. One is the horizon size r_{h}, which confines the region with possible casual interaction. Fluctuations with size $r \geq r_{\mathrm{h}}$ have to be studied in the framework of General Relativity. The other physical scale deals with the amplitude of the fluctuations. When the amplitude of the fluctuations of a certain size is much smaller than that of the mean field, perturbative methods can be used and the density inhomogeneities are well described by linear differential equations. However, below a certain scale called the non-linear scale r_{nl}, the linear approximation breaks down. At scales with $r \leq r_{\mathrm{nl}}$, no precise analytical method exists and the growth of fluctuations as well as the distribution of matter are usually studied with simulations.

We will focus first on perturbations that are well in between these two scales. In this case, Newtonian gravitational interaction and linear perturbation theory hold to be good approximations, and the evolution of dark matter phase-space density $f(\boldsymbol{r}, \boldsymbol{u}, t)$ can be described by the Vlasov equation:

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + \boldsymbol{u}\frac{\partial f}{\partial \boldsymbol{r}} - \nabla\phi\frac{\partial f}{\partial \boldsymbol{u}} = 0\,. \tag{2.28}$$

Here, $\boldsymbol{r}$ and $\boldsymbol{u}$ are the physical position and velocity, and ϕ is the Newtonian potential given by the modified Poisson equation:

$$\nabla^2\phi(\boldsymbol{r},t) = 4\pi G \int \mathrm{d}^3u\; m f(\boldsymbol{r},\boldsymbol{u},t) - \Lambda\,. \tag{2.29}$$

The first term in (2.29), with m indicating the mass of the matter particle, is the self-gravitation term. The second term is the acceleration provided by the cosmological constant, whose form is chosen to be consistent with the Friedmann equation (2.5). In this form of the Poisson equation we have assumed the dynamical effect of radiation to be negligible. This is valid after the Universe switched from radiation-dominated to matter-dominated at redshift $z_{\rm eq} = a_{\rm eq}^{-1} - 1 \simeq 3300$. We will discuss the situation in the radiation dominated era separately.

The Vlasov equation is also called the collisionless Boltzmann equation, which is the Boltzmann equation without the collision term. It conserves the phase-space density of dark matter.

2.2.2 Ideal fluid approximation

Due to the complicated non-linear structure and the high dimensionality of the Vlasov equation, it not only hinters analytical solutions but is also hard to analyze numerically. As a further simplification, we treat the matter as an ideal fluid with zero pressure. This fluid approximation is valid for the scales we are focusing on. However, since dark matter is collision-less and thus allowing multiple streams instead of a well-defined velocity field $\boldsymbol{u}(\boldsymbol{r})$, this simplification is sure to break down at non-linear scales with $r \leq r_{\rm nl}$ where multiple streams become important.

With the fluid approximation one considers the zeroth and the first momentum moments of the phase space distribution $f(\boldsymbol{r},\boldsymbol{u},t)$. The zeroth order moment gives the local mass density field ρ, and the first moment normalized by the zeroth moment yields the velocity of the flow $\bar{\boldsymbol{u}}$:

$$\int \mathrm{d}^3u\; m f(\boldsymbol{r},\boldsymbol{u},t) =: \rho(\boldsymbol{r},t)\,, \tag{2.30}$$

$$\frac{\int \mathrm{d}^3u\; m\boldsymbol{u} f(\boldsymbol{r},\boldsymbol{u},t)}{\int \mathrm{d}^3v\; m f(\boldsymbol{r},\boldsymbol{u},t)} =: \bar{\boldsymbol{u}}(\boldsymbol{r},t)\,. \tag{2.31}$$

Taking the zeroth and first momentum moment of the Vlasov equation (2.28) and inserting the definitions (2.30) and (2.31), one obtains

$$\frac{\partial\rho}{\partial t} + \nabla_r\cdot(\rho\bar{\boldsymbol{u}}) = 0 \qquad \text{Continuity equation}\,, \tag{2.32}$$

$$\frac{\partial\bar{\boldsymbol{u}}}{\partial t} + (\bar{\boldsymbol{u}}\cdot\nabla_r)\bar{\boldsymbol{u}} = -\nabla_r\phi \qquad \text{Euler equation}\,. \tag{2.33}$$

The Poisson equation now reads

$$\nabla^2\phi(\boldsymbol{r},t) = 4\pi G\rho - \Lambda\,. \tag{2.34}$$

As one can see, they are just the equations for an ideal fluid of zero pressure. There exists no general analytic solution to this set of equations. Nevertheless several perturbative techniques are available (see e.g. Zel'Dovich 1970; Bernardeau et al. 2002b; Szapudi & Kaiser 2003; Crocce & Scoccimarro 2006) which allow an analytic treatment in the linear and weakly non-linear regimes. Here we introduce the Eularian perturbation technique.

2.2.3 Density contrast and peculiar velocity field

To study the growth of inhomogeneities in the Universe, we subtract the mean field, and consider small perturbations to the density and velocity fields. We do so in the comoving coordinates $\boldsymbol{x} = \boldsymbol{r}/a(t)$:

$$\rho(\boldsymbol{r},t) = \hat{\rho}\left(\frac{\boldsymbol{r}}{a(t)},t\right) = \hat{\rho}\left(\boldsymbol{x},t\right) = \bar{\hat{\rho}} + \delta\hat{\rho}\left(\boldsymbol{x},t\right)\,, \tag{2.35}$$

$$\bar{\boldsymbol{u}}(\boldsymbol{r},t) = \hat{\bar{\boldsymbol{u}}}\left(\frac{\boldsymbol{r}}{a(t)},t\right) = \hat{\bar{\boldsymbol{u}}}\left(\boldsymbol{x},t\right) = \dot{a}\boldsymbol{x} + \boldsymbol{v}(\boldsymbol{x},t)\,, \tag{2.36}$$

where $\bar{\hat{\rho}}$ denotes the mean density of the matter component in the Universe, $\dot{a}\boldsymbol{x}$ is the Hubble flow, and $\boldsymbol{v}$ is the peculiar velocity. According to the previous chapter, we have

$$\bar{\hat{\rho}}(a) = \rho_{\mathrm{m}0}a^{-3} = \frac{3H_0^2\Omega_{\mathrm{m}}}{8\pi G a^3}\,. \tag{2.37}$$

We further define two useful quantities, the density contrast δ and the comoving gravitational potential Φ:

$$\delta(\boldsymbol{x},t) := \frac{\delta\hat{\rho}\left(\boldsymbol{x},t\right)}{\bar{\hat{\rho}}}\,, \tag{2.38}$$

$$\Phi(\boldsymbol{x},t) := \phi\left(a\boldsymbol{x},t\right) + \frac{a\ddot{a}}{2}|\boldsymbol{x}|^2\,. \tag{2.39}$$

Putting all these into the set of fluid equations (2.32)-(2.34) and making use of the Friedmann equations (2.5), we obtain the set of dynamical equations which govern the evolution of the density contrast and the peculiar velocity field:

$$\frac{\partial\delta}{\partial t} + \frac{1}{a}\nabla_x\cdot\left[(1+\delta)\boldsymbol{v}\right] = 0\,, \tag{2.40}$$

$$\frac{\partial\boldsymbol{v}}{\partial t} + \frac{\dot{a}}{a}\boldsymbol{v} + \frac{1}{a}\left(\boldsymbol{v}\cdot\nabla_x\right)\boldsymbol{v} = -\frac{1}{a}\nabla_x\Phi\,, \tag{2.41}$$

$$\nabla_x^2\Phi = \frac{3H_0^2\Omega_{\mathrm{m}}}{2a}\delta\,. \tag{2.42}$$

The peculiar velocity field is a vector field, which can be fully described by its divergence $\nabla_x\cdot\boldsymbol{v} =: \theta$ and its vorticity $\nabla\times\boldsymbol{v} =: \boldsymbol{w}$. Since the source term of (2.41) is a gradient, one can easily see that the equation of motion for $\boldsymbol{w}$ does not have a source, and that in the linear regime $\boldsymbol{w}$ decays away as the Universe expands. So we will focus on the evolution of the density contrast δ and velocity divergence θ.

2.2.4 Eularian perturbation theory

According to the idea of perturbation theory, we write δ and θ as

$$\delta(\boldsymbol{x},t) = \sum_{n=1}^{\infty}\delta^{(n)}(\boldsymbol{x},t),\;\; \theta(\boldsymbol{x},t) = \sum_{n=1}^{\infty}\theta^{(n)}(\boldsymbol{x},t)\,, \tag{2.43}$$

where $\delta^{(1)}$ and $\theta^{(1)}$ are linear in the initial density field, $\delta^{(2)}$ and $\theta^{(2)}$ are quadratic, and so on.

Linear growth

Linearizing (2.40) and (2.41), then combining (2.40)-(2.42) and eliminating θ, we obtain a second-order linear differential equation for the linear density contrast $\delta^{(1)}$:

$$\frac{\partial^2\delta^{(1)}}{\partial t^2} + \frac{2\dot{a}}{a}\frac{\partial\delta^{(1)}}{\partial t} - \frac{3H_0^2\Omega_{\rm m}}{2a^3}\delta^{(1)} = 0\,. \tag{2.44}$$

This equation does not contain derivatives with respect to spatial coordinates $\boldsymbol{x}$, nor does $\boldsymbol{x}$ appear explicitly in the equation. Thus one can separate the t and $\boldsymbol{x}$ dependence in the solution, and write it in a general form

$$\delta^{(1)}(\boldsymbol{x}, t) = D_+(t)\Delta_+(\boldsymbol{x}) + D_-(t)\Delta_-(\boldsymbol{x})\,, \tag{2.45}$$

with $D_\pm(t)$ being the two linearly independent solutions of

$$\ddot{D} + \frac{2\dot{a}}{a}\dot{D} - \frac{3H_0^2\Omega_{\rm m}}{2a^3}D = 0\,. \tag{2.46}$$

One solution of it is found to be the Hubble parameter $H(t)$. In a matter-dominated expanding universe, $H(t)$ decreases with time and thus represents a decreasing solution which we choose to be D_-. Since we are interested in the growth of structure, the growing solution D_+ is of greater concern. It can be constructed with the aid of the form of D_-, and is found to be

$$D_+(t) \propto H(t)H_0^2\int_0^t \frac{{\rm d}t'}{a^2(t')H^2(t')}\,. \tag{2.47}$$

This function D_+ is called the *growth factor*. According to (2.45), it describes the linear growth of the density contrast. It is usually normalized to $D_+(t_0) = 1$, i.e. it has value unity at the present time.

An exact scaling $D_+(t) = a(t)$ is found for an Einstein-de-Sitter (EdS) Universe ($\Omega_{\rm total} = \Omega_{\rm m} = 1$). In general, finding the form of $D_+(t)$ requires numerical integration.

Transfer function

So far we have dealt only with the growth of subhorizon matter perturbations when the Universe is *not* dominated by radiation. For superhorizon perturbations, the Newtonian description breaks down. A generalized treatment (see e.g. Dodelson 2003) shows that they grow as $\delta \propto a^2$ in the radiation-dominated era, and $\delta \propto a$ in the matter-dominated era. On the other hand, sub-horizon matter perturbations cease to grow in the radiation-dominated era due to the suppression by the radiation-dominated expansion of the Universe.

Generally speaking, perturbations of different comoving sizes L grow differently depending on when they enter the horizon. We denote the scale factor at horizon entering as $a_{\rm enter}(L)$. To put it more correctly, $a_{\rm enter}(L)$ is the scale factor when the horizon size $r_{\rm h}$ expands to the size of the perturbation L

$$r_{\rm h}(a_{\rm enter}) = L\,. \tag{2.48}$$

The growth of small-scale perturbations which enter the horizon before matter domination are suppressed compared to perturbations of larger scales, as depicted in Fig. 2.2.

Besides the different horizon-entry time, there exist several other effects that can break the scale-independence in the linear growth of the structure. For example, in case the Universe is dominated by hot dark matter (HDM), the mean free path of structure-building particles would be significant,

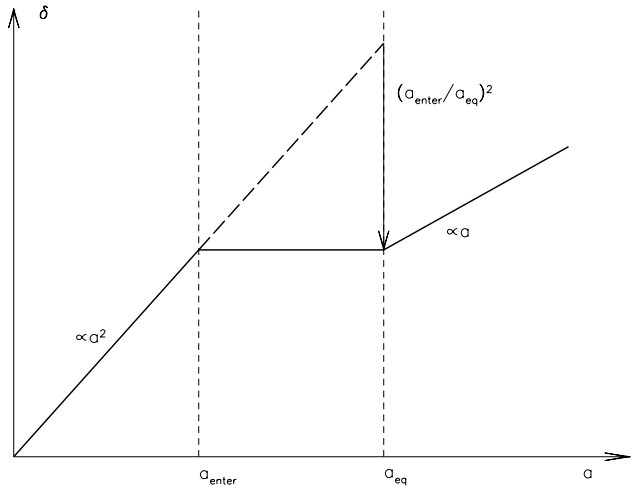


Figure 2.2: Sketch illustrating the growth of a perturbation which enters the horizon before matter domination. The growth is suppressed in the radiation-dominated phase by a factor of $f_{\rm sup} = (a_{\rm enter}/a_{\rm eq})^2$ compared to the case without suppression (figure taken from Bartelmann & Schneider, 2001)

.

and all perturbations below this size will be erased. Furthermore, scale-dependent corrections to the structure growth can also be introduced by baryons through acoustic oscillations (see Eisenstein & Hu, 1998).

To account for all these scale-dependent effects, one defines the *transfer function* T_k. It connects the ratio of perturbation amplitudes at present time ('0') to that at an initial epoch ('i')

$$\frac{\tilde{\delta}_0(k)}{\tilde{\delta}_0(k_{\rm s})} = T_k \frac{\tilde{\delta}_{\rm i}(k)}{\tilde{\delta}_{\rm i}(k_{\rm s})}\,, \tag{2.49}$$

where the wave vector k indicates the scale of interest, and it is compared to a scale $k_{\rm s}$ which is chosen to be large enough so that it entered the horizon only at late times.

The transfer function has an asymptotic behavior of $T_k \simeq 1$ for small k and $T_k \simeq k^{-2}$ for large k, with a turning point at $1/k \simeq r_{\rm h}(a_{\rm eq})$. That $T_k \simeq 1$ at small k follows directly from the definition. The behavior of $T_k \simeq k^{-2}$ at large k is because the large k modes correspond to small-scale fluctuations which entered the horizon at the radiation domination era. At that time, the horizon size grows proportionally to the scale factor $r_{\rm h}(a) \propto a$, which can be derived by inserting (2.9) to (2.20) and keeping only the radiation contribution in (2.9). Since a fluctuation enters the horizon when $k \cdot r_{\rm h} \simeq 1$, one has $a_{\rm enter} \propto k^{-1}$. Therefore the small-scale fluctuations are suppressed by $(a_{\rm enter}/a_{\rm eq})^2 \propto k^{-2}$.

A fitting formula of T_k has been derived by Bardeen et al. (1986) for a Cold Dark Matter-only universe. Another fitting formula which includes corrections due to baryonic effects is given in Eisenstein & Hu (1999).

Eularian non-linear perturbation theory

As the density contrast grows under gravity, at some point the non-linear terms in (2.40) and (2.41) begin to play an non-negligible role. In the Fourier domain this means different Fourier modes no longer grow independently but begin to interact with each other more and more strongly.

We now turn to work in the Fourier domain where the effects of the non-linear terms are easilier described. Writing the dynamical equations (2.40)-(2.42) in Fourier space, again assuming the velocity field to be curl-free, and substituting the Poisson equation (2.42) into (2.41) we get

$$\frac{\partial\tilde{\delta}(\boldsymbol{k})}{\partial t}+\frac{\tilde{\theta}(\boldsymbol{k})}{a}+\frac{1}{a}\int \mathrm{d}^3q\;\mathrm{d}^3p\;\delta_{\mathrm{D}}(\boldsymbol{k}-\boldsymbol{q}-\boldsymbol{p})\;\alpha(\boldsymbol{q},\boldsymbol{p})\;\tilde{\theta}(\boldsymbol{q},t)\;\tilde{\delta}(\boldsymbol{p},t)=0\,, \tag{2.50}$$

$$\frac{\partial\tilde{\theta}(\boldsymbol{k})}{\partial t}+\frac{\dot{a}}{a}\tilde{\theta}(\boldsymbol{k})+\frac{1}{a}\int \mathrm{d}^3q\;\mathrm{d}^3p\;\delta_{\mathrm{D}}(\boldsymbol{k}-\boldsymbol{q}-\boldsymbol{p})\;\beta(\boldsymbol{q},\boldsymbol{p})\;\tilde{\theta}(\boldsymbol{q},t)\;\tilde{\theta}(\boldsymbol{p},t)=0\,, \tag{2.51}$$

where $\tilde{g}$ indicates the Fourier counterpart of the quantity g, and the kernels α and β are defined as

$$\alpha(\boldsymbol{q},\boldsymbol{p})=\frac{(\boldsymbol{p}+\boldsymbol{q})\cdot\boldsymbol{q}}{q^2}\;,\quad \beta(\boldsymbol{q},\boldsymbol{p})=\frac{(\boldsymbol{p}+\boldsymbol{q})^2\boldsymbol{p}\cdot\boldsymbol{q}}{2p^2q^2}\;. \tag{2.52}$$

They describe the coupling between different Fourier modes which arises from the non-linear terms in (2.40) and (2.41).

To solve (2.50) and (2.51) perturbatively, we consider the Fourier transforms of $\delta^{(n)}$ and $\theta^{(n)}$. General solutions for $\tilde{\delta}^{(n)}$ and $\tilde{\theta}^{(n)}$ are hard to find due to the coupling of time and spatial dependence in the equations. However, in the special case of an EdS Universe, the time and spatial dependence can be separated thanks to the exact scaling $D_+(t)=a(t)$.

For an EdS Universe, making the ansatz

$$\tilde{\delta}^{(n)}(\boldsymbol{k},t)=D_+^n(t)\;\tilde{\delta}_n(\boldsymbol{k})\,, \tag{2.53}$$

$$\tilde{\theta}^{(n)}(\boldsymbol{k},t)=-\dot{a}\;D_+^n(t)\;\tilde{\theta}_n(\boldsymbol{k})\,, \tag{2.54}$$

one can see that the time dependence of (2.50) and (2.51) drops out. The general solutions for the spatial dependence can be written as

$$\tilde{\delta}^{(n)}(\boldsymbol{k})=\int \mathrm{d}^3\boldsymbol{q}_1...\mathrm{d}^3\boldsymbol{q}_n\;\delta_{\mathrm{D}}\left(\boldsymbol{k}-\sum_{i=1}^{n}\boldsymbol{q}_i\right)F_n(\boldsymbol{q}_1,...,\boldsymbol{q}_n)\;\delta^{(1)}(\boldsymbol{q}_1)...\delta^{(1)}(\boldsymbol{q}_n)\,, \tag{2.55}$$

$$\tilde{\theta}^{(n)}(\boldsymbol{k})=\int \mathrm{d}^3\boldsymbol{q}_1...\mathrm{d}^3\boldsymbol{q}_n\;\delta_{\mathrm{D}}\left(\boldsymbol{k}-\sum_{i=1}^{n}\boldsymbol{q}_i\right)G_n(\boldsymbol{q}_1,...,\boldsymbol{q}_n)\;\delta^{(1)}(\boldsymbol{q}_1)...\delta^{(1)}(\boldsymbol{q}_n)\,. \tag{2.56}$$

Here the functions F_n and G_n are constructed from the kernels α and β through recursion relations (see e.g. Goroff et al. 1986; Jain & Bertschinger 1994):

$$\begin{aligned}F_n(\boldsymbol{q}_1,\ldots,\boldsymbol{q}_n) &= \sum_{m=1}^{n-1}\frac{G_m(\boldsymbol{q}_1,\ldots,\boldsymbol{q}_m)}{(2n+3)(n-1)}\Big[(2n+1)\alpha(\boldsymbol{k}_1,\boldsymbol{k}_2)\;F_{n-m}(\boldsymbol{q}_{m+1},\ldots,\boldsymbol{q}_n)\\ &\quad +2\beta(\boldsymbol{k}_1,\boldsymbol{k}_2)\;G_{n-m}(\boldsymbol{q}_{m+1},\ldots,\boldsymbol{q}_n)\Big],\end{aligned} \tag{2.57}$$

$$\begin{aligned}G_n(\boldsymbol{q}_1,\ldots,\boldsymbol{q}_n) &= \sum_{m=1}^{n-1}\frac{G_m(\boldsymbol{q}_1,\ldots,\boldsymbol{q}_m)}{(2n+3)(n-1)}\Big[3\alpha(\boldsymbol{k}_1,\boldsymbol{k}_2)\;F_{n-m}(\boldsymbol{q}_{m+1},\ldots,\boldsymbol{q}_n)\\ &\quad +2n\beta(\boldsymbol{k}_1,\boldsymbol{k}_2)\;G_{n-m}(\boldsymbol{q}_{m+1},\ldots,\boldsymbol{q}_n)\Big],\end{aligned} \tag{2.58}$$

where $\boldsymbol{k}_1\equiv\boldsymbol{q}_1+\ldots+\boldsymbol{q}_m$, $\boldsymbol{k}_2\equiv\boldsymbol{q}_{m+1}+\ldots+\boldsymbol{q}_n$, $\boldsymbol{k}\equiv\boldsymbol{k}_1+\boldsymbol{k}_2$.

We have $F_1 = G_1 = 1$ by construction. For $n = 2$, one has (see Bernardeau et al., 2002a and references therein)

$$F_2(\boldsymbol{q}_1, \boldsymbol{q}_2) = \frac{5}{7} + \frac{1}{2}\frac{\boldsymbol{q}_1 \cdot \boldsymbol{q}_2}{q_1 q_2}\left(\frac{q_1}{q_2} + \frac{q_2}{q_1}\right) + \frac{2}{7}\frac{(\boldsymbol{q}_1 \cdot \boldsymbol{q}_2)^2}{q_1^2 q_2^2}, \tag{2.59}$$

$$G_2(\boldsymbol{q}_1, \boldsymbol{q}_2) = \frac{3}{7} + \frac{1}{2}\frac{\boldsymbol{q}_1 \cdot \boldsymbol{q}_2}{q_1 q_2}\left(\frac{q_1}{q_2} + \frac{q_2}{q_1}\right) + \frac{4}{7}\frac{(\boldsymbol{q}_1 \cdot \boldsymbol{q}_2)^2}{q_1^2 q_2^2}. \tag{2.60}$$

2.2.5 Two- and three-point statistics of the matter density field

So far we have considered the evolution of matter density fluctuations. The initial condition for this evolution, currently given by the inflation theory, can be formulated only in a statistical way. For this reason one describes the matter density contrast δ in the Universe by a *random field*, and studies only its statistical properties.

The density field in our current Universe corresponds to one realization of this random field. There is no way to apply an ensemble average to the observational data. However with the help of the ergodic hypothesis and the cosmological principle, one can perform a spatial average instead. The comparison of the results with theoretical predictions holds only in the limit of validity of the cosmological principle.

A general way to study a random field is to study its moments. In the case of the matter density contrast δ, its first moment $\langle\delta\rangle$ vanishes. A full description of its statistical properties would require all higher-order moments.

According to observations of the CMB as well as predictions from the simplest single-scalar-field inflation theory, the primordial (i.e. at a very high redshift) matter density field is very close to a Gaussian random field. For a Gaussian random field, all the statistical information is contained in its second-order moment. All odd-power higher order moments vanish, and all even-power higher order moments can be expressed in terms of the second-order moment.

As long as the density perturbations grow linearly, they remain Gaussian. But as gravitational clustering is non-linear in nature, non-Gaussianity will be generated inevitably, resulting in non-trivial higher-order moments. The matter density field of the Universe today has already significant deviations from a Gaussian field (Kayo et al. 2001; Ostriker et al. 2003). A common way of measuring non-Gaussianity nowadays is to use the third-order statistics. In this subsection, the second- and third-order statistics (the 2- and 3-pt correlation function and their Fourier counterparts) of the matter density field are introduced.

Two-point correlation function and the power spectrum

The 2-pt correlation function (2PCF) of the matter density field at position $\boldsymbol{x}$ and $\boldsymbol{y}$ is defined as

$$\langle\delta(\boldsymbol{x})\delta(\boldsymbol{y})\rangle =: C_{\delta\delta}(|\boldsymbol{x} - \boldsymbol{y}|)\,, \tag{2.61}$$

where $\langle\ \rangle$ indicates the ensemble average. The 2PCF depends only on the separation $|\boldsymbol{x} - \boldsymbol{y}|$ of the two points due to the assumption of δ being statistically homogeneous and isotropic.

The configuration space δ is a real quantity, but its Fourier counterpart, defined as

$$\tilde{\delta}(\boldsymbol{k}) := \int \mathrm{d}^3x\ \delta(\boldsymbol{x})\ \mathrm{e}^{-\mathrm{i}\boldsymbol{x}\cdot\boldsymbol{k}}\,, \tag{2.62}$$

is in general complex. Consider the second-order moment of $\tilde{\delta}(\boldsymbol{k})$:

$$\begin{aligned}\left\langle \tilde{\delta}(\boldsymbol{k})\tilde{\delta}^*(\boldsymbol{k}')\right\rangle &= \int \mathrm{d}^3x\ \ \mathrm{e}^{-\mathrm{i}\boldsymbol{x}\cdot\boldsymbol{k}} \int \mathrm{d}^3x'\ \ \mathrm{e}^{\mathrm{i}\boldsymbol{x}'\cdot\boldsymbol{k}'}\ \ \langle\delta(\boldsymbol{x})\delta(\boldsymbol{x}')\rangle \\ &= \int \mathrm{d}^3x\ \ \mathrm{e}^{-\mathrm{i}\boldsymbol{x}\cdot\boldsymbol{k}} \int \mathrm{d}^3y\ \ \mathrm{e}^{\mathrm{i}(\boldsymbol{x}+\boldsymbol{y})\cdot\boldsymbol{k}'}\ C_{\delta\delta}(|\boldsymbol{y}|) \\ &= (2\pi)^3\delta_{\mathrm{D}}(\boldsymbol{k}-\boldsymbol{k}') \int \mathrm{d}^3y\ \ \mathrm{e}^{-\mathrm{i}\boldsymbol{y}\cdot\boldsymbol{k}}\ C_{\delta\delta}(|\boldsymbol{y}|)\,,\end{aligned} \tag{2.63}$$

the form of this motivates one to define the *power spectrum* P_δ of the matter density contrast, as

$$\left\langle \tilde{\delta}(\boldsymbol{k})\tilde{\delta}^*(\boldsymbol{k}')\right\rangle := (2\pi)^3\delta_{\mathrm{D}}(\boldsymbol{k}-\boldsymbol{k}')P_\delta(|\boldsymbol{k}|)\,. \tag{2.64}$$

Comparing (2.63) and (2.64) one can easily see that the power spectrum and the 2PCF are Fourier transform pairs,

$$P_\delta(|\boldsymbol{k}|) = \int \mathrm{d}^3y\ \ \mathrm{e}^{-\mathrm{i}\boldsymbol{y}\cdot\boldsymbol{k}}\ C_{\delta\delta}(|\boldsymbol{y}|)\,. \tag{2.65}$$

Generally, since the density contrast δ evolves with time, the power spectrum and 2PCF are also functions of time. Using the knowledge of the linear growth of density contrast, one can readily express a late time linear power spectrum $P_{\delta,\mathrm{L}}(k,a)$ as a function of the power spectrum at a specified initial epoch with scale factor a_i,

$$P_{\delta,\mathrm{L}}(k,a) = \frac{D_+^2(a)}{D_+^2(a_\mathrm{i})}T_k^2P_\delta(k,a_\mathrm{i})\,. \tag{2.66}$$

A power law initial power spectrum is usually assumed,

$$P_\delta(k,a_\mathrm{i}) \propto k^{n_\mathrm{s}}\,, \tag{2.67}$$

where the *spectrum index* n_s is observationally found to be close to unity (e.g. Seljak et al. 2005; Sánchez et al. 2009).

The scale dependence of the linear power spectrum is totally contained in the transfer function T_k. Taking $n_\mathrm{s} = 1$, with the asymptotic behavior of T_k, one has

$$P_{\delta,\mathrm{L}}(k) \propto \begin{cases} k & \text{for } 1/k \gg r_\mathrm{h}(a_\mathrm{eq}) \\ k^{-3} & \text{for } 1/k \ll r_\mathrm{h}(a_\mathrm{eq})\,. \end{cases}$$

The turnover scale at $1/k \simeq r_\mathrm{h}(a_\mathrm{eq})$ is the only characteristic scale in the linear power spectrum. It corresponds to the scale of the fluctuation which enters the horizon at matter-radiation equality. The growth of perturbations with smaller sizes and which enter the horizon earlier is suppressed in the radiation-dominated era.

Non-linear growth of the density contrast adds an additional scale dependence to the power spectrum: it affects the small scales more. Theoretical treatments of non-linear power spectrum can be found in e.g. Peacock & Dodds (1996) and Smith et al. (2003).

The normalization of the power spectrum is fixed by the parameter σ_8, which is defined as the variance of density fluctuations in spheres of radius $8h^{-1}$Mpc.

The bispectrum

The bispectrum of the matter density field is defined as

$$\left\langle \tilde{\delta}(\boldsymbol{k_1})\tilde{\delta}(\boldsymbol{k_2})\tilde{\delta}(\boldsymbol{k_3})\right\rangle =: (2\pi)^3\ \delta_{\mathrm{D}}(\boldsymbol{k_1}+\boldsymbol{k_2}+\boldsymbol{k_3})B_\delta(k_1,k_2,k_3)\,. \tag{2.68}$$

The Dirac delta function here guarantees that the bispectrum is defined only when $\boldsymbol{k_1}$, $\boldsymbol{k_2}$ and $\boldsymbol{k_3}$ form a triangle. This is again related to the statistical homogeneity the Universe. After taking into account the statistical homogeneity and the isotropy, the bispectrum is characterized by only three real quantities, which are chosen to be the three side lengths of the triangle here. It is also common to choose two side lengths and the angle between them to parametrize the bispectrum.

If one assumes Gaussian initial conditions, then the bispectrum generated by gravitational instability at large scales can be given by second-order perturbation theory, and reads (Fry 1984)

$$B_\delta(k_1,k_2,k_3) = 2F_2(\boldsymbol{k_1},\boldsymbol{k_2})P_\delta(k_1)P_\delta(k_2) + \text{cyc.}, \tag{2.69}$$

with F_2 defined by (2.59).

Efforts have been made to modify (2.69) in order to fit the results from N-body simulations better (Scoccimarro & Frieman 1999), i.e. to interpolate between perturbative and the highly non-linear regimes. This is achieved by replacing the kernel (2.59) by an effective kernel

$$F_2^{\mathrm{eff}}(\boldsymbol{k_1},\boldsymbol{k_2}) = \frac{5}{7}a(n,k_1)a(n,k_2) + \frac{1}{2}\frac{\boldsymbol{k}_1\cdot\boldsymbol{k}_2}{k_1k_2}\left(\frac{k_1}{k_2}+\frac{k_2}{k_1}\right)b(n,k_1)b(n,k_2) + \frac{2}{7}\frac{(\boldsymbol{k}_1\cdot\boldsymbol{k}_2)^2}{k_1^2k_2^2}c(n,k_1)c(n,k_2)\,, \tag{2.70}$$

where n is the spectral index, usually chosen to be the local spectral index obtained from the linear power spectrum (see Scoccimarro & Couchman 2001). The functions $a(n,k_1)$, $b(n,k_1)$ and $c(n,k_1)$ are chosen to fit the N-body simulations for small scales and are approximating unity for large scales.

Currently, the commonly used fitting formula for the non-linear evolution of B_δ in CDM models is given by Scoccimarro & Couchman (2001). It fits the measurements in N-body simulations to an accuracy of 15%.

Chapter 3

Gravitational lensing

The physical law that governs the traveling of light is Fermat's principle. It states that out of all the possible paths connecting two points in space, the light follows those with stationary light-travel time. This means, in an Euclidian space, that light travels along straight lines; and in a spacetime described by GR, that light follows null geodesics.

Since spacetime is curved around a massive body in the picture of GR, light rays can be bent. Although the idea of light bent by massive bodies has a very early origin (Newton 1704), it was during the Solar eclipse in 1919 that the first observations (Dyson et al. 1920) were made. The measured light-deflection angles in these observations provided a strong support for GR.

The deflection of light by a gravitational field is analogous to the light deflection by an optical lens, thus the name of gravitational lensing. The gravitational lensing effect can act as a direct probe of the gravitational field, unlike most of the other astronomical probes which rely on luminous matter. With this advantage, it has been used to measure the matter distribution on all scales, from planets up to the LSS of the Universe.

According to the degree of distortion to the original light path, gravitational lensing is divided into *strong lensing* and *weak lensing*. Strong lensing occurs around high mass concentrations (e.g. compact objects, galaxies or clusters of galaxies), and is associated with phenomena like multiple images and giant arcs. Weak lensing, on the other hand, deals with mild distortions of the light bundle. Its observable effects can be caused by numerous objects along the light path, and studies are mostly done statistically. In this thesis we will concentrate on cosmic shear – weak gravitational lensing by the LSS.

Cosmic shear is sensitive to all cosmological parameters which have influence on the density perturbations and/or the geometry of the Universe, including those concerning properties of dark energy, which have been a key concern after the discoveries made by observations of supernovae, the cosmic microwave background, and the large-scale structure (for a review see e.g. Munshi et al. 2008). Since its first detection in 2000 (Bacon et al. 2000; Kaiser et al. 2000; Van Waerbeke et al. 2000; Wittman et al. 2000), cosmic shear has been developed into a competitive cosmological probe. Its constraining power on cosmological parameters is now comparable to other probes (e.g. Spergel et al. 2007; Fu et al. 2008). With forthcoming large-field multicolor imaging surveys, photometric redshift and shape information of a huge number of galaxies will be available, rendering cosmic shear even greater statistical power. In particular, cosmic shear is considered to be one of the most promising dark energy probes (Albrecht et al. 2006; Peacock et al. 2006) when the results of these surveys become available.

We introduce the basic theory of cosmic shear in this chapter, where Bartelmann & Schneider (2001) and Schneider et al. (1992) have been heavily referenced.

3.1 The geometry of gravitational lensing

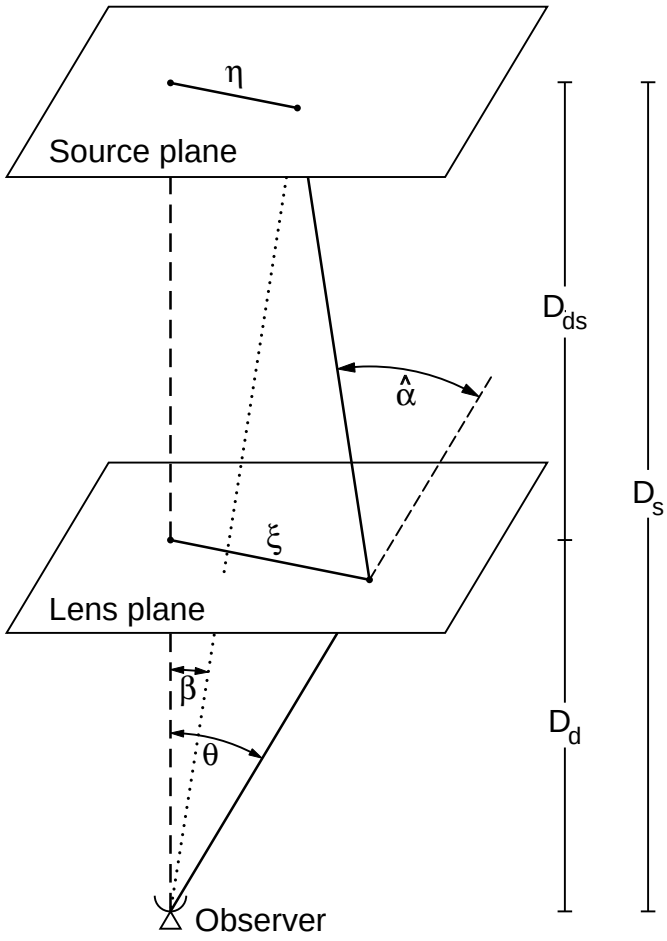


Figure 3.1: Sketch of the geometry of a gravitational lens system (from Bartelmann & Schneider 2001).

Consider a typical situation in gravitational lensing: the deflection of a light ray by a point mass as depicted in Fig. 3.1. The lens is located at the intersection of the dashed line and the 'Lens plane'; the light ray follows the solid line, is deflected at the lens plane by an angle $\hat{\boldsymbol{\alpha}}$, and is finally observed by the observer at an angular position $\boldsymbol{\theta}$. The distance between the lens plane and the 'Source plane' is $D_{\rm ds}$, and that between the lens plane and the observer is $D_{\rm d}$. Since these distances here relate physical transverse separations (e.g. $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$) to angles, they are the angular-diameter distances.

3.1.1 The lens equation and the deflection angle

A basic equation to describe this light deflection would be one that relates the true position of a source $\boldsymbol{\beta}$ to its observed position $\boldsymbol{\theta}$ on the sky. This equation, named the lens equation, can be derived from the geometrical relations shown in Fig .3.1 using the fact that the angles in consideration are small:

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \frac{D_{\rm ds}}{D_{\rm s}}\,\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) \equiv \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta})\ . \tag{3.1}$$

The mechanism of light deflection enters (3.1) in the form of the deflection angle $\hat{\boldsymbol{\alpha}}$. Using GR in the weak-field limit, the deflection angle for a ray with impact parameter $\boldsymbol{\xi}$ is linked to the point mass as

$$\hat{\alpha} = \frac{4GM}{c^2\,\xi} \,. \tag{3.2}$$

Note that even in the case of strong lensing, one still works in the weak-field limit. Light deflection in strong field, e.g. near black holes, is not a subject of the gravitational lensing theory that will be introduced in this chapter. The form of the deflection angle (3.2) and the form of the angular diameter distance (2.23) are the only places GR enters in the whole gravitational lensing theory.

When the deflecting mass is not point-like but spatially extended, the deflection angle can be calculated as the sum of deflections by its individual mass elements.

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \frac{4G}{c^2} \int \mathrm{d}^2\xi'\, \Sigma(\boldsymbol{\xi}') \, \frac{\boldsymbol{\xi} - \boldsymbol{\xi}'}{|\boldsymbol{\xi} - \boldsymbol{\xi}'|^2} \,, \tag{3.3}$$

where Σ is the matter density projected along the line-of-sight

$$\Sigma(\boldsymbol{\xi}) \equiv \int \mathrm{d}r_3\, \rho(\xi_1, \xi_2, r_3) \,. \tag{3.4}$$

Here the so-called Born approximation has been made. Analogous to the Born approximation in atomic and nuclear physics, we have assumed that the distribution of the deflecting mass is small ('thin') along the line-of-sight compared to the distance traveled by the light ray, so that the light ray can be approximated as a straight line in the neighborhood of the deflecting mass distribution. This approximation holds well for gravitational lensing by common extended astronomical objects, e.g. galaxies and clusters of galaxies.

Based on (3.4), we further find an expression for $\boldsymbol{\alpha}(\boldsymbol{\theta})$ in (3.1):

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{1}{\pi} \int \mathrm{d}^2\theta'\, \kappa(\boldsymbol{\theta}') \, \frac{\boldsymbol{\theta} - \boldsymbol{\theta}'}{|\boldsymbol{\theta} - \boldsymbol{\theta}'|^2} \,, \tag{3.5}$$

where the *dimensionless surface mass density* or *convergence* κ is defined as

$$\kappa(\boldsymbol{\theta}) = \frac{\Sigma(D_\mathrm{d}\boldsymbol{\theta})}{\Sigma_\mathrm{cr}} \quad \text{with} \quad \Sigma_\mathrm{cr} := \frac{c^2}{4\pi G} \, \frac{D_\mathrm{s}}{D_\mathrm{d}\, D_\mathrm{ds}} \,. \tag{3.6}$$

As will be shown, the *critical surface mass density* Σ_cr is also the characteristic value for the surface mass density which divides the 'weak' and 'strong' lensing regimes. By definition (3.6), it is distance dependent.

3.1.2 Jacobi matrix of lens mapping

For two-dimensional quantities like the angles defined ablove, we have the identities $\nabla \ln|\boldsymbol{\theta}| = \boldsymbol{\theta}/|\boldsymbol{\theta}|^2$ and $\nabla^2 \ln|\boldsymbol{\theta}| = \nabla \cdot \left(\boldsymbol{\theta}/|\boldsymbol{\theta}|^2\right) = 2\pi\delta_\mathrm{D}(\boldsymbol{\theta})$. This motivates one to define the deflection potential ψ as (e.g. Bartelmann & Schneider 2001)

$$\psi(\boldsymbol{\theta}) = \frac{1}{\pi} \int \mathrm{d}^2\theta'\, \kappa(\boldsymbol{\theta}') \, \ln|\boldsymbol{\theta} - \boldsymbol{\theta}'| \,, \tag{3.7}$$

so that we can express $\boldsymbol{\alpha}$ and κ as

$$\boldsymbol{\alpha} = \nabla\psi \,, \tag{3.8}$$

$$\kappa = \frac{1}{2}\nabla^2\psi = \frac{1}{2}\left(\psi_{,11} + \psi_{,22}\right) , \tag{3.9}$$

where $g_{,ij}$ indicates the derivative of g with respect to the basis vectors $\boldsymbol{e}_i$ and $\boldsymbol{e}_j$.

As depicted by Fig .3.1, the image of a light source at position $\boldsymbol{\beta}$ would be observed at a different position $\boldsymbol{\theta}$ due to the lensing effect. To describe the mapping from the original position to the observed position, one considers the Jacobi matrix of the mapping

$$\mathcal{A}(\boldsymbol{\theta}) := \frac{\partial\boldsymbol{\beta}}{\partial\boldsymbol{\theta}} = \left(\delta_{ij} - \frac{\partial^2\psi(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\right) = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix} , \tag{3.10}$$

where we have introduced the components of the *shear* $\gamma \equiv \gamma_1 + \mathrm{i}\gamma_2 = |\gamma|\mathrm{e}^{2\mathrm{i}\varphi}$, with

$$\gamma_1 = \frac{1}{2}(\psi_{,11} - \psi_{,22}) , \quad \gamma_2 = \psi_{,12} . \tag{3.11}$$

We can see that the matrix $\mathcal{A}$ is symmetric, with trace $\mathrm{tr}\mathcal{A} = 2(1-\kappa)$, and the shear γ describing its trace-less part. The ratio μ of image size and the size of the source is given by the determinant of $\mathcal{A}$,

$$\mu = \frac{1}{\det\mathcal{A}} = \frac{1}{(1-\kappa)^2 - |\gamma|^2} . \tag{3.12}$$

For the special case of a source with circular isophotes, its image would have elliptical isophotes. The ratio of the major and the minor axis of the ellipse to the radius of the circle would be $1-\kappa+|\gamma|$ and $1-\kappa-|\gamma|$, respectively, and the major axis of the ellipse would point into the direction of ϕ, with 2ϕ being the phase angle of γ.

The last feature described above is related to the fact that the shear is a spin-2 quantity. Here the spin of a quantity is based only on the rate its polar angle changes with respect to the rotation of the coordinate system. For example, if the coordinate systems turns an angle θ, a spin-2 quantity would rotate 2θ in the opposite direction, i.e. it gets multiplied by a phase factor $\mathrm{e}^{-2\mathrm{i}\theta}$. We write the two components of shear in terms of a complex number because the spin property is conveniently expressed in this way.

In the lens mapping, surface brightness is conserved according to Liouville's theorem. Thus μ, the ratio of image size and the size of the source, is also the flux ratio of the image and the source. In the weak lensing regime we have $\mu \approx 1$, the case $\mu \gg 1$ occurs by definition only in the strong lensing regime. In (3.12) one can see that $\mu \gg 1$ happens when the convergence κ is close to 1, i.e. when the surface mass density is close to Σ_{cr}.

The lensing magnification is one of the cosmological tools provided by gravitational lensing. We will focus on another tool, the image shape distortion quantified by the shear γ. In the case of a circular background source, the axis ratio of the lensed image is $b/a = 1-\kappa-|\gamma|/(1-\kappa+|\gamma|)$, which leads to

$$\frac{|\gamma(\boldsymbol{\theta})|}{1-\kappa(\boldsymbol{\theta})} = \frac{1-b/a}{1+b/a} , \tag{3.13}$$

i.e. it is the quantity $\gamma(\boldsymbol{\theta})/\left[1-\kappa(\boldsymbol{\theta})\right]$, called the reduced shear, that is determined by the axis ratio of the lensed image. Generally, a background source is not circular, but is elliptical to the first order. We denote this intrinsic ellipticity as ϵ_{I}. If the intrinsic ellipticities of a population of galaxies are statistically uncorrelated, then the reduced shear can be estimated statistically from the observed ellipticities of these galaxies. In this thesis we take the approximation $\gamma \simeq g$, which holds well in the limit of weak lensing.

Since both the convergence κ and the shear γ can be written as combinations of second derivatives of the deflection potential ϕ, see (3.9) and (3.11), they are interrelated. With (3.7) it is straight forward to find

$$\gamma(\boldsymbol{\theta}) = \frac{1}{\pi}\int \mathrm{d}^2\theta'\, \mathcal{D}(\boldsymbol{\theta}-\boldsymbol{\theta}')\,\kappa(\boldsymbol{\theta}')\,, \tag{3.14}$$

where

$$\mathcal{D}(\boldsymbol{\theta}) \equiv \frac{\theta_2^2-\theta_1^2-2\mathrm{i}\theta_1\theta_2}{|\boldsymbol{\theta}|^4} = \frac{-1}{(\theta_1-\mathrm{i}\theta_2)^2}\,. \tag{3.15}$$

We will call this convolution kernel $\mathcal{D}$ the Kaiser-Squires (K-S) kernel and (3.14) the K-S relation, since it was first recognized in Kaiser & Squires (1993) that the inverse relation of (3.14) suggests one can reconstruct the projected mass distribution from the shear signal. The Fourier counterpart of the K-S kernel can be obtained by Fourier transforming (3.9) and (3.11), as

$$\tilde{\mathcal{D}}(\boldsymbol{\ell}) = \pi\frac{\ell_1^2-\ell_2^2+2\mathrm{i}\ell_1\ell_2}{|\boldsymbol{\ell}|^2} = \pi\,\mathrm{e}^{2\mathrm{i}\beta} \quad \text{for} \quad \ell\neq 0\,, \tag{3.16}$$

where β is the polar angle of $\boldsymbol{\ell}$. With this, one can express the Fourier counterpart of (3.14) as

$$\tilde{\gamma}(\boldsymbol{\ell}) = \mathrm{e}^{2\mathrm{i}\beta}\tilde{\kappa}(\boldsymbol{\ell}) \quad \text{for} \quad \ell\neq 0\,, \tag{3.17}$$

i.e. the Fourier counterpart of the convergence and the shear are the same up to a phase factor.

The Fourier space relation (3.17) further allows one to invert the relation (3.14). Since one has

$$\tilde{\kappa}(\boldsymbol{\ell}) = \mathrm{e}^{-2\mathrm{i}\beta}\tilde{\gamma}(\boldsymbol{\ell}) \quad \text{for} \quad \ell\neq 0\,, \tag{3.18}$$

It is straightforward to see that

$$\kappa(\boldsymbol{\theta})-\kappa_0 = \frac{1}{\pi}\int \mathrm{d}^2\theta'\, \mathcal{D}^*(\boldsymbol{\theta}-\boldsymbol{\theta}')\,\gamma(\boldsymbol{\theta}')\,. \tag{3.19}$$

An arbitrary constant κ_0 occurs since the $\ell = 0$ mode is not determined. This means if the shear field is known, the convergence field can be determined, but only up to an additive constant. This arbitrary constant causes a major problem in using the shear signal to reconstruct the projected mass distribution, e.g. in the case of inferring the mass profile of a galaxy cluster with the shear signal in the field. Especially in the inner regions of a galaxy cluster, the approximation $\gamma \simeq g$ does not hold well any more. In this case the problem of the arbitrary constant translates to the fact that, if one scales the convergence field as $[1-\kappa'(\boldsymbol{\theta})] = \lambda[1-\kappa(\boldsymbol{\theta})]$, the reduced shear g stays invariant. This so-called *mass-sheet degeneracy* can be removed with the aid of additional observables or physical assumptions. The cosmic shear study, which is based on the statistical properties of the shear and convergence field, is not affected by the mass sheet degeneracy.

3.2 Cosmic Shear

Cosmic shear refers to the coherent shape distortion of distant sources by the LSS. This distortion is usually very mild, typically of the order of a few percent. Thus cosmic shear has to be detected and studied in a statistical way, using images of a large number of distant galaxies.

For cosmic shear, the thin lens approximation fails, since the 'lens' here is no longer a concentrated object, but all the intervening matter between the source and the observer. Amazingly, the formalism of gravitational lensing as presented in the previous section still holds for cosmic shear

to the linear order. We will show this by considering the light propagation in a 3D distribution of matter, as is the case for the LSS of the Universe. Then we will introduce the concept of E- and B-mode for the shear signal, and show the commonly used statistical measures of cosmic shear and some relations between them.

3.2.1 Light propagation in a three-dimensional matter distribution

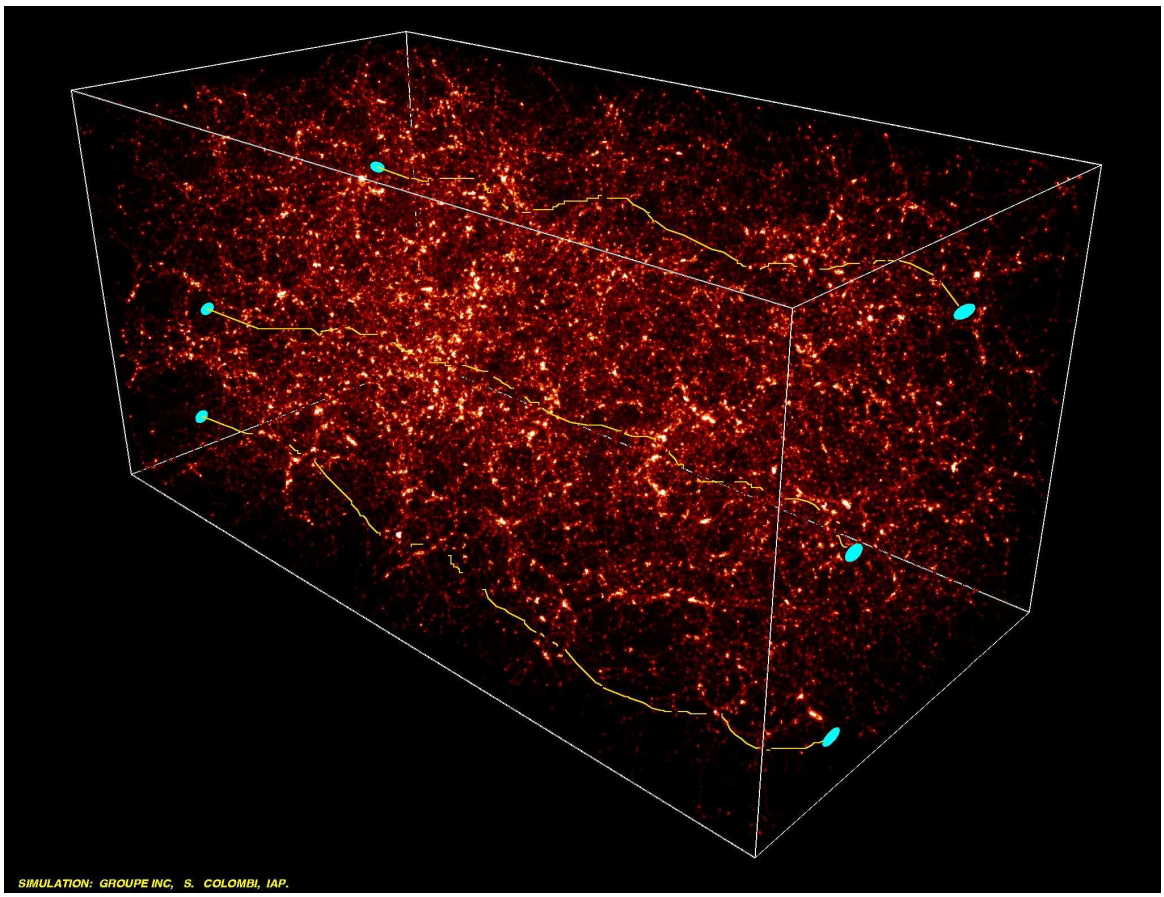


Figure 3.2: The paths of three light rays travelling through a simulated mass distribution. The deflections have been greatly exaggerated. Credit: S. Colombi (IAP), CFHT Team.

In an inhomogeneous universe, light of background galaxies is distorted continuously on its path by the intervening matter. The equation governing the propagation of thin light bundles through an arbitrary spacetime is the *equation of geodesic deviation*. It describes how the comoving separation vector $\boldsymbol{x}(\boldsymbol{\theta},\chi)$ between a ray separated by an angle $\boldsymbol{\theta}$ at the observer from a fiducial ray evolves:

$$\frac{\mathrm{d}^2\boldsymbol{x}}{\mathrm{d}\chi^2} + K\boldsymbol{x} = -\frac{2}{c^2}\left[\nabla_\perp\Phi\left(\boldsymbol{x}(\boldsymbol{\theta},\chi),\chi\right) - \nabla_\perp\Phi^{(0)}(\chi)\right], \tag{3.20}$$

where χ is the comoving radial distance, K is the curvature signature (see Sect. 2.1.1), and Φ is the comoving gravitational potential as defined in (2.39); $\nabla_\perp$ is the 2D comoving gradient operator acting on the plane perpendicular to the fiducial light ray; '0' denotes the fiducial light ray.

Equation (3.20) can be solved with the Green's function method, yielding

$$\boldsymbol{x}(\boldsymbol{\theta},\chi) = f_K(\chi)\boldsymbol{\theta} - \frac{2}{c^2}\int_0^\chi \mathrm{d}\chi'\ f_K(\chi-\chi')\ \left[\nabla_\perp\Phi\left(\boldsymbol{x}(\boldsymbol{\theta},\chi'),\chi'\right) - \nabla_\perp\Phi^{(0)}(\chi')\right], \tag{3.21}$$

with $f_K(\chi)$ defined by (2.3).

The true angular position of a source at $\boldsymbol{x}(\boldsymbol{\theta},\chi)$ is $\boldsymbol{\beta} = \boldsymbol{x}/f_K(\chi)$. Thus (3.21) leads to an expression of the Jacobi matrix of lens mapping as defined in (3.10):

$$\mathcal{A}_{ij}(\boldsymbol{\theta},\chi) = \delta_{ij} - \frac{2}{c^2}\int_0^\chi \mathrm{d}\chi'\ \frac{f_K(\chi-\chi')f_K(\chi')}{f_K(\chi)}\ \Phi_{,ik}\left(\boldsymbol{x}(\boldsymbol{\theta},\chi'),\chi'\right)\ \mathcal{A}_{kj}(\boldsymbol{\theta},\chi')\,. \tag{3.22}$$

This equation is exact in the limit of a weakly inhomogeneous universe. From it we can see that the source of the image distortion and magnification is the second order derivative of the gravitational potential along the actual light path $\boldsymbol{x}(\boldsymbol{\theta},\chi)$. The fact that the Jacobi matrix also enters the r.h.s. of the equation suggests that the light deflection at one comoving distance χ is dependent on that at comoving distance $\chi' < \chi$ (lens-lens coupling).

Now we apply two approximations. Firstly we replace $\Phi_{,ik}\left(\boldsymbol{x}(\boldsymbol{\theta},\chi'),\chi'\right)$ by $\Phi_{,ik}\left(f_K(\chi')\boldsymbol{\theta},\chi'\right)$ in (3.22), i.e. we evaluate the potential derivative along the fiducial light ray. This approximation is again called the Born approximation; it holds well if the second derivative of the gravitational potential is smooth within the scale of the separation vector $\boldsymbol{x}$. Secondly we keep the Jacobi matrix on the r.h.s. of (3.22) only to the zeroth order of Φ, i.e. we approximate it by δ_{kj}. In this way we neglect the lens-lens coupling. Discussion of validity of these two approximations can be found in Schneider et al. (1998); Cooray & Hu (2002); Shapiro & Cooray (2006) and Hilbert et al. (2009). A general conclusion is that the corrections to them are not important for cosmic shear surveys which will be performed in the near future.

The Jacobi matrix after applying these two assumptions is

$$\mathcal{A}_{ij}(\boldsymbol{\theta},\chi) = \delta_{ij} - \frac{2}{c^2}\int_0^\chi \mathrm{d}\chi'\ \frac{f_K(\chi-\chi')f_K(\chi')}{f_K(\chi)}\ \Phi_{,ij}\left(f_K(\chi')\boldsymbol{\theta},\chi'\right)\,. \tag{3.23}$$

We can match the form $\mathcal{A}_{ij} = \delta_{ij} - \psi_{,ij}$ by redefining the potential

$$\psi(\boldsymbol{\theta},\chi) := \frac{2}{c^2}\int_0^\chi \mathrm{d}\chi'\ \frac{f_K(\chi-\chi')f_K(\chi')}{f_K(\chi)}\ \Phi\left(f_K(\chi')\boldsymbol{\theta},\chi'\right)\,. \tag{3.24}$$

Now we see that lensing by the 3D matter distribution, under the two approximations introduced, can be treated in the same way as in the case of a thin lens. The equivalent lens plane has a deflection potential of ψ as defined in (3.24), convergence $\kappa = \nabla^2\psi/2$, and shear $\gamma = (\psi_{,11} - \psi_{,22})/2 + \mathrm{i}\psi_{,12}$.

This convergence can be further related to the density fluctuation in the Universe δ by making use of (2.42), yielding

$$\kappa(\boldsymbol{\theta},\chi) = \frac{3H_0^2\Omega_\mathrm{m}}{2c^2}\int_0^\chi \mathrm{d}\chi'\ \frac{f_K(\chi-\chi')f_K(\chi')}{f_K(\chi)}\ \frac{\delta\left(f_K(\chi')\boldsymbol{\theta},\chi'\right)}{a(\chi')}\,. \tag{3.25}$$

Equation (3.25) gives the equivalent convergence for a source at comoving distance χ which is observed at position $\boldsymbol{\theta}$. Again, it receives contributions from all the intervening matter. In cosmic shear studies, one usually considers the lensing distortion of a large number of galaxies spread along the line of sight. Denoting the distance probability distribution of a population i of source galaxies by $p_\mathrm{s}^{(i)}(\chi)$, one obtains a source-averaged convergence

$$\kappa(\boldsymbol{\theta}) = \int_0^{r_\mathrm{h}} \mathrm{d}\chi\ p_\mathrm{s}^{(i)}(\chi)\kappa(\boldsymbol{\theta},\chi) = \frac{3H_0^2\Omega_\mathrm{m}}{2c^2}\int_0^{r_\mathrm{h}} \mathrm{d}\chi\ g(\chi)f_K(\chi)\ \frac{\delta\left(f_K(\chi)\boldsymbol{\theta},\chi\right)}{a(\chi)}\,, \tag{3.26}$$

where we have defined the lensing efficiency

$$g(\chi) = \int_{\chi}^{r_h} \mathrm{d}\chi' \; p_s^{(i)}(\chi') \; \frac{f_K(\chi' - \chi)}{f_K(\chi')} \; . \tag{3.27}$$

The lensing efficiency $g(\chi)$ can be regarded as D_{ds}/D_s weighted over the source population for a lens at comoving distance χ.

3.2.2 E- and B-modes

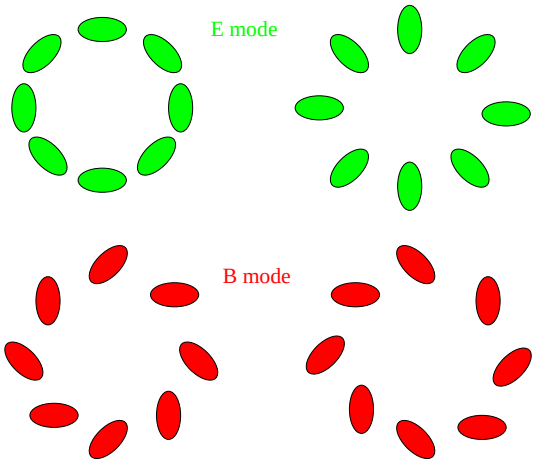


Figure 3.3: *Upper row*: A point mass (left) or underdensity (right) produces tangential, curl-free shear patterns called the E-mode. *Lower row*: Divergence-free pattern obtained from rotating all shears by 45 degrees. These so-called B-mode patterns cannot be produced from gravitational lensing (figure from Van Waerbeke & Mellier, 2003).

The shear can be seen as the lensing contribution to galaxy shapes, which can be described by a symmetric and trace-less 2D tensor field, i.e. a polarization field. Such a field can be decomposed into a curl-free and a divergence-free component, in analogy with the decomposition of the electromagnetic field into a electric E-mode and a magnetic B-mode. Therefore this decomposition is also called the E/B-mode decomposition.

Defined as the second-order derivatives of the deflection potential ψ (3.11), the shear generated by gravitational lensing is a pure E-mode field as the deflection potential is a scalar field. The measured B-mode comes only from systematics and higher-order effects. Thus performing an E/B-mode decomposition provides a check on the possible systematics in the measured shear signal (e.g. Crittenden et al. 2002; Pen et al. 2002). In Chap. 4 we will present some results concerning E/B-mode decomposition for 3-pt statistics.

3.2.3 Two- and three-point cosmic shear statistics

The two-point correlation functions and the power spectrum

Since the shear field is a polarization field which we describe by a complex number at each spatial position, the 2-pt correlator of shear corresponds to more than one real function. Consider the correlation of shear at two positions separated by $\boldsymbol{\theta} = \theta\, \mathrm{e}^{\mathrm{i}\varphi}$. It is convenient to define the tangential and cross-component of the shear regarding to this pair as $\gamma_t = -\mathrm{Re}(\gamma\, \mathrm{e}^{-2\mathrm{i}\varphi})$ and $\gamma_\times = -\mathrm{Im}(\gamma\, \mathrm{e}^{-2\mathrm{i}\varphi})$. Then one can proceed to define three real correlation functions of these shear components,

$$\xi_+(\theta) = \langle \gamma_t(\boldsymbol{\vartheta} + \boldsymbol{\theta}) \gamma_t(\boldsymbol{\vartheta}) \rangle + \langle \gamma_\times(\boldsymbol{\vartheta} + \boldsymbol{\theta}) \gamma_\times(\boldsymbol{\vartheta}) \rangle \; , \tag{3.28}$$

$$\xi_-(\theta) = \langle \gamma_t(\boldsymbol{\vartheta} + \boldsymbol{\theta}) \gamma_t(\boldsymbol{\vartheta}) \rangle - \langle \gamma_\times(\boldsymbol{\vartheta} + \boldsymbol{\theta}) \gamma_\times(\boldsymbol{\vartheta}) \rangle \; , \tag{3.29}$$

$$\xi_\times(\theta) = \langle \gamma_\mathrm{t}(\boldsymbol{\vartheta}+\boldsymbol{\theta})\gamma_\times(\boldsymbol{\vartheta})\rangle \ . \tag{3.30}$$

Thanks to the simple relation (3.17) between the shear and the convergence in Fourier space, one has

$$(2\pi)^2\delta_\mathrm{D}(\boldsymbol{\ell}-\boldsymbol{\ell}')P_\gamma(\ell) := \langle\tilde{\gamma}(\boldsymbol{\ell})\tilde{\gamma}^*(\boldsymbol{\ell}')\rangle = \langle\tilde{\kappa}(\boldsymbol{\ell})\tilde{\kappa}^*(\boldsymbol{\ell}')\rangle =: (2\pi)^2\delta_\mathrm{D}(\boldsymbol{\ell}-\boldsymbol{\ell}')P_\kappa^\mathrm{E}(\ell)\,, \tag{3.31}$$

i.e. one does not need to distinguish between convergence power spectrum and shear power spectrum. Note that $\xi_+(\theta) = \langle\gamma\gamma^*\rangle(\theta)$, one also has $\langle\kappa\kappa^*\rangle(\theta) = \xi_+(\theta)$.

We have distinguished $\kappa^*(\theta)$ and $\kappa(\theta)$ although the physical convergence field has no imaginary part. We did so because if one reconstructs the convergence field from shear signals using (3.19), the resulting convergence could have a complex part due to the B-modes in the shear, which can come from systematical errors, higher-order effects, and noise. Here we have introduced $\kappa = \kappa^\mathrm{E} + \mathrm{i}\kappa^\mathrm{B}$, to let the complex part κ_B account for the convergence corresponding to B-mode shear signal. We shall omit the B-mode in this thesis as a default and deal mainly with the physical E-mode cosmic shear signal.

The power spectrum defined in (3.31) corresponds to the E-mode convergence only, thus we denoted it by P_κ^E. We further define

$$\left\langle\tilde{\kappa}^\mathrm{B}(\boldsymbol{\ell})\tilde{\kappa}^{\mathrm{B}*}(\boldsymbol{\ell}')\right\rangle := (2\pi)^2\delta_\mathrm{D}(\boldsymbol{\ell}-\boldsymbol{\ell}')P_\kappa^\mathrm{B}(\ell)\,, \text{ and} \tag{3.32}$$

$$\left\langle\tilde{\kappa}^\mathrm{E}(\boldsymbol{\ell})\tilde{\kappa}^{\mathrm{B}*}(\boldsymbol{\ell}')\right\rangle := (2\pi)^2\delta_\mathrm{D}(\boldsymbol{\ell}-\boldsymbol{\ell}')P_\kappa^\mathrm{EB}(\ell)\,. \tag{3.33}$$

Then one can relate the 2-pt functions to the power spectrum. Their relations are found to be

$$\begin{aligned}
\xi_+(\theta) =& \frac{1}{2\pi}\int \mathrm{d}\ell\ \ell\ J_0(\ell\theta)\ \left[P_\kappa^\mathrm{E}(\ell)+P_\kappa^\mathrm{B}(\ell)\right]\\
\xi_-(\theta) =& \frac{1}{2\pi}\int \mathrm{d}\ell\ \ell\ J_4(\ell\theta)\ \left[P_\kappa^\mathrm{E}(\ell)-P_\kappa^\mathrm{B}(\ell)\right]\\
\xi_\times(\theta) =& \frac{1}{2\pi}\int \mathrm{d}\ell\ \ell\ J_4(\ell\theta)\ P_\kappa^\mathrm{EB}(\ell)\,.
\end{aligned} \tag{3.34}$$

In general the mix term P_κ^EB should vanish since it has an odd parity (Schneider 2003), thus $\xi_\times(\theta)$ vanishes as well.

The natural components and the bispectrum

At the 3-pt level there are more combinations of shear correlation functions. Schneider & Lombardi (2003) studied these combinations of them which have simple behavior under general rotations of the coordinates. They found four combinations which should be seen as the fundamental 3-pt configuration space cosmic shear statistics, and gave them the name 'the natural components'.

Consider three points $\boldsymbol{X}_i$, $i \le l \le 3$. In general they form a triangle with sides $\boldsymbol{x}_1 = \boldsymbol{X}_3 - \boldsymbol{X}_2$, $\boldsymbol{x}_2 = \boldsymbol{X}_1 - \boldsymbol{X}_3$, and $\boldsymbol{x}_3 = \boldsymbol{X}_2 - \boldsymbol{X}_1$. The directions of the sides are defined so that $\boldsymbol{x}_1 + \boldsymbol{x}_2 + \boldsymbol{x}_3 = \boldsymbol{0}$. Unlike in the 2-pt case where one can define γ_t and $\gamma_\times$ with respect to the line linking the two points, there exists no unique natural choice of a reference point to define γ_t and $\gamma_\times$. The reference point could be chosen as any point inside the triangle. The middle of one side, the centroid and the

orthocenter have all been used in practice. For any choice, the natural components are defined as

$$\begin{aligned}
\Gamma^{(0)} &= \gamma_{\mathrm{ttt}} - \gamma_{\mathrm{t}\times\times} - \gamma_{\times\mathrm{t}\times} - \gamma_{\times\times\mathrm{t}} + \mathrm{i}\left[\gamma_{\mathrm{tt}\times} + \gamma_{\mathrm{t}\times\mathrm{t}} + \gamma_{\times\mathrm{tt}} - \gamma_{\times\times\times}\right] ,\\
\Gamma^{(1)} &= \gamma_{\mathrm{ttt}} - \gamma_{\mathrm{t}\times\times} + \gamma_{\times\mathrm{t}\times} + \gamma_{\times\times\mathrm{t}} + \mathrm{i}\left[\gamma_{\mathrm{tt}\times} + \gamma_{\mathrm{t}\times\mathrm{t}} - \gamma_{\times\mathrm{tt}} + \gamma_{\times\times\times}\right] ,\\
\Gamma^{(2)} &= \gamma_{\mathrm{ttt}} + \gamma_{\mathrm{t}\times\times} - \gamma_{\times\mathrm{t}\times} + \gamma_{\times\times\mathrm{t}} + \mathrm{i}\left[\gamma_{\mathrm{tt}\times} - \gamma_{\mathrm{t}\times\mathrm{t}} + \gamma_{\times\mathrm{tt}} + \gamma_{\times\times\times}\right] ,\\
\Gamma^{(3)} &= \gamma_{\mathrm{ttt}} + \gamma_{\mathrm{t}\times\times} + \gamma_{\times\mathrm{t}\times} - \gamma_{\times\times\mathrm{t}} + \mathrm{i}\left[-\gamma_{\mathrm{tt}\times} + \gamma_{\mathrm{t}\times\mathrm{t}} + \gamma_{\times\mathrm{tt}} + \gamma_{\times\times\times}\right] .
\end{aligned} \tag{3.35}$$

These natural components do not mix under a rotation of the coordinates, they change only by a phase factor.

Under cyclic permutation of the arguments, they behave as

$$\begin{aligned}
\Gamma^{(0)}(x_1,x_2,x_3) &= \Gamma^{(0)}(x_2,x_3,x_1) = \Gamma^{(0)}(x_3,x_1,x_2) ,\\
\Gamma^{(1)}(x_1,x_2,x_3) &= \Gamma^{(3)}(x_2,x_3,x_1) = \Gamma^{(2)}(x_3,x_1,x_2) ,\\
\Gamma^{(2)}(x_1,x_2,x_3) &= \Gamma^{(1)}(x_2,x_3,x_1) = \Gamma^{(3)}(x_3,x_1,x_2) ,\\
\Gamma^{(3)}(x_1,x_2,x_3) &= \Gamma^{(2)}(x_2,x_3,x_1) = \Gamma^{(1)}(x_3,x_1,x_2) ,
\end{aligned} \tag{3.36}$$

i.e. $\Gamma^{(0)}$ stays invariant, and the other natural components transform into each other. This suggests that if a permutation of the arguments is allowed, the four complex natural components cannot be considered as independent quantities. The degree of freedom corresponds to two complex quantities, or four real quantities.

The 3-pt shear and convergence statistics are again simply related in Fourier space. So here we introduce only the convergence bispectrum which is defined via

$$\langle\kappa(\boldsymbol{\ell_1})\kappa(\boldsymbol{\ell_2})\kappa(\boldsymbol{\ell_3})\rangle =: (2\pi)^2\,\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell_1}+\boldsymbol{\ell_2}+\boldsymbol{\ell_3})\,B(\ell_1,\ell_2,\ell_3)\,, \tag{3.37}$$

where the Dirac delta function imposes the condition that $\boldsymbol{\ell_1}+\boldsymbol{\ell_2}+\boldsymbol{\ell_3}=0$, which we call the *triangle condition*. Here we have considered the field to be both statistically homogeneous and isotropic, thus the bispectrum can be quantified by three real arguments which we have chosen to be the three side-lengths of the triangle. Another commonly used choice is two side-lengths and the angle between them.

The relations between the natural components and the convergence bispctrum have been studied and given in Schneider et al. (2005). These relations have complicated dependences on their arguments and have highly oscillatory integration kernels, which largely limits their application. In Chap 4 we will give the relations between the natural components and the 3-pt convergence correlation functions, which provides another way to link the observable shear statistics and the underlying matter density field.

The aperture mass statistics

The *aperture mass* M_{ap} is one important measure of the shear and the convergence introduced by Kaiser et al. (1994) and Schneider (1996) to circumvent the mass-sheet degeneracy problem. The aperture mass within an aperture of size θ centering on $\boldsymbol{\theta_0}$ is defined as

$$M_{\mathrm{ap}}(\theta,\boldsymbol{\theta_0}) = \int_{A_\theta} \mathrm{d}^2\vartheta\;\kappa(\boldsymbol{\vartheta})\;U_\theta(|\boldsymbol{\vartheta}-\boldsymbol{\theta_0}|)\,, \tag{3.38}$$

where A_θ is the area of the aperture, and $U_\theta(\vartheta)$ is a compensated filter function, i.e.

$$\int_{A_\theta} \mathrm{d}^2\vartheta\; U_\theta(|\boldsymbol{\vartheta}|) = 2\pi \int_0^\theta \mathrm{d}\vartheta\; \vartheta U_\theta(\vartheta) = 0\,. \tag{3.39}$$

Note that $U_\theta(\vartheta)$ is a function of both ϑ and θ. To ensure that its functional form satisfy the condition (3.39) for any ϑ and θ, it has to have the scaling

$$U_\theta(\vartheta) = \frac{1}{\theta^2}\bar{U}\left(\frac{\vartheta}{\theta}\right), \tag{3.40}$$

where $\bar{U}$ is a single argument function which needs to satisfy

$$\int_0^1 \mathrm{d}x\; x\bar{U}(x) = 0\,. \tag{3.41}$$

A valuable property of the aperture mass is that it can also written directly in terms of the shear, as

$$M_{\rm ap}(\theta, \boldsymbol{\theta_0}) = \int_{A_\theta} \mathrm{d}^2\vartheta\; Q_\theta(|\boldsymbol{\vartheta}|)\; \gamma_{\rm t}(\boldsymbol{\vartheta}; \boldsymbol{\theta_0})\,, \tag{3.42}$$

where $\gamma_{\rm t}(\boldsymbol{\vartheta}; \boldsymbol{\theta_0})$ denotes the tangential shear at position $\boldsymbol{\vartheta}$ relative to the point $\boldsymbol{\theta_0}$, and

$$Q_\theta(\vartheta) = \frac{2}{\vartheta^2}\int_0^\vartheta \mathrm{d}\vartheta'\; \vartheta' U_\theta(\vartheta') - U(\vartheta)\,. \tag{3.43}$$

The form of the Q filter can be derived from the relation between the shear and the convergence (see Squires & Kaiser 1996, Schneider 1996, and Schneider & Bartelmann 1997).

The aperture mass $M_{\rm ap}$ is a real quantity and is sensitive only to the tangential shear. One can expand it to a complex quantity as ($\boldsymbol{\theta_0}$ has been put to the origin for simplicity)

$$\begin{aligned} M(\theta) :=& M_{\rm ap}(\theta) + \mathrm{i}M_\perp(\theta) = \int_{A_\theta} \mathrm{d}^2\vartheta\; Q_\theta(|\boldsymbol{\vartheta}|)\,[\gamma_{\rm t}(\boldsymbol{\vartheta}) + \mathrm{i}\gamma_\times(\boldsymbol{\vartheta})] \\ =& -\int_{A_\theta} \mathrm{d}^2\vartheta\; Q_\theta(|\boldsymbol{\vartheta}|)\; \gamma(\boldsymbol{\vartheta})\; \mathrm{e}^{-2\mathrm{i}\phi}\,, \end{aligned} \tag{3.44}$$

with ϕ being the polar angle of $\boldsymbol{\vartheta}$. While the real part of $M(\theta)$ corresponds to the physical convergence $\kappa^{\rm E}$ (3.38), the imaginary part of it corresponds to $\kappa^{\rm B}$, which vanishes in the absence of a B-mode. This suggests that the aperture mass statistics naturally allows a E/B-mode decomposition. The lensing signal which has no B-mode component enters only the $M_{\rm ap}(\theta)$. The $M_\perp(\theta)$, on the other hand, is a measure of the B-mode which quantifies systematic errors and noises.

At the 2-pt level, the dispersions of the aperture measures are related to the power spectrum by (Schneider et al. 1998)

$$\begin{aligned} \left\langle M_{\rm ap}^2(\theta)\right\rangle =& \frac{1}{2\pi}\int_0^\infty \mathrm{d}\ell\; \ell\; P_{\rm E}(\ell)\; W_{\rm ap}(\ell\theta) \\ \left\langle M_\perp^2(\theta)\right\rangle =& \frac{1}{2\pi}\int_0^\infty \mathrm{d}\ell\; \ell\; P_{\rm B}(\ell)\; W_{\rm ap}(\ell\theta)\,, \end{aligned} \tag{3.45}$$

which follow directly from (3.38) with

$$\begin{aligned} W_{\rm ap}(\ell\theta) =& \left(2\pi\int_0^\theta \mathrm{d}\vartheta\; \vartheta\; U_\theta(\vartheta) J_0(\ell\vartheta)\right)^2 \\ =& \left(2\pi\int_0^1 \mathrm{d}x\; x\; \bar{U}(x) J_0(\ell\theta x)\right)^2\,, \end{aligned} \tag{3.46}$$

where in the second step we have inserted (3.40).

One can see that the dispersion of aperture mass $\left\langle M_{\rm ap}^2 \right\rangle$ is sensitive only to the E-mode power spectrum, and $\left\langle M_\perp^2 \right\rangle$ is sensitive only to the B-mode one, as expected.

The statistics to be applied directly to weak lensing survey data is the shear correlation function since it deals easily with the complex survey geometry a lensing survey usually has. The other statistics, e.g. that of the aperture measures, need to be obtained from the shear correlation functions. At the 2-pt level, the relations between these two statistics are (Schneider et al. 2002)

$$\begin{aligned}\left\langle M_{\rm ap}^2 \right\rangle(\theta) =& \frac{1}{2}\int_0^\infty \frac{{\rm d}\vartheta\,\vartheta}{\theta^2}\left[\xi_+(\vartheta)T_+^{\rm ap}\left(\frac{\vartheta}{\theta}\right)+\xi_-(\vartheta)T_-^{\rm ap}\left(\frac{\vartheta}{\theta}\right)\right],\\ \left\langle M_\perp^2 \right\rangle(\theta) =& \frac{1}{2}\int_0^\infty \frac{{\rm d}\vartheta\,\vartheta}{\theta^2}\left[\xi_+(\vartheta)T_+^{\rm ap}\left(\frac{\vartheta}{\theta}\right)-\xi_-(\vartheta)T_-^{\rm ap}\left(\frac{\vartheta}{\theta}\right)\right].\end{aligned} \tag{3.47}$$

The forms of the filters $T_+^{\rm ap}$ and $T_-^{\rm ap}$ can be derived by combining (3.34) and (3.45), yielding

$$\begin{aligned}T_+^{\rm ap}(x) =& \int {\rm d}\eta\;\eta\;J_0(x\eta)W_{\rm ap}(\eta)\,,\\ T_-^{\rm ap}(x) =& \int {\rm d}\eta\;\eta\;J_4(x\eta)W_{\rm ap}(\eta)\,.\end{aligned} \tag{3.48}$$

For most choices of the filter U, one cannot obtain an analytical form for $T_+^{\rm ap}$ and $T_-^{\rm ap}$.

Note that the integrals in (3.47) extend in principle from zero to infinity, which causes a problem since in practice one cannot measure the shear correlation functions out to very small or very large angular separations. The difficulty at small angular separations arises since galaxies are extended sources, and it is impossible to precisely measure galaxy shapes if the images of the galaxies overlap. The difficulty at large angular separations, on the other hand, is due to the finite size of the survey. For the choices of the filter U made in Kaiser (1995); Schneider (1996); Crittenden et al. (2002), the $T_+^{\rm ap}$ and $T_-^{\rm ap}$ filters do not extend to infinity and thus remove the problem there (see e.g. Schneider et al. 2002), but they do extend to $x \to 0$, which leads to a certain mixing of E- and B-modes (Kilbinger et al. 2006). We shall give an overview of the solutions to this problem at the 2-pt level and the current situation for 3-pt statistics in Chap. 4. The relations between the 3-pt aperture statistics, the natural components and the bispectrum are given in Schneider et al. (2005).

Chapter 4

Relations between three-point configuration space shear and convergence statistics

The convergence κ and the shear γ are two basic quantities considered in gravitational lens theory. Defined as the dimensionless surface mass density, κ is a weighted projection of the 3D matter density contrast δ (3.25). The shear γ, on the other hand, is directly accessible from observations (see Sect. 3.1.2). Therefore, the theoretical framework of gravitational lensing should include the relation between configuration space κ and γ statistics as well as the one relating configuration space statistics to their Fourier space counterparts. At the level of 2-pt statistics, such relations have already been established. For 3-pt statistics, the relation between the shear 3-pt correlation functions (γ3PCFs) and the convergence bispectrum, which is the Fourier counterpart of the 3-pt convergence correlation function (κ3PCF), has been derived by Schneider et al. (2005). The other non-trivial relation, the one between γ3PCFs and κ3PCFs, is still missing. One purpose of the work described in this chapter is to establish this missing link.

How to perform E/B-mode decomposition is also a major concern of the weak lensing community. For observational data an E/B-mode decomposition provides a necessary check on the possible systematics (e.g. Crittenden et al. 2002; Pen et al. 2002). In recent years there have been several efforts to construct better statistics which allow for an E/B-mode decomposition at the 2-pt level (Schneider & Kilbinger 2007; Eifler et al. 2010; Fu & Kilbinger 2010; Schneider et al. 2010). They all use weight functions to filter the shear 2-pt correlation functions (γ2PCFs), and the condition for E/B-mode decomposition transforms to a condition on the weight functions. Such a condition at the 3-pt level is also missing so far. We will see that with the aid of the relation between the γ3PCFs and the κ3PCFs, one can easily formulate this condition.

In the first section of this chapter we show how the relation between the γ3PCF and the κ3PCF is obtained. In Sect. 4.2 we investigate the correspondence between the derived relation and already established results. We then extend our results to other γ3PCFs in Sect. 4.3, and in Sect. 4.4 we present an application of the 3-pt relations, deriving the condition for E/B-mode separation of 3-pt shear statistics. How these relations can be numerically evaluated is demonstrated in Sect. 4.5, and we conclude in Sect. 4.6. In the first two sections we ignore the B-mode and consider the convergence κ to be a real quantity, starting from Sect. 4.3 we extend the consideration to B-mode and complex κ. The content of this chapter is based on Shi et al. (2011).

4.1 Relation between three-point γ and κ correlation functions

4.1.1 The form of the relation

At the 2-pt level, the relation between the configuration space shear and convergence statistics is the $\xi_+ - \xi_-$ relation (Crittenden et al. 2002; Schneider et al. 2002)

$$\xi_-(x) = \int \mathrm{d}y\, y\, \xi_+(y) \left[\frac{4x^2 - 12y^2}{x^4} H(x-y) + \frac{\delta_\mathrm{D}^{(1)}(x-y)}{x} \right], \tag{4.1}$$

where H and $\delta_\mathrm{D}^{(1)}$ are Heaviside function and 1D Dirac delta function, respectively. The functions ξ_+ and ξ_- are defined in (3.28) and (3.29). They can also be written as $\xi_+(x) = \langle\kappa\kappa\rangle(|\boldsymbol{x}|) = \langle\gamma\gamma^*\rangle(|\boldsymbol{x}|)$ and $\xi_-(x) = \langle\gamma\gamma\rangle(\boldsymbol{x})\, \mathrm{e}^{-4\mathrm{i}\phi_x}$, with ϕ_x being the polar angle of the separation vector $\boldsymbol{x}$ (see Sect. 3.2.3). Note that the shear γ is a spin-2 quantity and thus $\langle\gamma\gamma\rangle(\boldsymbol{x})$ has a spin of 4. Being the product of $\langle\gamma\gamma\rangle(\boldsymbol{x})$ and a phase factor of $\mathrm{e}^{-4\mathrm{i}\phi_x}$, the quantity $\xi_-(x)$ no longer depends on the polar angle of $\boldsymbol{x}$.

The relation (4.1) has already taken both the statistical homogeneity and isotropy of the shear field into account and is therefore a one-dimensional relation of quantities on the real domain. The derivation of the $\xi_+ - \xi_-$ relation originates from the relation between ξ_+ and ξ_- and the convergence power spectrum P_κ (3.34), or equivalently

$$P_\kappa(\ell) = 2\pi \int_0^\infty \mathrm{d}x\, x\, \xi_+(x)\, J_0(\ell x) = 2\pi \int_0^\infty \mathrm{d}x\, x\, \xi_-(x)\, J_4(\ell x)\,. \tag{4.2}$$

Inverting one of the relations in (4.2) one can write ξ_+ and ξ_- in terms of each other, e.g.

$$\xi_-(x) = \int_0^\infty \frac{\mathrm{d}\ell\, \ell}{2\pi}\, J_4(\ell x)\, P_\kappa(\ell) = \int_0^\infty \mathrm{d}y\, y\, \xi_+(y) \int_0^\infty \mathrm{d}\ell\, \ell\, J_4(\ell x)\, J_0(\ell y)\,, \tag{4.3}$$

and the final form of the relation (4.1) can be reached by performing the 1D Bessel integral whose result can be obtained from Gradshteyn et al. (2000).

The same procedure, however, fails to work for 3-pt statistics since the corresponding Bessel integral actually consists of three integrals, and they have highly complicated dependencies on the arguments (see Schneider et al. 2005). A brute force numerical evaluation of these integrals is also extremely challenging due to the oscillatory behaviour of the Bessel functions.

Since the advantage of transforming to the Fourier plane and back no longer holds for 3-pt statistics, we attempt to stay in configuration space, which at least avoids the problem of oscillatory integrals. One can see from (4.1) that the result of the Bessel integral in (4.3) is actually not oscillatory, as expected.

The configuration space 3-pt shear correlator can be written as $\langle\gamma(\boldsymbol{X_1})\gamma(\boldsymbol{X_2})\gamma(\boldsymbol{X_3})\rangle$, with $\boldsymbol{X_i}$ being the positions on the two-dimensional (2D) plane where the shear signals are evaluated. Following the assumed statistical homogeneity of the shear field, the correlator depends only on the separations of these three positions. We choose $\boldsymbol{x_1} \equiv \boldsymbol{X_1} - \boldsymbol{X_3}$ and $\boldsymbol{x_2} \equiv \boldsymbol{X_2} - \boldsymbol{X_3}$ to be its arguments (see the leftmost sketch of Fig. 4.1) and write the correlator as $\langle\gamma\gamma\gamma\rangle(\boldsymbol{x_1}, \boldsymbol{x_2})$. After the same procedure is applied to the 3-pt convergence correlator, the relation we are interested in will be shown to be of the form

$$\langle\gamma\gamma\gamma\rangle(\boldsymbol{x_1}, \boldsymbol{x_2}) = -\frac{1}{\pi^3} \int \mathrm{d}^2y_1 \int \mathrm{d}^2y_2\, \langle\kappa\kappa\kappa\rangle(\boldsymbol{y_1}, \boldsymbol{y_2})\, G_0(\boldsymbol{x_1} - \boldsymbol{y_1}, \boldsymbol{x_2} - \boldsymbol{y_2})\,, \tag{4.4}$$

where we have defined the convolution kernel G_0 for which we need to find an explicit expression.

Writing the relation in the form of a convolution is motivated by the K-S relation (3.14) between the convergence and the shear (Kaiser & Squires 1993), which yields the result (4.4) and also allows us to express the kernel G_0 as

$$G_0(\boldsymbol{a},\boldsymbol{b}) = -\int \mathrm{d}^2v\ \mathcal{D}(\boldsymbol{v})\ \mathcal{D}(\boldsymbol{v}-\boldsymbol{a})\ \mathcal{D}(\boldsymbol{v}-\boldsymbol{b}) = \int \mathrm{d}^2v\ \frac{1}{\boldsymbol{v}^{*2}}\ \frac{1}{(\boldsymbol{v}^*-\boldsymbol{a}^*)^2}\ \frac{1}{(\boldsymbol{v}^*-\boldsymbol{b}^*)^2}\ . \tag{4.5}$$

Here, for simplicity, we have used the complex notation for the K-S kernel (3.15)

$$\mathcal{D}(\boldsymbol{z}) = -1/\boldsymbol{z}^{*2}\,, \tag{4.6}$$

i.e. we have identified the 2D separation vectors with complex numbers. Throughout the text we will use the vector and complex notations interchangeably, and use $\boldsymbol{x}$ to indicate a complex quantity, x for its absolute value, and $\boldsymbol{x}^*$ for its complex conjugate.

The integral in (4.5) is difficult to perform directly, so we first take a look at the more studied 2-pt case. The relation between 2-pt γ and κ correlation functions can be written in the same way as

$$\langle\gamma\gamma\rangle(\boldsymbol{x}) = \frac{1}{\pi^2}\int \mathrm{d}^2y\ \langle\kappa\kappa\rangle(\boldsymbol{y})\ F(\boldsymbol{x}-\boldsymbol{y})\,, \tag{4.7}$$

with

$$F(\boldsymbol{z}) = \int \mathrm{d}^2v\ \mathcal{D}(\boldsymbol{v})\ \mathcal{D}(\boldsymbol{v}-\boldsymbol{z}) = \int \mathrm{d}^2v\ \frac{1}{\boldsymbol{v}^{*2}}\ \frac{1}{(\boldsymbol{v}^*-\boldsymbol{z}^*)^2}\,. \tag{4.8}$$

Unlike the case of the $\xi_+ - \xi_-$ relation, we have not assumed a statistically isotropic field for (4.4) or (4.7). The $\xi_+ - \xi_-$ relation is actually what one should obtain after adding the assumption of isotropy to (4.7).

4.1.2 The form of the convolution kernels

Now we aim for obtaining the forms of the F and G_0 kernels, which can be seen as the 2- and 3-pt equivalence of the K-S kernel (4.6). Introducing the symbols $\partial \equiv \partial_1 + \mathrm{i}\partial_2$ and $\nabla^2 \equiv \partial_1^2 + \partial_2^2 = \partial\partial^*$, one can write (3.9) and (3.11) as

$$\kappa = \frac{1}{2}\nabla^2\psi\,, \quad \boldsymbol{\gamma} = \frac{1}{2}\partial^2\psi\,, \tag{4.9}$$

which clearly shows that both the convergence κ and the shear $\boldsymbol{\gamma}$ are second-order derivatives of the deflection potential ψ. It is then convenient to use ψ as a link between κ and γ. Using the identities $\nabla \ln|\boldsymbol{x}| = \boldsymbol{x}/|\boldsymbol{x}|^2$ and $\nabla^2 \ln|\boldsymbol{x}| = 2\pi\delta_{\mathrm{D}}^{(2)}(\boldsymbol{x})$ which hold for a 2D $\boldsymbol{x}$, one can easily verify the consistency of (4.9) with the relation (3.7) between ψ and κ (e.g. Bartelmann & Schneider 2001) which we re-write here as

$$\psi(\boldsymbol{x}) = \frac{1}{\pi}\int \mathrm{d}^2y\ \kappa(\boldsymbol{y})\ln|\boldsymbol{x}-\boldsymbol{y}|\,. \tag{4.10}$$

Applying the operator ∂^2 on both sides of (4.10) and taking (4.9) into account, one reaches the K-S relation (3.14), since $\mathcal{D}(\boldsymbol{z}) = \partial^2 \ln|\boldsymbol{z}|$.

The same procedure can be generalized to second-order statistics. The 2-pt equivalence of (4.10) is

$$\langle\psi(\boldsymbol{x_1})\psi(\boldsymbol{x_2})\rangle = \frac{1}{\pi^2}\int \mathrm{d}^2y_1\ \ln|\boldsymbol{x_1}-\boldsymbol{y_1}|\ \int \mathrm{d}^2y_2\ \ln|\boldsymbol{x_2}-\boldsymbol{y_2}|\ \langle\kappa(\boldsymbol{y_1})\kappa(\boldsymbol{y_2})\rangle\ . \tag{4.11}$$

Using the statistical homogeneity of the κ field, and re-defining the integration variables, (4.11) reduces to

$$\begin{aligned}\langle\psi(\boldsymbol{x_1})\psi(\boldsymbol{x_2})\rangle &= \frac{1}{\pi^2}\int \mathrm{d}^2y\ \langle\kappa\kappa\rangle(\boldsymbol{y})\ \int \mathrm{d}^2u\ \ln|\boldsymbol{u}|\ln|\boldsymbol{x_1}-\boldsymbol{x_2}-\boldsymbol{y}-\boldsymbol{u}| \\ &= \frac{1}{\pi^2}\int \mathrm{d}^2y\ \langle\kappa\kappa\rangle(\boldsymbol{y})\,\mathcal{F}'(\boldsymbol{x_1}-\boldsymbol{x_2}-\boldsymbol{y})\ ,\end{aligned} \tag{4.12}$$

where we have defined

$$\mathcal{F}'(\boldsymbol{z}) = \int \mathrm{d}^2u\ \ln|\boldsymbol{u}|\ln|\boldsymbol{z}-\boldsymbol{u}|\ . \tag{4.13}$$

Obviously, $\mathcal{F}'$ is infinite at every $\boldsymbol{z}$, which is related to the fact that ψ is defined only up to an additive constant. However, we shall only need the derivatives of $\mathcal{F}'$. So we define

$$\mathcal{F}(\boldsymbol{z}) = \mathcal{F}'(\boldsymbol{z}) - \mathcal{F}'(\boldsymbol{0}) = \int \mathrm{d}^2u\ \ln|\boldsymbol{u}|\ln\left(\frac{|\boldsymbol{z}-\boldsymbol{u}|}{|\boldsymbol{u}|}\right), \tag{4.14}$$

and will use $\mathcal{F}$ and $\mathcal{F}'$ interchangeably. Let φ denote the angle between $\boldsymbol{u}$ and $\boldsymbol{z}$, (4.14) can be rewritten as

$$\mathcal{F}(\boldsymbol{z}) = \frac{1}{2}\int_0^\infty \mathrm{d}u\ u\ \ln u\ \int_0^{2\pi}\mathrm{d}\varphi\ \ln\left(1-\frac{2|\boldsymbol{z}|}{u}\cos\varphi+\frac{|\boldsymbol{z}|^2}{u^2}\right). \tag{4.15}$$

The integral over φ yields zero if $|\boldsymbol{z}| < u$, and $4\pi\ln(|\boldsymbol{z}|/u)$ otherwise. Thus

$$\mathcal{F}(\boldsymbol{z}) = 2\pi\int_0^{|\boldsymbol{z}|}\mathrm{d}u\ u\ \ln u\ \ln(|\boldsymbol{z}|/u) = \frac{\pi}{2}|\boldsymbol{z}|^2(\ln|\boldsymbol{z}|-1)\ . \tag{4.16}$$

We are now ready to apply differential operators to (4.12) to get the relations of 2-pt shear and convergence statistics. As a consistency check, we first apply two ∇^2 operators to (4.12), one acting on $\boldsymbol{x_1}$ and the other on $\boldsymbol{x_2}$. According to (4.9), this turns the l.h.s. of (4.12) into $4\langle\kappa(\boldsymbol{x_1})\kappa(\boldsymbol{x_2})\rangle$. On the r.h.s. of (4.12) the operators act exclusively on $\mathcal{F}$,

$$\nabla^2_{x_1}\nabla^2_{x_2}\mathcal{F}(\boldsymbol{x_1}-\boldsymbol{x_2}-\boldsymbol{y}) = \nabla^2\nabla^2\mathcal{F}(\boldsymbol{z}) = \nabla^2(2\pi\ln|\boldsymbol{z}|) = 4\pi^2\delta_{\mathrm{D}}^{(2)}(\boldsymbol{z}), \tag{4.17}$$

with $\boldsymbol{z} = \boldsymbol{x_1}-\boldsymbol{x_2}-\boldsymbol{y}$ here. Using (4.17), one easily sees that the r.h.s. of (4.12) after the operation gives $4\langle\kappa\kappa\rangle(\boldsymbol{x_1}-\boldsymbol{x_2})$, which is equivalent to $4\langle\kappa(\boldsymbol{x_1})\kappa(\boldsymbol{x_2})\rangle$ under the assumption of statistical homogeneity of the κ field.

Now we apply the operator $\partial^2_{x_1}\partial^2_{x_2}/4$ on (4.12), which turns the l.h.s. of (4.12) into $\langle\gamma(\boldsymbol{x_1})\gamma(\boldsymbol{x_2})\rangle$. On the r.h.s. the operation again acts only on $\mathcal{F}$,

$$\frac{1}{4}\partial^2_{x_1}\partial^2_{x_2}\mathcal{F}(\boldsymbol{x_1}-\boldsymbol{x_2}-\boldsymbol{y}) = \frac{1}{4}\partial^4\mathcal{F}(\boldsymbol{z})\ , \tag{4.18}$$

also with $\boldsymbol{z} = \boldsymbol{x_1}-\boldsymbol{x_2}-\boldsymbol{y}$. Remembering the definition of the kernel F (4.7), this leads to

$$F(\boldsymbol{z}) = \int \mathrm{d}^2v\ \frac{1}{\boldsymbol{v}^{*2}}\ \frac{1}{(\boldsymbol{v}^*-\boldsymbol{z}^*)^2} = \frac{1}{4}\partial^4\mathcal{F}(\boldsymbol{z}) = 2\pi\frac{\boldsymbol{z}}{\boldsymbol{z}^{*3}}\ . \tag{4.19}$$

For the 3-pt kernel G_0 we split the integral in (4.5) into

$$\begin{aligned}&\int \mathrm{d}^2v\ \frac{1}{\boldsymbol{v}^{*2}}\ \frac{1}{(\boldsymbol{v}^*-\boldsymbol{a}^*)^2}\ \frac{1}{(\boldsymbol{v}^*-\boldsymbol{b}^*)^2} \\ &= \frac{1}{(\boldsymbol{a}^*-\boldsymbol{b}^*)^2}\int \mathrm{d}^2v\ \frac{1}{\boldsymbol{v}^{*2}}\left[\frac{1}{(\boldsymbol{v}^*-\boldsymbol{a}^*)^2}+\frac{1}{(\boldsymbol{v}^*-\boldsymbol{b}^*)^2}\right] - \frac{2}{(\boldsymbol{a}^*-\boldsymbol{b}^*)^3}\int \mathrm{d}^2v\ \frac{1}{\boldsymbol{v}^{*2}}\left[\frac{1}{\boldsymbol{v}^*-\boldsymbol{a}^*}-\frac{1}{\boldsymbol{v}^*-\boldsymbol{b}^*}\right],\end{aligned} \tag{4.20}$$

where we have assumed $\boldsymbol{a} \neq \boldsymbol{b}$. From (4.19) as well as

$$\int \mathrm{d}^2 v \, \frac{1}{\boldsymbol{v}^{*2}} \, \frac{1}{\boldsymbol{v}^* - \boldsymbol{z}^*} = \frac{1}{2} \partial^3 \mathcal{F}(\boldsymbol{z}) = -\pi \frac{\boldsymbol{z}}{\boldsymbol{z}^{*2}} \,, \tag{4.21}$$

we obtain

$$G_0(\boldsymbol{a}, \boldsymbol{b}) = \int \mathrm{d}^2 v \, \frac{1}{\boldsymbol{v}^{*2}} \, \frac{1}{(\boldsymbol{v}^* - \boldsymbol{a}^*)^2} \, \frac{1}{(\boldsymbol{v}^* - \boldsymbol{b}^*)^2} = \frac{2\pi}{(\boldsymbol{a}^* - \boldsymbol{b}^*)^2} \left(\frac{\boldsymbol{a}}{\boldsymbol{a}^{*3}} + \frac{\boldsymbol{b}}{\boldsymbol{b}^{*3}} \right) + \frac{2\pi}{(\boldsymbol{a}^* - \boldsymbol{b}^*)^3} \left(\frac{\boldsymbol{a}}{\boldsymbol{a}^{*2}} - \frac{\boldsymbol{b}}{\boldsymbol{b}^{*2}} \right) . \tag{4.22}$$

The forms of the kernels (4.19) and (4.22) hold rigorously outside their singularities (at $\boldsymbol{z} = 0$ for F; at $\boldsymbol{a} = 0$, $\boldsymbol{b} = 0$, and $\boldsymbol{a} = \boldsymbol{b}$ for G_0). One may wonder if additional delta functions exist at these singularities. We will show in Sect. 4.2 that this is not the case.

The method we used to derive the forms of the kernels (4.19) and (4.22) also allows one to derive the relations between other correlation functions of weak lensing quantities in a systematic way. We present explicit forms of some of the relations in Appendix 4.9.

4.1.3 The relations

To summarize, we have obtained:

$$\langle \gamma\gamma \rangle(\boldsymbol{x}) = \frac{2}{\pi} \int \mathrm{d}^2 y \, \langle \kappa\kappa \rangle(\boldsymbol{y}) \, \frac{\boldsymbol{y} - \boldsymbol{x}}{(\boldsymbol{y}^* - \boldsymbol{x}^*)^3} \,, \tag{4.23}$$

and

$$\begin{aligned} \langle \gamma\gamma\gamma \rangle(\boldsymbol{x_1}, \boldsymbol{x_2}) = -\frac{2}{\pi^2} \int \mathrm{d}^2 y_1 \int \mathrm{d}^2 y_2 \, \langle \kappa\kappa\kappa \rangle(\boldsymbol{y_1}, \boldsymbol{y_2}) \Bigg[\frac{1}{(\boldsymbol{y_1}^* - \boldsymbol{x_1}^* - \boldsymbol{y_2}^* + \boldsymbol{x_2}^*)^2} \\ \times \left(\frac{\boldsymbol{y_1} - \boldsymbol{x_1}}{(\boldsymbol{y_1}^* - \boldsymbol{x_1}^*)^3} + \frac{\boldsymbol{y_2} - \boldsymbol{x_2}}{(\boldsymbol{y_2}^* - \boldsymbol{x_2}^*)^3} \right) + \frac{1}{(\boldsymbol{y_1}^* - \boldsymbol{x_1}^* - \boldsymbol{y_2}^* + \boldsymbol{x_2}^*)^3} \left(\frac{\boldsymbol{y_1} - \boldsymbol{x_1}}{(\boldsymbol{y_1}^* - \boldsymbol{x_1}^*)^2} - \frac{\boldsymbol{y_2} - \boldsymbol{x_2}}{(\boldsymbol{y_2}^* - \boldsymbol{x_2}^*)^2} \right) \Bigg] . \end{aligned} \tag{4.24}$$

In these relations we have applied the statistical homogeneity of the convergence field, but not the statistical isotropy. Making use of the latter, one can derive the $\xi_+ - \xi_-$ relation from (4.23), as will be shown in Sect. 4.2.2.

4.2 Consistency checks

4.2.1 The case of uniform κ

There is a physical condition which will directly serve as a test of the $\kappa - \gamma$ relations (3.14), (4.23) and (4.24). At the 1-pt level, for the K-S relation, a uniform convergence field does not result in any shear. At the 2- and 3-pt level, the physical condition could be that a uniform $\langle \kappa\kappa \rangle$ ($\langle \kappa\kappa\kappa \rangle$) field leads to a vanishing shear correlation $\langle \gamma\gamma \rangle$ ($\langle \gamma\gamma\gamma \rangle$).

One can easily see that both the F and G_0 kernel we obtained satisfy this condition. If there are additional terms at the singularities of the kernels which contribute to the integral, a non-zero $\langle \gamma\gamma \rangle$ ($\langle \gamma\gamma\gamma \rangle$) term would be generated and the condition would not be satisfied anymore. Thus we argue that the expressions (4.23) and (4.24) are already complete.

4.2.2 Consistency with the $\xi_+ - \xi_-$ relation

Now we consider whether (4.23) is consistent with the $\xi_+ - \xi_-$ relation (4.1), which can be regarded as the isotropic form of (4.23). That the two relations are consistent is equivalent to

$$\int_0^{2\pi} \mathrm{d}\phi_y \, \frac{\boldsymbol{y}-\boldsymbol{x}}{(\boldsymbol{y}^*-\boldsymbol{x}^*)^3} = \frac{\pi}{2} \, \mathrm{e}^{4\mathrm{i}\phi_x} \left[\frac{4x^2-12y^2}{x^4} H(x-y) + \frac{\delta(x-y)}{x} \right] . \tag{4.25}$$

To verify that (4.25) indeed holds, we attempt to solve the ϕ_y-integral on the l.h.s.,

$$\int_0^{2\pi} \mathrm{d}\phi_y \, \frac{\boldsymbol{y}-\boldsymbol{x}}{(\boldsymbol{y}^*-\boldsymbol{x}^*)^3} = \mathrm{e}^{4\mathrm{i}\phi_x} \int_0^{2\pi} \mathrm{d}\phi \, \frac{y\,\mathrm{e}^{\mathrm{i}\phi}-x}{(y\,\mathrm{e}^{-\mathrm{i}\phi}-x)^3} , \tag{4.26}$$

where $\phi = \phi_y - \phi_x$ has been defined. The ϕ-integral can be carried out using the residual theorem, yielding $2\pi(x^2-3y^2)/x^4$ when $x > y$ and zero when $x < y$. One can see that this result corresponds to the Heaviside function on the r.h.s. of (4.25).

At the singularity $\boldsymbol{x} = \boldsymbol{y}$ the ϕ-integral is not well defined, which means one cannot rule out the existence of additional delta function at $x = y$ in the result of the ϕ-integral. This ambiguity can again be eliminated by using the physical condition *'a uniform $\langle\kappa\kappa\rangle$ field leads to a null shear correlation $\langle\gamma\gamma\rangle$'*, which translates to *'a constant ξ_+ yields vanishing ξ_-'* here. In this case, a delta function is indeed required to satisfy this condition, and the prefactor of the delta function can be determined to be $\pi/2x$, in consistency with (4.25).

The Kaiser-Squires relation and its isotropic form

A similar consistency exists between the K-S relation and its isotropic form. As both forms are already well-known, they can serve as a further support for our argument.

For an axisymmetric distribution of matter, i.e. $\kappa(\boldsymbol{x}) = \kappa(x)$, the following relation is established between the shear and the convergence (see e.g. Schneider et al. 1992)

$$\gamma(\boldsymbol{x}) = [\kappa(x) - \overline{\kappa}(x)] \;\; \mathrm{e}^{2\mathrm{i}\phi_x} , \tag{4.27}$$

with $\overline{\kappa}$ defined as

$$\overline{\kappa}(x) := \frac{2}{x^2} \int_0^x y \, \mathrm{d}y \, \kappa(y) . \tag{4.28}$$

This is equivalent to

$$\gamma(\boldsymbol{x}) = -\frac{1}{\boldsymbol{x}^{*2}} \int y \, \mathrm{d}y \, \kappa(y) \, \left[2H(x-y) - x \, \delta_{\mathrm{D}}^{(1)}(x-y) \right] . \tag{4.29}$$

In the case of a uniform convergence field $\kappa(x) = \text{const.}$, one can see that the integral of the Heaviside function and the delta function parts cancel each other.

The similarity between (4.29) and the $\xi_+ - \xi_-$ relation (4.1) is remarkable: they both have integrals of a Heaviside function part and a delta function part which cancel each other for constant κ and $\langle\kappa\kappa\rangle$, respectively, and the 2D correspondences of both do not have an additional delta function at their singularities.

4.2.3 Fourier transformations

In the Fourier plane the shear and the convergence differ only by a phase factor $e^{2i\beta}$ (3.17). This directly reflects the fact that γ is spin-2 while κ is spin-0, and leads to the well-known result $P_\gamma = P_\kappa$ (3.31) as well as $e^{2i\beta} = \tilde{\mathcal{D}}/\pi$ for $\boldsymbol{\ell} \neq \mathbf{0}$ (3.16).

The Fourier plane correspondences of (4.23) and (4.24) are also readily obtainable from the identity (3.17), as

$$\langle\tilde{\gamma}\tilde{\gamma}\rangle(\boldsymbol{\ell}) = e^{4i\beta}\,\langle\tilde{\kappa}\tilde{\kappa}\rangle(\boldsymbol{\ell})\,, \quad \text{for } \boldsymbol{\ell} \neq \mathbf{0}\,, \tag{4.30}$$

and

$$\langle\tilde{\gamma}\tilde{\gamma}\tilde{\gamma}\rangle(\boldsymbol{\ell_1}, \boldsymbol{\ell_2}, \boldsymbol{\ell_3}) = e^{2i(\beta_1+\beta_2+\beta_3)}\,\langle\tilde{\kappa}\tilde{\kappa}\tilde{\kappa}\rangle(\boldsymbol{\ell_1}, \boldsymbol{\ell_2}, \boldsymbol{\ell_3})\,, \quad \text{for } \boldsymbol{\ell_1}, \boldsymbol{\ell_2}, \boldsymbol{\ell_3} \neq \mathbf{0}\,, \tag{4.31}$$

with β_i denoting the polar angle of $\boldsymbol{\ell_i}$. These equations show that $e^{4i\beta} = \tilde{F}(\boldsymbol{\ell})/\pi^2$ for $\boldsymbol{\ell} \neq \mathbf{0}$, and that $e^{2i(\beta_1+\beta_2+\beta_3)} = -\tilde{G}_0(\boldsymbol{\ell_1}, \boldsymbol{\ell_2})/\pi^3$ for $\boldsymbol{\ell_1} \neq \mathbf{0} \neq \boldsymbol{\ell_2}$ and $\boldsymbol{\ell_3} = -\boldsymbol{\ell_1} - \boldsymbol{\ell_2} \neq 0$, since F/π^2 and $-G_0/\pi^3$ are the convolution kernels for the configuration space relations by their definitions.

In Appendix 4.8 we show explicitly that the Fourier transforms of F/π^2 and $-G_0/\pi^3$, with F and G_0 given in (4.19) and (4.22), are indeed $e^{4i\beta}$ and $e^{2i(\beta_1+\beta_2+\beta_3)}$, respectively. However one cannot obtain the forms of F and G_0 kernels simply through inverse Fourier transforming the phase factors $e^{4i\beta}$ and $e^{2i(\beta_1+\beta_2+\beta_3)}$. This is due to the fact that the Fourier inversion theorem is valid strictly only for square-integrable functions, which is not the case for the phase factors. The same situation occurs for the K-S kernel $\mathcal{D}$.

4.3 The other shear three-point functions

Until now we have considered only the 3PCF of shear itself $\langle\gamma(\boldsymbol{X_1})\gamma(\boldsymbol{X_2})\gamma(\boldsymbol{X_3})\rangle$, which is one of the four independent possible combinations considering that γ is a complex quantity. The other three are $\langle\gamma^*(\boldsymbol{X_1})\gamma(\boldsymbol{X_2})\gamma(\boldsymbol{X_3})\rangle$, $\langle\gamma(\boldsymbol{X_1})\gamma^*(\boldsymbol{X_2})\gamma(\boldsymbol{X_3})\rangle$, and $\langle\gamma(\boldsymbol{X_1})\gamma(\boldsymbol{X_2})\gamma^*(\boldsymbol{X_3})\rangle$, according to the choice made in Schneider & Lombardi (2003). Following Schneider et al. (2005), we denote these four γ3PCFs by $\Gamma^{(i)}_{\rm cart}(\boldsymbol{x_1}, \boldsymbol{x_2})$ ($i = 0, 1, 2, 3$), with 'cart' emphasizing that the shear is measured in Cartesian coordinates, $\Gamma^{(0)}_{\rm cart} \equiv \langle\gamma\gamma\gamma\rangle$, and $\Gamma^{(i)}_{\rm cart}$ ($i = 1, 2, 3$) corresponding to the γ3PCF with γ^* at position $\boldsymbol{X_i}$. Since we have considered statistical homogeneity of the shear field, the $\Gamma_{\rm cart}$'s depend only on the separation vectors of the position $\boldsymbol{X_1}$, $\boldsymbol{X_2}$, and $\boldsymbol{X_3}$. The other γ3PCFs, i.e. those with two or three γ^*'s, can be obtained by taking the complex conjugate of the $\Gamma_{\rm cart}$'s.

Note that $\Gamma^{(1)}_{\rm cart}(\boldsymbol{x_1}, \boldsymbol{x_2}) \equiv \langle\gamma^*\gamma\gamma\rangle(\boldsymbol{x_1}, \boldsymbol{x_2})$, $\Gamma^{(2)}_{\rm cart}(\boldsymbol{x_1}, \boldsymbol{x_2}) \equiv \langle\gamma\gamma^*\gamma\rangle(\boldsymbol{x_1}, \boldsymbol{x_2})$, and $\Gamma^{(3)}_{\rm cart}(\boldsymbol{x_1}, \boldsymbol{x_2}) \equiv \langle\gamma\gamma\gamma^*\rangle(\boldsymbol{x_1}, \boldsymbol{x_2})$ are different functions, since $\boldsymbol{x_1}$ ($\boldsymbol{x_2}$) is defined to be the difference of the positions of the first (second) and the third γ in the bracket. Due to the same reason, they can be transformed into each other through permutations and flips of the vertices of the triangle formed by their arguments (see Fig. 4.1), and thus are not independent if argument permutations and flips are allowed. As an example, one has

$$\begin{aligned}
&\langle\gamma^*(\boldsymbol{X_1})\gamma(\boldsymbol{X_2})\gamma(\boldsymbol{X_3})\rangle \equiv \Gamma^{(1)}_{\rm cart}(\boldsymbol{x_1}, \boldsymbol{x_2})\\
=&\langle\gamma^*(\boldsymbol{X_1})\gamma(\boldsymbol{X_3})\gamma(\boldsymbol{X_2})\rangle \equiv \Gamma^{(1)}_{\rm cart}(\boldsymbol{x_1} - \boldsymbol{x_2}, -\boldsymbol{x_2})\\
=&\langle\gamma(\boldsymbol{X_2})\gamma^*(\boldsymbol{X_1})\gamma(\boldsymbol{X_3})\rangle \equiv \Gamma^{(2)}_{\rm cart}(\boldsymbol{x_2}, \boldsymbol{x_1})\\
=&\langle\gamma(\boldsymbol{X_3})\gamma^*(\boldsymbol{X_1})\gamma(\boldsymbol{X_2})\rangle \equiv \Gamma^{(2)}_{\rm cart}(-\boldsymbol{x_2}, \boldsymbol{x_1} - \boldsymbol{x_2})\\
=&\langle\gamma(\boldsymbol{X_2})\gamma(\boldsymbol{X_3})\gamma^*(\boldsymbol{X_1})\rangle \equiv \Gamma^{(3)}_{\rm cart}(\boldsymbol{x_2} - \boldsymbol{x_1}, -\boldsymbol{x_1})\\
=&\langle\gamma(\boldsymbol{X_3})\gamma(\boldsymbol{X_2})\gamma^*(\boldsymbol{X_1})\rangle \equiv \Gamma^{(3)}_{\rm cart}(-\boldsymbol{x_1}, \boldsymbol{x_2} - \boldsymbol{x_1})\,,
\end{aligned} \tag{4.32}$$

where different lines correspond to different ways of labeling the same triangle with side lengths x_1, x_2, and $|\boldsymbol{x_1} - \boldsymbol{x_2}|$. The same permutations and flips also reveal the inherent symmetry of $\Gamma^{(0)}_{\rm cart}$,

$$\begin{aligned}&\Gamma^{(0)}_{\rm cart}(\boldsymbol{x_1}, \boldsymbol{x_2}) = \Gamma^{(0)}_{\rm cart}(\boldsymbol{x_1} - \boldsymbol{x_2}, -\boldsymbol{x_2}) = \Gamma^{(0)}_{\rm cart}(\boldsymbol{x_2}, \boldsymbol{x_1})\\ =&\Gamma^{(0)}_{\rm cart}(-\boldsymbol{x_2}, \boldsymbol{x_1} - \boldsymbol{x_2}) = \Gamma^{(0)}_{\rm cart}(\boldsymbol{x_2} - \boldsymbol{x_1}, -\boldsymbol{x_1}) = \Gamma^{(0)}_{\rm cart}(-\boldsymbol{x_1}, \boldsymbol{x_2} - \boldsymbol{x_1})\,.\end{aligned} \tag{4.33}$$

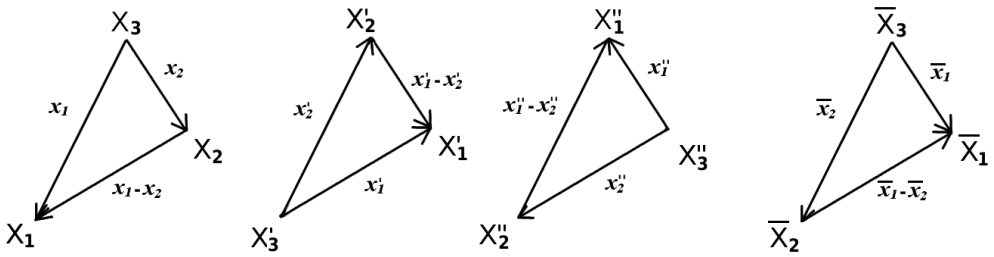


Figure 4.1: Definition of the geometry of a triangle (the leftmost sketch) and how it changes under permutations (the first three sketches from the left) and flip (the leftmost and the rightmost sketch) of the vertices.

In the case that the shear is measured relative to a center of the triangle, $\Gamma^{(i)}_{\rm cart}$ transforms to $\Gamma^{(i)}$, the natural components of the γ3PCF as defined in Schneider & Lombardi (2003). For a general triangle configuration, all four $\Gamma_{\rm cart}$'s are expected to be non-zero, thus all of them should be used to exploit the full 3-pt information of cosmic shear.

Before relating the other $\Gamma_{\rm cart}$'s to the κ3PCFs, we extend κ to a complex quantity $\kappa = \kappa^{\rm E} + {\rm i}\kappa^{\rm B}$. Although the physical convergence is a real quantity, the convergence field corresponding to the measured shear signals can have an imaginary part due to e.g. systematical errors and noise. The shear component which corresponds to this unphysical imaginary part of the convergence field is identified as the B-mode, on which we will elaborate more in Sect. 4.5. When taking the B-mode into consideration, the 3-pt correlation functions of the convergence field can be written as

$$\begin{aligned}K^{(0)} \equiv \langle\kappa\kappa\kappa\rangle &= \langle\kappa^{\rm E}\kappa^{\rm E}\kappa^{\rm E}\rangle + {\rm i}\langle\kappa^{\rm B}\kappa^{\rm E}\kappa^{\rm E}\rangle + {\rm i}\langle\kappa^{\rm E}\kappa^{\rm B}\kappa^{\rm E}\rangle + {\rm i}\langle\kappa^{\rm E}\kappa^{\rm E}\kappa^{\rm B}\rangle\\ &\quad - \langle\kappa^{\rm E}\kappa^{\rm B}\kappa^{\rm B}\rangle - \langle\kappa^{\rm B}\kappa^{\rm E}\kappa^{\rm B}\rangle - \langle\kappa^{\rm B}\kappa^{\rm B}\kappa^{\rm E}\rangle - {\rm i}\langle\kappa^{\rm B}\kappa^{\rm B}\kappa^{\rm B}\rangle\,,\\ K^{(1)} \equiv \langle\kappa^*\kappa\kappa\rangle &= \langle\kappa^{\rm E}\kappa^{\rm E}\kappa^{\rm E}\rangle - {\rm i}\langle\kappa^{\rm B}\kappa^{\rm E}\kappa^{\rm E}\rangle + {\rm i}\langle\kappa^{\rm E}\kappa^{\rm B}\kappa^{\rm E}\rangle + {\rm i}\langle\kappa^{\rm E}\kappa^{\rm E}\kappa^{\rm B}\rangle\\ &\quad - \langle\kappa^{\rm E}\kappa^{\rm B}\kappa^{\rm B}\rangle + \langle\kappa^{\rm B}\kappa^{\rm E}\kappa^{\rm B}\rangle + \langle\kappa^{\rm B}\kappa^{\rm B}\kappa^{\rm E}\rangle + {\rm i}\langle\kappa^{\rm B}\kappa^{\rm B}\kappa^{\rm B}\rangle\,,\\ K^{(2)} \equiv \langle\kappa\kappa^*\kappa\rangle &= \langle\kappa^{\rm E}\kappa^{\rm E}\kappa^{\rm E}\rangle + {\rm i}\langle\kappa^{\rm B}\kappa^{\rm E}\kappa^{\rm E}\rangle - {\rm i}\langle\kappa^{\rm E}\kappa^{\rm B}\kappa^{\rm E}\rangle + {\rm i}\langle\kappa^{\rm E}\kappa^{\rm E}\kappa^{\rm B}\rangle\\ &\quad + \langle\kappa^{\rm E}\kappa^{\rm B}\kappa^{\rm B}\rangle - \langle\kappa^{\rm B}\kappa^{\rm E}\kappa^{\rm B}\rangle + \langle\kappa^{\rm B}\kappa^{\rm B}\kappa^{\rm E}\rangle + {\rm i}\langle\kappa^{\rm B}\kappa^{\rm B}\kappa^{\rm B}\rangle\,,\\ K^{(3)} \equiv \langle\kappa\kappa\kappa^*\rangle &= \langle\kappa^{\rm E}\kappa^{\rm E}\kappa^{\rm E}\rangle + {\rm i}\langle\kappa^{\rm B}\kappa^{\rm E}\kappa^{\rm E}\rangle + {\rm i}\langle\kappa^{\rm E}\kappa^{\rm B}\kappa^{\rm E}\rangle - {\rm i}\langle\kappa^{\rm E}\kappa^{\rm E}\kappa^{\rm B}\rangle\\ &\quad + \langle\kappa^{\rm E}\kappa^{\rm B}\kappa^{\rm B}\rangle + \langle\kappa^{\rm B}\kappa^{\rm E}\kappa^{\rm B}\rangle - \langle\kappa^{\rm B}\kappa^{\rm B}\kappa^{\rm E}\rangle + {\rm i}\langle\kappa^{\rm B}\kappa^{\rm B}\kappa^{\rm B}\rangle\,.\end{aligned} \tag{4.34}$$

Apart from the E-mode $\langle\kappa^{\rm E}\kappa^{\rm E}\kappa^{\rm E}\rangle$ term, there are still additional B-mode contributions to the real parts of the K's, namely $\langle\kappa^{\rm E}\kappa^{\rm B}\kappa^{\rm B}\rangle$, $\langle\kappa^{\rm B}\kappa^{\rm E}\kappa^{\rm B}\rangle$, and $\langle\kappa^{\rm B}\kappa^{\rm B}\kappa^{\rm E}\rangle$. The imaginary part of the K's are composed of the parity violating terms which are expected to vanish due to parity symmetry (Schneider 2003). The property of $K^{(i)}$ under permutations and flips of the vertices of the triangle formed by their arguments is the same as that of $\Gamma^{(i)}_{\rm cart}$.

Similar to (4.4), the relations between $\Gamma^{(i)}_{\rm cart}$ and $K^{(i)}$ for $i = 1, 2, 3$ can be written as

$$\Gamma^{(1)}_{\rm cart}(\boldsymbol{x_1}, \boldsymbol{x_2}) = -\frac{1}{\pi^3} \int \mathrm{d}^2 y_1 \int \mathrm{d}^2 y_2 \; K^{(1)}(\boldsymbol{y_1}, \boldsymbol{y_2}) \; G_1(\boldsymbol{x_1} - \boldsymbol{y_1}, \boldsymbol{x_2} - \boldsymbol{y_2}) \,, \tag{4.35}$$

$$\Gamma^{(2)}_{\rm cart}(\boldsymbol{x_1}, \boldsymbol{x_2}) = -\frac{1}{\pi^3} \int \mathrm{d}^2 y_1 \int \mathrm{d}^2 y_2 \; K^{(2)}(\boldsymbol{y_1}, \boldsymbol{y_2}) \; G_2(\boldsymbol{x_1} - \boldsymbol{y_1}, \boldsymbol{x_2} - \boldsymbol{y_2}) \,, \tag{4.36}$$

and

$$\Gamma^{(3)}_{\rm cart}(\boldsymbol{x_1}, \boldsymbol{x_2}) = -\frac{1}{\pi^3} \int \mathrm{d}^2 y_1 \int \mathrm{d}^2 y_2 \; K^{(3)}(\boldsymbol{y_1}, \boldsymbol{y_2}) \; G_3(\boldsymbol{x_1} - \boldsymbol{y_1}, \boldsymbol{x_2} - \boldsymbol{y_2}) \,, \tag{4.37}$$

where the convolution kernels G_1, G_2, and G_3 have been defined. Again with the aid of the K-S relation, we can write these convolution kernels as

$$\begin{aligned} G_1(\boldsymbol{a}, \boldsymbol{b}) = -\int \mathrm{d}^2 \nu \; \mathcal{D}(\boldsymbol{\nu}) \; \mathcal{D}^*(\boldsymbol{\nu} - \boldsymbol{a}) \; \mathcal{D}(\boldsymbol{\nu} - \boldsymbol{b}) &= \int \mathrm{d}^2 \nu \; \frac{1}{\boldsymbol{\nu}^{*2}} \; \frac{1}{(\boldsymbol{\nu} - \boldsymbol{a})^2} \; \frac{1}{(\boldsymbol{\nu}^* - \boldsymbol{b}^*)^2} \\ &= \int \mathrm{d}^2 \nu \; \frac{1}{\boldsymbol{\nu}^2} \; \frac{1}{(\boldsymbol{\nu}^* + \boldsymbol{a}^*)^2} \; \frac{1}{(\boldsymbol{\nu}^* + \boldsymbol{a}^* - \boldsymbol{b}^*)^2} \,, \end{aligned} \tag{4.38}$$

$$\begin{aligned} G_2(\boldsymbol{a}, \boldsymbol{b}) = -\int \mathrm{d}^2 \nu \; \mathcal{D}(\boldsymbol{\nu}) \; \mathcal{D}(\boldsymbol{\nu} - \boldsymbol{a}) \; \mathcal{D}^*(\boldsymbol{\nu} - \boldsymbol{b}) &= \int \mathrm{d}^2 \nu \; \frac{1}{\boldsymbol{\nu}^{*2}} \; \frac{1}{(\boldsymbol{\nu}^* - \boldsymbol{a}^*)^2} \; \frac{1}{(\boldsymbol{\nu} - \boldsymbol{b})^2} \\ &= \int \mathrm{d}^2 \nu \; \frac{1}{\boldsymbol{\nu}^2} \; \frac{1}{(\boldsymbol{\nu}^* - \boldsymbol{a}^* + \boldsymbol{b}^*)^2} \; \frac{1}{(\boldsymbol{\nu}^* + \boldsymbol{b}^*)^2} \,, \end{aligned} \tag{4.39}$$

and

$$G_3(\boldsymbol{a}, \boldsymbol{b}) = -\int \mathrm{d}^2 \nu \; \mathcal{D}^*(\boldsymbol{\nu}) \; \mathcal{D}(\boldsymbol{\nu} - \boldsymbol{a}) \; \mathcal{D}(\boldsymbol{\nu} - \boldsymbol{b}) = \int \mathrm{d}^2 \nu \; \frac{1}{\boldsymbol{\nu}^2} \; \frac{1}{(\boldsymbol{\nu}^* - \boldsymbol{a}^*)^2} \; \frac{1}{(\boldsymbol{\nu}^* - \boldsymbol{b}^*)^2} \,. \tag{4.40}$$

When $\boldsymbol{a} \neq \boldsymbol{b}$, the product of the three terms in the integrand of (4.40) can be split into products of two, as

$$\begin{aligned} &\frac{1}{\boldsymbol{\nu}^2} \; \frac{1}{(\boldsymbol{\nu}^* - \boldsymbol{a}^*)^2} \; \frac{1}{(\boldsymbol{\nu}^* - \boldsymbol{b}^*)^2} \\ =& \frac{1}{(\boldsymbol{a}^* - \boldsymbol{b}^*)^2} \frac{1}{\boldsymbol{\nu}^2} \left[\frac{1}{(\boldsymbol{\nu}^* - \boldsymbol{a}^*)^2} + \frac{1}{(\boldsymbol{\nu}^* - \boldsymbol{b}^*)^2} \right] - \frac{2}{(\boldsymbol{a}^* - \boldsymbol{b}^*)^3} \frac{1}{\boldsymbol{\nu}^2} \left[\frac{1}{\boldsymbol{\nu}^* - \boldsymbol{a}^*} - \frac{1}{\boldsymbol{\nu}^* - \boldsymbol{b}^*} \right] . \end{aligned} \tag{4.41}$$

These terms are also obtainable from doing derivatives to the kernel $\mathcal{F}$,

$$\int \mathrm{d}^2 \nu \; \frac{1}{\boldsymbol{\nu}^2} \frac{1}{(\boldsymbol{\nu}^* - \boldsymbol{a}^*)^2} = \frac{1}{4} \partial^2 \partial^{*2} \mathcal{F}(\boldsymbol{a}) = \pi^2 \delta^{(2)}(\boldsymbol{a}) \,, \tag{4.42}$$

$$\int \mathrm{d}^2 \nu \; \frac{1}{\boldsymbol{\nu}^2} \frac{1}{\boldsymbol{\nu}^* - \boldsymbol{a}^*} = \frac{1}{2} \partial \partial^{*2} \mathcal{F}(\boldsymbol{a}) = \frac{\pi}{\boldsymbol{a}} \,. \tag{4.43}$$

This way we obtain the form of the convolution kernel G_3. The forms for the kernel G_1 and G_2 can be obtained likewise. The results are

$$G_1(\boldsymbol{a}, \boldsymbol{b}) = \frac{\pi^2}{\boldsymbol{b}^{*2}} \left[\delta^{(2)}_{\rm D}(\boldsymbol{a}) + \delta^{(2)}_{\rm D}(\boldsymbol{b} - \boldsymbol{a}) \right] - \frac{2\pi}{\boldsymbol{b}^{*3}} \left(\frac{1}{\boldsymbol{a}} + \frac{1}{\boldsymbol{b} - \boldsymbol{a}} \right) , \tag{4.44}$$

$$G_2(\boldsymbol{a}, \boldsymbol{b}) = \frac{\pi^2}{\boldsymbol{a}^{*2}} \left[\delta^{(2)}_{\rm D}(\boldsymbol{a} - \boldsymbol{b}) + \delta^{(2)}_{\rm D}(\boldsymbol{b}) \right] - \frac{2\pi}{\boldsymbol{a}^{*3}} \left(\frac{1}{\boldsymbol{a} - \boldsymbol{b}} + \frac{1}{\boldsymbol{b}} \right) , \tag{4.45}$$

$$G_3(\boldsymbol{a},\boldsymbol{b}) = \frac{\pi^2}{(\boldsymbol{a}^*-\boldsymbol{b}^*)^2}\left[\delta_{\rm D}^{(2)}(\boldsymbol{a})+\delta_{\rm D}^{(2)}(\boldsymbol{b})\right] - \frac{2\pi}{(\boldsymbol{a}^*-\boldsymbol{b}^*)^3}\left(\frac{1}{\boldsymbol{a}}-\frac{1}{\boldsymbol{b}}\right) . \tag{4.46}$$

The symmetries in the $\Gamma_{\rm cart}$'s and K's are also reflected in the G kernels. One can verify that $G_2(\boldsymbol{a},\boldsymbol{b}) = G_1(\boldsymbol{b}-\boldsymbol{a},-\boldsymbol{a})$, $G_3(\boldsymbol{a},\boldsymbol{b}) = G_1(-\boldsymbol{b},\boldsymbol{a}-\boldsymbol{b})$ as results of the symmetry under permutations, and $G_2(\boldsymbol{a},\boldsymbol{b}) = G_1(\boldsymbol{b},\boldsymbol{a})$, $G_3(\boldsymbol{a},\boldsymbol{b}) = G_3(\boldsymbol{b},\boldsymbol{a})$ as results of the symmetry under flips, in consistency with (4.32). Similarly, one has $G_0(\boldsymbol{a},\boldsymbol{b}) = G_0(\boldsymbol{b}-\boldsymbol{a},-\boldsymbol{a}) = G_0(-\boldsymbol{b},\boldsymbol{a}-\boldsymbol{b}) = G_0(\boldsymbol{b},\boldsymbol{a})$, in consistency with (4.33).

4.4 Inverse relations

So far we have obtained the expressions of the four γ3PCFs as functions of the κ3PCFs. Written in a short form, they are

$$\Gamma_{\rm cart}^{(i)} = -\frac{1}{\pi^3} G_i * K^{(i)} , \tag{4.47}$$

where i runs from 0 to 3. The forms of the G_i kernels are given by (4.22), (4.44), (4.45), and (4.46).

These relations can be inverted. We define the kernels of the inverse relations to be G_i', i.e.

$$K^{(i)} = -\frac{1}{\pi^3} G_i' * \Gamma_{\rm cart}^{(i)} . \tag{4.48}$$

Using the convolution theorem, it is apparent from (4.47) and (4.48) that

$$\left(-\frac{1}{\pi^3}\tilde{G}_i\right)\cdot\left(-\frac{1}{\pi^3}\tilde{G}_i'\right) = 1 . \tag{4.49}$$

From the corresponding Fourier plane relations of (4.47), we also know

$$-\frac{\tilde{G}_0}{\pi^3} = {\rm e}^{2{\rm i}(\beta_1+\beta_2+\beta_3)} , \quad -\frac{\tilde{G}_1}{\pi^3} = {\rm e}^{2{\rm i}(-\beta_1+\beta_2+\beta_3)} , \quad -\frac{\tilde{G}_2}{\pi^3} = {\rm e}^{2{\rm i}(\beta_1-\beta_2+\beta_3)} , \quad -\frac{\tilde{G}_3}{\pi^3} = {\rm e}^{2{\rm i}(\beta_1+\beta_2-\beta_3)} , \tag{4.50}$$

which implies

$$\tilde{G}_0\tilde{G}_0^* = \tilde{G}_1\tilde{G}_1^* = \tilde{G}_2\tilde{G}_2^* = \tilde{G}_3\tilde{G}_3^* = \pi^6 . \tag{4.51}$$

Comparing (4.49) and (4.51) one has

$$\tilde{G}_i' = \tilde{G}_i^* , \tag{4.52}$$

and further,

$$G_i' = G_i^* , \tag{4.53}$$

i.e. the convolution kernel for the inverse relation is the complex conjugate of the original kernel.

This property of the convolution kernel has its root in the fact that $\tilde{\gamma}$ and $\tilde{\kappa}$ differ only by a phase factor. This fact also endows the convolution kernels in the 1-pt and 2-pt relations between γ and κ with the same property. As is well known for the 1-pt relation, the inverse relation of the K-S relation (3.14) is (3.19), where the convolution kernel is the complex conjugate of the K-S kernel $\mathcal{D}$. The inverse relation of the 2-pt relation (4.7) can also be shown to be

$$\langle\kappa\kappa\rangle(\boldsymbol{x}) = \frac{1}{\pi^2}\int {\rm d}^2y\, \langle\gamma\gamma\rangle(\boldsymbol{y})\, F^*(\boldsymbol{x}-\boldsymbol{y}) . \tag{4.54}$$

4.5 Condition of three-point E/B decomposition

Being mathematically a polarization field, the shear field can be decomposed into a E-mode and a B-mode (see Sect. 3.2.2). Performing such a decomposition when treating cosmic shear data has long been recognized as a necessity, since it provides a valuable check on the possible systematics (e.g. Crittenden et al. 2002; Pen et al. 2002).

The E/B-mode decomposition can be done either on the shear field itself (e.g. Bunn et al. 2003; Bunn 2010), or at the level of correlation functions (e.g. Schneider 2006). The complex survey geometry after masking, which is especially characteristic for a lensing survey (e.g. Erben et al. 2009), renders the first option barely feasible, and singles out the correlation function as the basic statistic to be applied directly to the data. Thus the natural way to perform the E/B-mode decomposition on cosmic shear data is to derive statistics based on the shear correlation functions.

A commonly used statistic for this purpose is the aperture mass statistic (see Sect. 3.2.3) which can be expressed as a linear combination of ξ_+ and ξ_- (3.47), where the forms of the weight functions $T_+^{\rm ap}$ and $T_-^{\rm ap}$ are given explicitly in Schneider et al. (2002). The chosen forms of the weight functions guarantee that $\left\langle M_{\rm ap}^2 \right\rangle$ responds only to the E-mode and $\left\langle M_\perp^2 \right\rangle$ only to the B-mode.

The aperture mass statistics has been generalized to 3-pt level by Jarvis et al. (2004) and Kilbinger & Schneider (2005), and is the only statistics available up to now which allows an E/B-mode decomposition at the 3-pt level. However, as found by Kilbinger et al. (2006), it cannot ensure a clean E/B-mode decomposition when applied to real data. The lack of shear-correlation measurements on small and large scales, which arises from the inability of shape measurement for close projected galaxy pairs and the finite field size, prohibits one from performing the integral in (3.47) from zero to infinity, and thus introduces a mixing of the E- and B-modes.

In recent years, there have been several efforts to construct better statistics which allow E/B-mode decomposition (Schneider & Kilbinger 2007; Eifler et al. 2010; Fu & Kilbinger 2010; Schneider et al. 2010), all of them focusing on the cosmic shear 2-pt statistics. These new statistics are based on the idea that the weight functions $T_+^{\rm ap}$ and $T_-^{\rm ap}$ used in the aperture mass statistics are just one example out of the many possibilities. In general one can define second-order statistics in the form (Schneider & Kilbinger 2007)

$$\begin{aligned} \mathrm{EE} &= \int_0^\infty \vartheta \, \mathrm{d}\vartheta \, \left[\xi_+(\vartheta) T_+(\vartheta) + \xi_-(\vartheta) T_-(\vartheta)\right] , \\ \mathrm{BB} &= \int_0^\infty \vartheta \, \mathrm{d}\vartheta \, \left[\xi_+(\vartheta) T_+(\vartheta) - \xi_-(\vartheta) T_-(\vartheta)\right] , \end{aligned} \tag{4.55}$$

for which the condition that EE responds only to E-mode and BB only to B-mode is found to be

$$\begin{aligned} &\int_0^\infty \vartheta \, \mathrm{d}\vartheta \; T_+(\vartheta) J_0(\ell\vartheta) = \int_0^\infty \vartheta \, \mathrm{d}\vartheta \; T_-(\vartheta) J_4(\ell\vartheta) \,, \text{ or equivalently} \\ &\qquad T_+(\vartheta) = T_-(\vartheta) + \int_\vartheta^\infty \theta \, \mathrm{d}\theta \; T_-(\theta) \left(\frac{4}{\theta^2} - \frac{12\vartheta^2}{\theta^4} \right) . \end{aligned} \tag{4.56}$$

Note that instead of being functions of the separation length as the aperture mass statistics, EE and BB are just numbers. At first sight one seems to have reduced the information quantity by integrating over the scale dependence in (4.55). In fact, the information can be easily regained by constructing a set of weight functions satisfying (4.56). As one example, $\left\langle M_{\rm ap}^2 \right\rangle(\theta)$ and $\left\langle M_\perp^2 \right\rangle(\theta)$ for any θ value can be reconstructed in the framework of (4.55) by specifying $T_+(\vartheta) = T_+^{\rm ap}(\vartheta/\theta)/\theta^2$ and $T_-(\vartheta) = T_-^{\rm ap}(\vartheta/\theta)/\theta^2$.

Since the condition for E/B-mode decomposition (4.56) still leaves large freedom for the choice of the weight functions, one can construct statistics which fulfill additional constraints, e.g. a finite support over the separation length. If one requires that EE and BB respond only to $\xi_+(\vartheta)$ and $\xi_-(\vartheta)$ with $\vartheta_{\rm min} < \vartheta < \vartheta_{\rm max}$, where $\vartheta_{\rm min}$ and $\vartheta_{\rm max}$ are the chosen small- and large-scale cutoff, $T_+(\vartheta)$ and $T_-(\vartheta)$ must vanish outside the same range. Since T_+ and T_- are interrelated by (4.56), one can specify only one of them to satisfy this constraint. The requirement that the other weight function also vanishes outside the specified range needs to be put as additional integral constraints. As shown by Schneider & Kilbinger (2007), if one chooses T_- to vanish for $\vartheta < \vartheta_{\rm min}$ and $\vartheta > \vartheta_{\rm max}$, then to allow an E/B-mode decomposition on a finite interval $\vartheta_{\rm min} < \vartheta < \vartheta_{\rm max}$, T_- has to satisfy additionally,

$$\int_{\vartheta_{\rm min}}^{\vartheta_{\rm max}} \frac{{\rm d}\vartheta}{\vartheta}\, T_-(\vartheta) = 0 = \int_{\vartheta_{\rm min}}^{\vartheta_{\rm max}} \frac{{\rm d}\vartheta}{\vartheta^3}\, T_-(\vartheta)\,, \tag{4.57}$$

which would guarantee that T_+ vanishes for $\vartheta < \vartheta_{\rm min}$ and $\vartheta > \vartheta_{\rm max}$.

Similar statistics are needed at the 3-pt level as well. The first step required is to formulate the conditions for 3-pt weight functions to allow E/B-mode decomposition, in analogy to (4.56). As we will show in this section, the relations between the γ3PCFs and κ3PCFs that we derived provide a natural way of formulating such conditions.

A pure E-mode shear 3-pt statistics is related only to the E-mode κ3PCF $\left\langle \kappa^{\rm E}\kappa^{\rm E}\kappa^{\rm E}\right\rangle$ but not to other 3PCFs with $\kappa^{\rm B}$ contribution. Therefore we first write the 3PCFs of $\kappa^{\rm E}$ and $\kappa^{\rm B}$ as linear combinations of the real and imaginary parts of the $K^{(i)}$'s, using (4.34), and then relate them with the $\Gamma_{\rm cart}$'s through (4.48), as the $\Gamma_{\rm cart}$'s are the directly measurable statistics from a lensing survey. The results read

$$\begin{aligned}
\langle\kappa^{\rm E}\kappa^{\rm E}\kappa^{\rm E}\rangle &= \frac{1}{4}Re\left[\, K^{(0)} + K^{(1)} + K^{(2)} + K^{(3)}\,\right] \\
&= -\frac{1}{4\pi^3}Re\left[\, G_0^* * \Gamma^{(0)}_{\rm cart} + G_1^* * \Gamma^{(1)}_{\rm cart} + G_2^* * \Gamma^{(2)}_{\rm cart} + G_3^* * \Gamma^{(3)}_{\rm cart}\,\right], \\
\langle\kappa^{\rm E}\kappa^{\rm B}\kappa^{\rm B}\rangle &= \frac{1}{4}Re\left[-K^{(0)} - K^{(1)} + K^{(2)} + K^{(3)}\right] \\
&= -\frac{1}{4\pi^3}Re\left[-G_0^* * \Gamma^{(0)}_{\rm cart} - G_1^* * \Gamma^{(1)}_{\rm cart} + G_2^* * \Gamma^{(2)}_{\rm cart} + G_3^* * \Gamma^{(3)}_{\rm cart}\right], \\
\langle\kappa^{\rm B}\kappa^{\rm E}\kappa^{\rm B}\rangle &= \frac{1}{4}Re\left[-K^{(0)} + K^{(1)} - K^{(2)} + K^{(3)}\right] \\
&= -\frac{1}{4\pi^3}Re\left[-G_0^* * \Gamma^{(0)}_{\rm cart} + G_1^* * \Gamma^{(1)}_{\rm cart} - G_2^* * \Gamma^{(2)}_{\rm cart} + G_3^* * \Gamma^{(3)}_{\rm cart}\right], \\
\langle\kappa^{\rm B}\kappa^{\rm B}\kappa^{\rm E}\rangle &= \frac{1}{4}Re\left[-K^{(0)} + K^{(1)} + K^{(2)} - K^{(3)}\right] \\
&= -\frac{1}{4\pi^3}Re\left[-G_0^* * \Gamma^{(0)}_{\rm cart} + G_1^* * \Gamma^{(1)}_{\rm cart} + G_2^* * \Gamma^{(2)}_{\rm cart} - G_3^* * \Gamma^{(3)}_{\rm cart}\right],
\end{aligned} \tag{4.58}$$

and

$$
\begin{aligned}
\langle\kappa^{\rm B}\kappa^{\rm E}\kappa^{\rm E}\rangle &= \frac{1}{4}Im\left[\, K^{(0)} - K^{(1)} + K^{(2)} + K^{(3)} \,\right] \\
&= -\frac{1}{4\pi^3}Im\left[\, G_0^* * \Gamma^{(0)}_{\rm cart} - G_1^* * \Gamma^{(1)}_{\rm cart} + G_2^* * \Gamma^{(2)}_{\rm cart} + G_3^* * \Gamma^{(3)}_{\rm cart} \,\right], \\
\langle\kappa^{\rm E}\kappa^{\rm B}\kappa^{\rm E}\rangle &= \frac{1}{4}Im\left[\, K^{(0)} + K^{(1)} - K^{(2)} + K^{(3)} \,\right] \\
&= -\frac{1}{4\pi^3}Im\left[\, G_0^* * \Gamma^{(0)}_{\rm cart} + G_1^* * \Gamma^{(1)}_{\rm cart} - G_2^* * \Gamma^{(2)}_{\rm cart} + G_3^* * \Gamma^{(3)}_{\rm cart} \,\right], \\
\langle\kappa^{\rm E}\kappa^{\rm E}\kappa^{\rm B}\rangle &= \frac{1}{4}Im\left[\, K^{(0)} + K^{(1)} + K^{(2)} - K^{(3)} \,\right] \\
&= -\frac{1}{4\pi^3}Im\left[\, G_0^* * \Gamma^{(0)}_{\rm cart} + G_1^* * \Gamma^{(1)}_{\rm cart} + G_2^* * \Gamma^{(2)}_{\rm cart} - G_3^* * \Gamma^{(3)}_{\rm cart} \,\right], \\
\langle\kappa^{\rm B}\kappa^{\rm B}\kappa^{\rm B}\rangle &= \frac{1}{4}Im\left[-K^{(0)} + K^{(1)} + K^{(2)} + K^{(3)}\right] \\
&= -\frac{1}{4\pi^3}Im\left[-G_0^* * \Gamma^{(0)}_{\rm cart} + G_1^* * \Gamma^{(1)}_{\rm cart} + G_2^* * \Gamma^{(2)}_{\rm cart} + G_3^* * \Gamma^{(3)}_{\rm cart}\right],
\end{aligned}
\tag{4.59}
$$

which shows how the E- and B-mode κ3PCFs can be computed when the full information of the $\Gamma_{\rm cart}$'s is available. In the ideal case that there exists no noise or systematical effects, only the E-mode term κ3PCF $\left\langle\kappa^{\rm E}\kappa^{\rm E}\kappa^{\rm E}\right\rangle$ is expected to be non-zero, since it corresponds to the correlation in the physical density field which leads to the correlation in the shear signal.

Following the ideas of Schneider & Kilbinger (2007), we construct a new statistic

$$
{\rm EEE} = \int {\rm d}^2x_1 \int {\rm d}^2x_2 \; \left\langle\kappa^{\rm E}\kappa^{\rm E}\kappa^{\rm E}\right\rangle(\boldsymbol{x_1}, \boldsymbol{x_2})\; U(\boldsymbol{x_1}, \boldsymbol{x_2}) \,, \tag{4.60}
$$

which by definition responds only to E-mode. With the help of (4.58) we can link EEE to the observable $\Gamma_{\rm cart}$'s, as

$$
\begin{aligned}
{\rm EEE} &= -\frac{1}{4\pi^3}\int {\rm d}^2x_1 \int {\rm d}^2x_2 \; U(\boldsymbol{x_1}, \boldsymbol{x_2})\, Re\left[\int {\rm d}^2y_1 \int {\rm d}^2y_2 \; \sum_{i=0}^{3} G_i^*(\boldsymbol{x_1} - \boldsymbol{y_1}, \boldsymbol{x_2} - \boldsymbol{y_2})\, \Gamma^{(i)}_{\rm cart}(\boldsymbol{y_1}, \boldsymbol{y_2})\right] \\
&= -\frac{1}{4\pi^3}Re\left[\int {\rm d}^2y_1 \int {\rm d}^2y_2 \; \sum_{i=0}^{3} \Gamma^{(i)}_{\rm cart}(\boldsymbol{y_1}, \boldsymbol{y_2}) \int {\rm d}^2x_1 \int {\rm d}^2x_2 \; U(\boldsymbol{x_1}, \boldsymbol{x_2})\, G_i^*(\boldsymbol{x_1} - \boldsymbol{y_1}, \boldsymbol{x_2} - \boldsymbol{y_2})\right] \\
&= -\frac{1}{4\pi^3}Re\left[\int {\rm d}^2y_1 \int {\rm d}^2y_2 \; \sum_{i=0}^{3} \Gamma^{(i)}_{\rm cart}(\boldsymbol{y_1}, \boldsymbol{y_2}) \left(G_i^* * U\right)(\boldsymbol{y_1}, \boldsymbol{y_2})\right],
\end{aligned}
\tag{4.61}
$$

where in the first equation we have specified U to be a real function, and in the second equation we have used the fact that $G_i(-\boldsymbol{a}, -\boldsymbol{b}) = G_i(\boldsymbol{a}, \boldsymbol{b})$.

Denoting

$$
G_i^* * U =: T^{(i)} \,, \tag{4.62}
$$

the expression of EEE (4.61) has a similar form as (4.55). We can see in this form that EEE responds only to the E-mode if the weight function T's satisfy

$$
T^{(0)} * G_0 = U = T^{(1)} * G_1 = T^{(2)} * G_2 = T^{(3)} * G_3 \,. \tag{4.63}
$$

with U being a real function. One can easily verify that these conditions are also necessary conditions. Noticing that $T^{(i)} * G_i$ is the corresponding weight on $K^{(i)}$, the condition that these functions being purely real is required to separate the parity-violating and non-violating terms in (4.34). In addition, (4.63) is required to cancel the parity non-violating B-mode terms $\langle\kappa^{\rm E}\kappa^{\rm B}\kappa^{\rm B}\rangle$, $\langle\kappa^{\rm B}\kappa^{\rm E}\kappa^{\rm B}\rangle$ and $\langle\kappa^{\rm B}\kappa^{\rm B}\kappa^{\rm E}\rangle$.

The statistics containing the contribution from only one of the B-mode terms can be constructed in the same way. Omitting the arguments for notational simplicity, they can be expressed as

$$\begin{aligned}
\mathrm{EBB} =& \frac{1}{4} Re\left[\iint \sum_{i=0}^{3} T^{(i)}\,\Gamma^{(i)}_{\rm cart}\right] \text{ with } T^{(0)} * G_0 = T^{(1)} * G_1 = -T^{(2)} * G_2 = -T^{(3)} * G_3\,,\\
\mathrm{BEB} =& \frac{1}{4} Re\left[\iint \sum_{i=0}^{3} T^{(i)}\,\Gamma^{(i)}_{\rm cart}\right] \text{ with } T^{(0)} * G_0 = -T^{(1)} * G_1 = T^{(2)} * G_2 = -T^{(3)} * G_3\,,\\
\mathrm{BBE} =& \frac{1}{4} Re\left[\iint \sum_{i=0}^{3} T^{(i)}\,\Gamma^{(i)}_{\rm cart}\right] \text{ with } T^{(0)} * G_0 = -T^{(1)} * G_1 = -T^{(2)} * G_2 = T^{(3)} * G_3\,,
\end{aligned} \tag{4.64}$$

and

$$\begin{aligned}
\mathrm{BEE} =& \frac{1}{4} Im\left[\iint \sum_{i=0}^{3} T^{(i)}\,\Gamma^{(i)}_{\rm cart}\right] \text{ with } T^{(0)} * G_0 = -T^{(1)} * G_1 = T^{(2)} * G_2 = T^{(3)} * G_3\,,\\
\mathrm{EBE} =& \frac{1}{4} Im\left[\iint \sum_{i=0}^{3} T^{(i)}\,\Gamma^{(i)}_{\rm cart}\right] \text{ with } T^{(0)} * G_0 = T^{(1)} * G_1 = -T^{(2)} * G_2 = T^{(3)} * G_3\,,\\
\mathrm{EEB} =& \frac{1}{4} Im\left[\iint \sum_{i=0}^{3} T^{(i)}\,\Gamma^{(i)}_{\rm cart}\right] \text{ with } T^{(0)} * G_0 = T^{(1)} * G_1 = T^{(2)} * G_2 = -T^{(3)} * G_3\,,\\
\mathrm{BBB} =& \frac{1}{4} Im\left[\iint \sum_{i=0}^{3} T^{(i)}\,\Gamma^{(i)}_{\rm cart}\right] \text{ with } T^{(0)} * G_0 = -T^{(1)} * G_1 = -T^{(2)} * G_2 = -T^{(3)} * G_3\,.
\end{aligned} \tag{4.65}$$

For all of them the condition that $T^{(0)} * G_0$ is real has been imposed.

The four parity violating statistics (4.65) can be used as a check on parity violating systematical errors, while the other B-mode statistics (4.64) allows for a further examination of the B-modes.

With (4.51) one can easily invert the conditions on the weight functions to express the weight function T's directly in terms of each other. Take the conditions for EEE (4.63) for example. One can write the Fourier transforms of $T^{(1)}$, $T^{(2)}$ and $T^{(3)}$ as functions of the Fourier transform of $T^{(0)}$ as

$$\tilde{T}^{(1)} = \frac{1}{\pi^6}\tilde{T}^{(0)}\,\tilde{G}_0\tilde{G}_1^*\,,\quad \tilde{T}^{(2)} = \frac{1}{\pi^6}\tilde{T}^{(0)}\,\tilde{G}_0\tilde{G}_2^*\,,\quad \tilde{T}^{(3)} = \frac{1}{\pi^6}\tilde{T}^{(0)}\,\tilde{G}_0\tilde{G}_3^*\,. \tag{4.66}$$

In order to simplify these relations, we now attempt to give simple expressions for $\tilde{G}_0\tilde{G}_i^*$ for $i = 1, 2, 3$. With (4.50) one has $\tilde{G}_0(\boldsymbol{\ell_1}, \boldsymbol{\ell_2})\tilde{G}_1^*(\boldsymbol{\ell_1}, \boldsymbol{\ell_2}) = \pi^6\,\mathrm{e}^{4\mathrm{i}\beta_1}$, which does not depend on $\boldsymbol{\ell_2}$. Noticing that $\tilde{F} = \pi^2\,\mathrm{e}^{4\mathrm{i}\beta}$ (see Sect. 4.2.3), we actually have $\tilde{G}_0(\boldsymbol{\ell_1}, \boldsymbol{\ell_2})\tilde{G}_1^*(\boldsymbol{\ell_1}, \boldsymbol{\ell_2}) = \pi^4\tilde{F}(\boldsymbol{\ell_1})$, which yields in configuration space

$$\left(G_0 * G_1^*\right)(\boldsymbol{a}, \boldsymbol{b}) = \pi^4 F(\boldsymbol{a})\,\delta_{\rm D}^{(2)}(\boldsymbol{b})\,. \tag{4.67}$$

Similarly one can derive that

$$\left(G_0 * G_2^*\right)(\boldsymbol{a}, \boldsymbol{b}) = \pi^4 F(\boldsymbol{b})\,\delta_{\rm D}^{(2)}(\boldsymbol{a})\,,\quad \left(G_0 * G_3^*\right)(\boldsymbol{a}, \boldsymbol{b}) = \pi^4 F(\boldsymbol{a})\,\delta_{\rm D}^{(2)}(\boldsymbol{b} - \boldsymbol{a})\,. \tag{4.68}$$

Inserting (4.67) and (4.68) into (4.66), one obtains

$$T^{(1)}(\boldsymbol{x_1},\boldsymbol{x_2}) = \frac{1}{\pi^2}\int \mathrm{d}^2y\, T^{(0)}(\boldsymbol{x_1}-\boldsymbol{y},\boldsymbol{x_2})F(\boldsymbol{y}) = \frac{2}{\pi}\int \mathrm{d}^2y\, T^{(0)}(\boldsymbol{x_1}-\boldsymbol{y},\boldsymbol{x_2})\frac{\boldsymbol{y}}{\boldsymbol{y}^{*3}}\,, \tag{4.69}$$

$$T^{(2)}(\boldsymbol{x_1},\boldsymbol{x_2}) = \frac{1}{\pi^2}\int \mathrm{d}^2y\, T^{(0)}(\boldsymbol{x_1},\boldsymbol{x_2}-\boldsymbol{y})F(\boldsymbol{y}) = \frac{2}{\pi}\int \mathrm{d}^2y\, T^{(0)}(\boldsymbol{x_1},\boldsymbol{x_2}-\boldsymbol{y})\frac{\boldsymbol{y}}{\boldsymbol{y}^{*3}}\,, \tag{4.70}$$

$$T^{(3)}(\boldsymbol{x_1},\boldsymbol{x_2}) = \frac{1}{\pi^2}\int \mathrm{d}^2y\, T^{(0)}(\boldsymbol{x_1}-\boldsymbol{y},\boldsymbol{x_2}-\boldsymbol{y})F(\boldsymbol{y}) = \frac{2}{\pi}\int \mathrm{d}^2y\, T^{(0)}(\boldsymbol{x_1}-\boldsymbol{y},\boldsymbol{x_2}-\boldsymbol{y})\frac{\boldsymbol{y}}{\boldsymbol{y}^{*3}}\,, \tag{4.71}$$

where we have used the expression of F (4.19).

To summarize, if one chooses an arbitrary form of $T^{(0)}$ which makes $T^{(0)} * G_0$ real, constructs $T^{(1)}$, $T^{(2)}$ and $T^{(3)}$ according to (4.69), (4.70) and (4.71), and uses these weight functions to weight the measured $\Gamma^{(i)}_{\rm cart}$'s, then the resulting statistic EEE as defined in (4.61) receives contributions only from $\langle\kappa^{\rm E}\kappa^{\rm E}\kappa^{\rm E}\rangle$ but not the terms affected by the B-mode. The B-mode statistics can be obtained from the measured $\Gamma^{(i)}_{\rm cart}$'s through a similar procedure. Equations (4.69)-(4.71) are the analogue of (4.56) for 3-pt functions.

We note again that $\Gamma^{(1)}_{\rm cart}$, $\Gamma^{(2)}_{\rm cart}$ and $\Gamma^{(3)}_{\rm cart}$ are not independent under transformation of their arguments. Thus the statistics EEE (4.61) as well as the B-mode statistics (4.64) and (4.65) can all be written in terms of linear combinations of $\Gamma^{(0)}_{\rm cart}$ and $\Gamma^{(1)}_{\rm cart}$ alone. However we shall keep the current redundancy since it allows for simple analytical expressions of the relations between the weight functions.

So far one still has much freedom in choosing the form of $T^{(0)}(\boldsymbol{x_1},\boldsymbol{x_2})$. This freedom can be exploited to construct statistics which do not respond to the γ3PCF at smaller or larger angular separations than can be probed by the survey. We leave this to future work.

4.6 Numerical evaluation

4.6.1 Design of the sampling grid

We have written configuration space relations between weak lensing statistics in the form of convolutions, e.g. (4.23) and (4.24), where the convolution kernels are complex, have non-trivial spin numbers, and feature singularities. The convolutions can be performed numerically, but special care must be taken of these properties of the integration kernel.

One can take the K-S kernel $-1/\boldsymbol{z}^{*2}$ as an example of this kind of convolution kernel. The K-S kernel has an integer spin of 2, so an azimuthal integration of the kernel around its singularity at $z = 0$ should give zero, i.e. the values of the kernel along the circle cancel themselves due to their opposite phases. This property renders its singularity harmless, but entails the condition that the sampling grid should guarantee the cancellation. Such a requirement of a special grid design has already been realized in early lensing mass reconstruction works (e.g. Seitz & Schneider 1996). For a spin-2 kernel like the K-S kernel, a common square grid already suffices if the singularity is placed at a center of rotational symmetry, i.e. either onto a grid point or at the center of four grid points. In the case of the former, the grid point at the singularity has to be discarded. The latter, as shown by the left panel in Fig. 4.2, is a better choice considering the sampling homogeneity.

In general we need to deal with convolution kernels with different spin numbers. For example, the kernel F between γ2PCF and κ2PCF (4.19), which is proportional to $\boldsymbol{z}/\boldsymbol{z}^{*3}$, has a spin of 4. In this case the square grid cannot guarantee the phase cancellation around the singularity any more (see Fig. 4.3). With a similar analysis as shown in Fig. 4.3, one can see that a triangular grid (middle

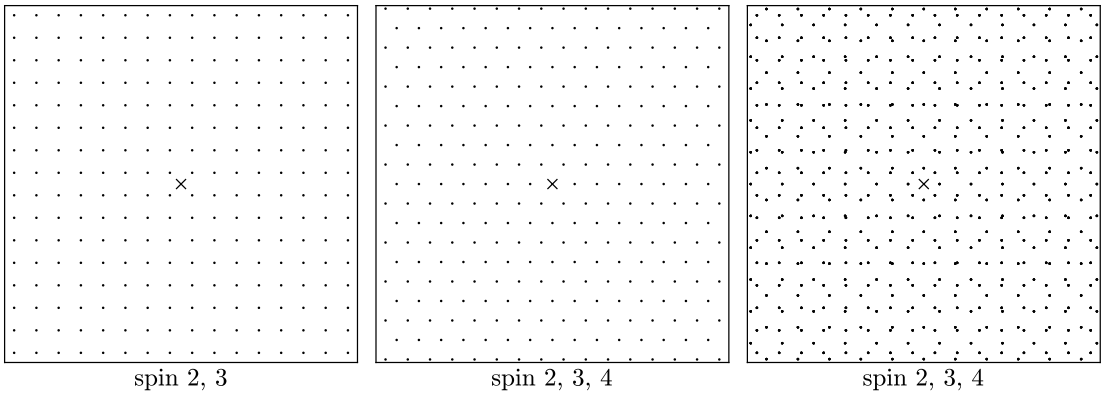


Figure 4.2: Examples of sampling grids applicable for 2D integration over singular kernels of different spin values. The cross in the center indicates the position of the singularity in the integration kernel.

panel, Fig. 4.2) can actually guarantee the phase cancellation around the singularity of a spin-4 kernel. When using a triangular grid, the singularity can also be put either on a grid point or at the center of three grid points. The former choice loses the grid point at the singularity, but is applicable to spin-3 kernels where the latter fails.

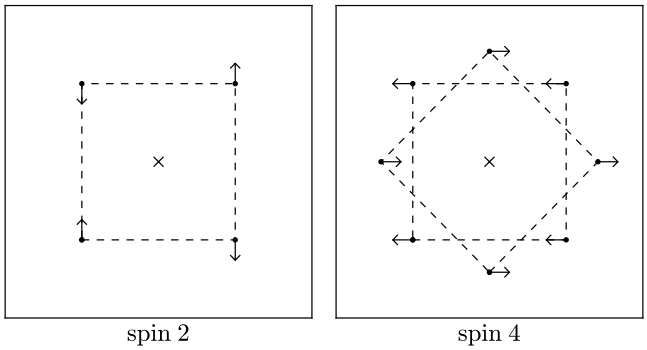


Figure 4.3: A square grid guarantees phase cancellation around the singularity of spin-2 kernels (left panel), but not that of spin-4 kernels (right panel). The crosses indicate the position of the singularity of the integration kernel, while the dots are the grid points closest to the singularity. The polar angles of the kernel at the grid points are indicated by the directions of the arrows. For a spin-2 kernel they cancel each other on a square grid already. For a spin-4 kernel they cancel each other if the square grid is duplicated, rotated 45 degrees and put on top of the original grid.

To achieve a high numerical accuracy, it is required that the circle integral around the singularity is well-sampled. If one uses a square (triangular) grid, the innermost circle is only sampled by four (six) grid points, which is not enough for many κ-models. To remedy this, one can duplicate the sampling grid, rotate it around the singularity, and put it on top of the original grid. We show the result with the square grid and 45 degrees of rotation in the right panel of Fig. 4.2. The resulting grid is also applicable for spin-4 kernels, as shown in the right panel of Fig. 4.3. This non-standard way of constructing sampling grids, although not creating the best grids in terms of sampling efficiency, deals very well with the singularity of the integration kernel, and can easily generate sampling grids

applicable for kernels of any spin number. We use this kind of grid in our numerical sampling.

Additional complications arise when performing the integration for 3-pt statistics. There are two 2D integrals in (4.4), (4.35), (4.36) and (4.37). The corresponding integration kernels (4.22), (4.44), (4.45) and (4.46) all have three singularities. Luckily we can split each integration kernel into four additive terms and perform the integrals over each of them separately. Moreover, one can apply translational shifts to the integrands so that in each 2D integral there is only one singularity in the integration kernel. The singularity can also be shifted to the origin of the grids for numerical simplicity. After all these procedures, the four relations can be written as

$$\Gamma^{(0)}_{\rm cart}(\boldsymbol{x_1},\boldsymbol{x_2}) = \frac{2}{\pi^2}\int \mathrm{d}^2y_1 \int \mathrm{d}^2y_2 \left\{ -\left[\langle\kappa\kappa\kappa\rangle(\boldsymbol{y_2}+\boldsymbol{x_1},\boldsymbol{y_1}+\boldsymbol{y_2}+\boldsymbol{x_2}) + \langle\kappa\kappa\kappa\rangle(\boldsymbol{y_1}+\boldsymbol{x_1}+\boldsymbol{y_2},\boldsymbol{y_2}+\boldsymbol{x_2})\right] \right.$$
$$\left. \times \frac{1}{\boldsymbol{y_1}^{*2}}\,\frac{\boldsymbol{y_2}}{\boldsymbol{y_2}^{*3}} + \left[\langle\kappa\kappa\kappa\rangle(\boldsymbol{y_2}+\boldsymbol{x_1},\boldsymbol{y_1}+\boldsymbol{y_2}+\boldsymbol{x_2}) + \langle\kappa\kappa\kappa\rangle(\boldsymbol{y_1}+\boldsymbol{x_1}+\boldsymbol{y_2},\boldsymbol{y_2}+\boldsymbol{x_2})\right]\,\frac{1}{\boldsymbol{y_1}^{*3}}\,\frac{\boldsymbol{y_2}}{\boldsymbol{y_2}^{*2}} \right\}, \tag{4.72}$$

$$\Gamma^{(1)}_{\rm cart}(\boldsymbol{x_1},\boldsymbol{x_2}) = -\frac{1}{\pi}\int \mathrm{d}^2y\ \left(\langle\kappa^*\kappa\kappa\rangle(\boldsymbol{x_1},\boldsymbol{y}+\boldsymbol{x_2}) + \langle\kappa^*\kappa\kappa\rangle(\boldsymbol{y}+\boldsymbol{x_1},\boldsymbol{y}+\boldsymbol{x_2})\right)\frac{1}{\boldsymbol{y}^{*2}}$$
$$+\frac{2}{\pi^2}\int \mathrm{d}^2y_1 \int \mathrm{d}^2y_2\ \left(\langle\kappa^*\kappa\kappa\rangle(\boldsymbol{y_2}+\boldsymbol{x_1},\boldsymbol{y_1}+\boldsymbol{x_2}) - \langle\kappa^*\kappa\kappa\rangle(\boldsymbol{y_1}+\boldsymbol{y_2}+\boldsymbol{x_1},\boldsymbol{y_1}+\boldsymbol{x_2})\right)\frac{1}{\boldsymbol{y_1}^{*3}}\frac{1}{\boldsymbol{y_2}}, \tag{4.73}$$

$$\Gamma^{(2)}_{\rm cart}(\boldsymbol{x_1},\boldsymbol{x_2}) = -\frac{1}{\pi}\int \mathrm{d}^2y\ \left(\langle\kappa\kappa^*\kappa\rangle(\boldsymbol{y}+\boldsymbol{x_1},\boldsymbol{y}+\boldsymbol{x_2}) + \langle\kappa\kappa^*\kappa\rangle(\boldsymbol{y}+\boldsymbol{x_1},\boldsymbol{x_2})\right)\frac{1}{\boldsymbol{y}^{*2}}$$
$$-\frac{2}{\pi^2}\int \mathrm{d}^2y_1 \int \mathrm{d}^2y_2\ \left(\langle\kappa\kappa^*\kappa\rangle(\boldsymbol{y_1}+\boldsymbol{x_1},\boldsymbol{y_1}+\boldsymbol{y_2}+\boldsymbol{x_2}) - \langle\kappa\kappa^*\kappa\rangle(\boldsymbol{y_1}+\boldsymbol{x_1},\boldsymbol{y_2}+\boldsymbol{x_2})\right)\frac{1}{\boldsymbol{y_1}^{*3}}\frac{1}{\boldsymbol{y_2}}, \tag{4.74}$$

$$\Gamma^{(3)}_{\rm cart}(\boldsymbol{x_1},\boldsymbol{x_2}) = -\frac{1}{\pi}\int \mathrm{d}^2y\ \left(\langle\kappa\kappa\kappa^*\rangle(\boldsymbol{x_1},\boldsymbol{y}+\boldsymbol{x_2}) + \langle\kappa\kappa\kappa^*\rangle(\boldsymbol{y}+\boldsymbol{x_1},\boldsymbol{x_2})\right)\frac{1}{\boldsymbol{y}^{*2}}$$
$$-\frac{2}{\pi^2}\int \mathrm{d}^2y_1 \int \mathrm{d}^2y_2\ \left(\langle\kappa\kappa\kappa^*\rangle(\boldsymbol{x_1}+\boldsymbol{y_2},\boldsymbol{y_1}+\boldsymbol{y_2}+\boldsymbol{x_2}) + \langle\kappa\kappa\kappa^*\rangle(\boldsymbol{y_1}+\boldsymbol{y_2}+\boldsymbol{x_1},\boldsymbol{x_2}+\boldsymbol{y_2})\right)\frac{1}{\boldsymbol{y_1}^{*3}}\frac{1}{\boldsymbol{y_2}}. \tag{4.75}$$

We can see that the integration kernels are either spin-2, spin-4, or spin-3. All sampling grids shown in Fig. 4.2 are applicable to spin-3 kernels.

4.6.2 Numerical results for two-point functions

We now construct several toy models for 2-pt and 3-pt convergence and shear correlations, and use them to test the relations derived as well as the numerical sampling method.

In the 2-pt case, we build two models for the convergence correlation function: $\langle\kappa\kappa\rangle(\boldsymbol{r}) = 1/r$, and $\langle\kappa\kappa\rangle(\boldsymbol{r}) = \mathrm{e}^{-r^2}$. Using the well-established $\xi_+ - \xi_-$ relation (4.1) we can obtain the corresponding models for the shear correlation function: $\langle\gamma\gamma\rangle(\boldsymbol{r}) = \mathrm{e}^{4\mathrm{i}\phi_r}/r$ for $\langle\kappa\kappa\rangle(\boldsymbol{r}) = 1/r$, and $\langle\gamma\gamma\rangle(\boldsymbol{r}) = \mathrm{e}^{4\mathrm{i}\phi_r}\left|(r^4+4r^2+6)\,\mathrm{e}^{-r^2}+2r^2-6\right|/r^4$ for $\langle\kappa\kappa\rangle(\boldsymbol{r}) = \mathrm{e}^{-r^2}$.

Fig. 4.4 shows the comparison between these shear correlation function models and the shear correlation functions sampled using (4.23), with the corresponding convergence correlation function models as input. The numerically sampled values match the analytical models very well. Fig. 4.4 shows the comparison between these shear correlation function models and the shear correlation functions sampled using (4.23), with the corresponding convergence correlation function models as input. The numerically sampled values match the analytical models very well.

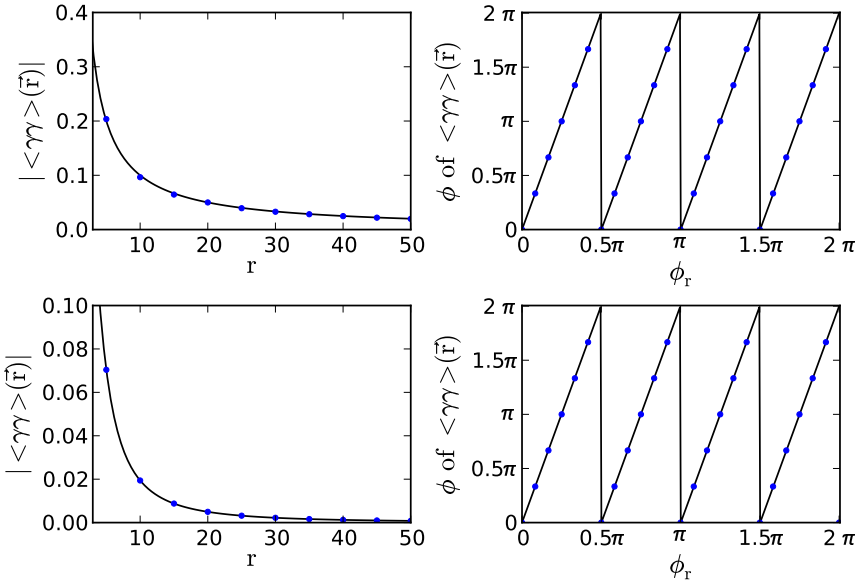


Figure 4.4: *Left panels:* $\langle\gamma\gamma\rangle(\boldsymbol{r})$ with fixed $\phi_r = \pi/6$ as a function of r. *Right panels:* Polar angle of $\langle\gamma\gamma\rangle(\boldsymbol{r})$ with fixed $r = 20$ as a function of ϕ_r. The black curves are the expected values while the dots are the results from numerical evaluation of 2D integral in (4.23). The $\langle\kappa\kappa\rangle$ used in the upper panels is $\langle\kappa\kappa\rangle(\boldsymbol{r}) = 1/r$ and in the lower panels $\langle\kappa\kappa\rangle(\boldsymbol{r}) = \exp(-r^2)$.

4.6.3 Numerical results for three-point functions

We build models for 3pt shear and convergence correlation functions via the 3pt correlation function of the deflection potential ψ. Suppose we evaluate the fields at positions X, Y and Z. By definition we have

$$\langle\kappa(\boldsymbol{X})\kappa(\boldsymbol{Y})\kappa(\boldsymbol{Z})\rangle = \left(\frac{1}{2}\nabla_X^2\right)\left(\frac{1}{2}\nabla_Y^2\right)\left(\frac{1}{2}\nabla_Z^2\right)\,\langle\psi(\boldsymbol{X})\psi(\boldsymbol{Y})\psi(\boldsymbol{Z})\rangle\,, \tag{4.76}$$

$$\langle\gamma(\boldsymbol{X})\gamma(\boldsymbol{Y})\gamma(\boldsymbol{Z})\rangle = \left(\frac{1}{2}\partial_X^2\right)\left(\frac{1}{2}\partial_Y^2\right)\left(\frac{1}{2}\partial_Z^2\right)\,\langle\psi(\boldsymbol{X})\psi(\boldsymbol{Y})\psi(\boldsymbol{Z})\rangle\,, \tag{4.77}$$

and

$$\langle\gamma(\boldsymbol{X})\gamma(\boldsymbol{Y})\gamma^*(\boldsymbol{Z})\rangle = \left(\frac{1}{2}\partial_X^2\right)\left(\frac{1}{2}\partial_Y^2\right)\left(\frac{1}{2}\partial_Z^{*2}\right)\,\langle\psi(\boldsymbol{X})\psi(\boldsymbol{Y})\psi(\boldsymbol{Z})\rangle\,. \tag{4.78}$$

Now we assume a model for $\langle\psi(\boldsymbol{X})\psi(\boldsymbol{Y})\psi(\boldsymbol{Z})\rangle$ as

$$\langle\psi(\boldsymbol{X})\psi(\boldsymbol{Y})\psi(\boldsymbol{Z})\rangle = \frac{1}{8\alpha^6}\,\mathrm{e}^{-\alpha(x^2+y^2)}\,, \tag{4.79}$$

with $\boldsymbol{x} = \boldsymbol{X} - \boldsymbol{Z}$ and $\boldsymbol{y} = \boldsymbol{Y} - \boldsymbol{Z}$. This model is special in the sense that it does not depend on the angle between $\boldsymbol{x}$ and $\boldsymbol{y}$, but is nevertheless simple and rather local. The statistical homogeneity of the field enables us to write the 3pt correlation function as a function of two 2D spatial coordinates, which we have chosen here to be $\boldsymbol{x}$ and $\boldsymbol{y}$.

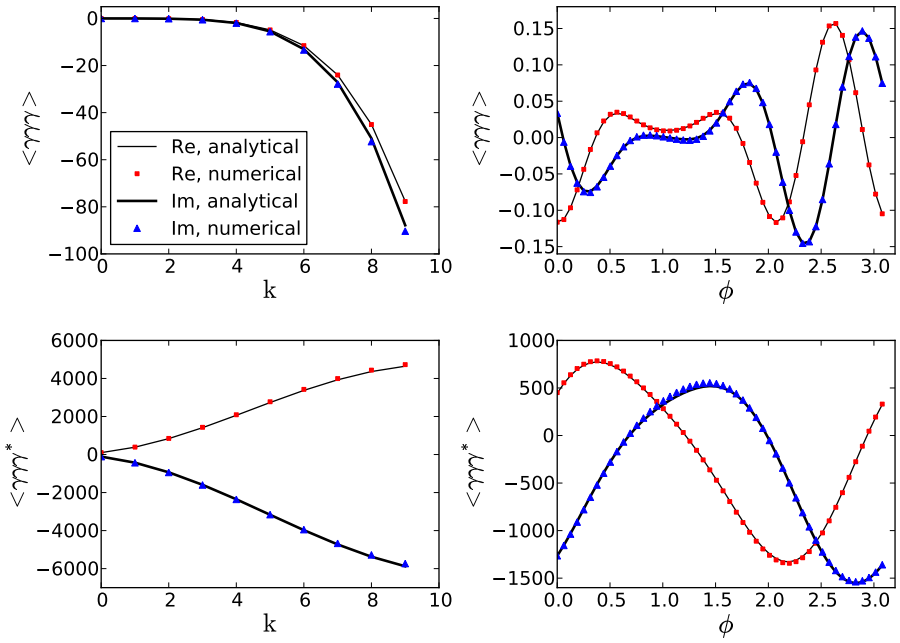


Figure 4.5: Comparison of numerically evaluated $\langle\gamma\gamma\gamma\rangle(\boldsymbol{x},\boldsymbol{y})$ and $\langle\gamma\gamma\gamma^*\rangle(\boldsymbol{x},\boldsymbol{y})$ with their analytical toy models described in this section adopting $\alpha = 0.05$. 'Re' and 'Im' indicate the real and imaginary parts of the shear correlation functions. *Left panels:* The functions are evaluated at $\boldsymbol{x} = 0.25k\ \mathrm{e}^{-\mathrm{i}\pi/3}$ and $\boldsymbol{y} = 0.17k\ \mathrm{e}^{\mathrm{i}\pi/8}$ for different values of k; *Right panels:* The functions are evaluated at $\boldsymbol{x} = 0.75\ \mathrm{e}^{-\mathrm{i}\pi/3}$ and $\boldsymbol{y} = 0.45\ \mathrm{e}^{\mathrm{i}\phi}$ for 50 equally spaced ϕ values.

Performing the derivatives, we find the corresponding 3pt shear and convergence correlation functions also depend only on $\boldsymbol{X} - \boldsymbol{Z} = \boldsymbol{x}$ and $\boldsymbol{Y} - \boldsymbol{Z} = \boldsymbol{y}$, and read

$$\langle\kappa\kappa\kappa\rangle(\boldsymbol{x},\boldsymbol{y}) = \frac{\mathrm{e}^{-\alpha(x^2+y^2)}}{\alpha^3} \times\Big[-4+\alpha\Big(6(x^2+y^2)+8\boldsymbol{x}\cdot\boldsymbol{y}\Big)-\alpha^2\Big((x^2+y^2)^2+4(x^2+y^2)\boldsymbol{x}\cdot\boldsymbol{y}+6x^2y^2\Big)+\alpha^3x^2y^2|\boldsymbol{x}+\boldsymbol{y}|^2\Big], \tag{4.80}$$

$$\langle\gamma\gamma\gamma\rangle(\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{x}^2\boldsymbol{y}^2(\boldsymbol{x}+\boldsymbol{y})^2\ \mathrm{e}^{-\alpha(x^2+y^2)}, \tag{4.81}$$

and

$$\langle\gamma\gamma\gamma^*\rangle(\boldsymbol{x},\boldsymbol{y}) = \Big[2\boldsymbol{x}^2+8\boldsymbol{x}\boldsymbol{y}+2\boldsymbol{y}^2-4\alpha\boldsymbol{x}\boldsymbol{y}\,|\boldsymbol{x}+\boldsymbol{y}|^2+\alpha^2\boldsymbol{x}^2\boldsymbol{y}^2(\boldsymbol{x}^*+\boldsymbol{y}^*)^2\Big]\ \mathrm{e}^{-\alpha(x^2+y^2)}. \tag{4.82}$$

Using (4.80) as the input model for $\langle\kappa\kappa\kappa\rangle$, we numerically evaluate $\langle\gamma\gamma\gamma\rangle(\boldsymbol{x},\boldsymbol{y})$ and $\langle\gamma\gamma\gamma^*\rangle(\boldsymbol{x},\boldsymbol{y})$ using (4.72) and (4.75). The results are then compared to the analytical models for the 3pt shear correlation functions (4.81) and (4.82). As shown in Fig. 4.5, the numerical evaluations closely match the analytical models.

Hence, we have proven numerically that the relations between γ3PCFs and κ3PCFs (4.72) and (4.75) are correct, no additional delta functions are needed. At the same time we have shown that these relations can be numerically evaluated with a high accuracy. Therefore these relations provide

a better way of relating the observable shear signal to the underlying matter density field than the original way, i.e. using the relations between the γ3PCFs and the κ bispectrum, since the latter does not allow for an easy accurate numerical evaluation.

4.7 Conclusion

We derived the relations between the 3-pt shear and convergence correlation functions which had been an important missing link between weak lensing three-point statistics. As an intermediate step, we found the 2-pt analogue of these relations and proved that it is the non-symmetrized form of the existing $\xi_+ - \xi_-$ relation. By drawing analogy to the corresponding 1-pt relations, namely the Kaiser-Squires relation and its isotropic form, we have further revealed that the newly derived relations and already established results fit into the same theoretical framework. The consistency of the configuration space relations with the known Fourier space relations have also been shown.

The 3-pt relations derived are simple both analytically and numerically. They can be used as an alternative way of relating the measurable 3-pt weak lensing statistics with the statistics of the underlying matter density field. Up to now one has to use the relations between the γ3PCFs and the convergence bispectrum to link theory to the observable 3-pt shear signal. Since the γ3PCFs are very oscillatory and complicated functions of the convergence bispectrum (Schneider et al. 2005), it is hard to study the behavior of the 3-pt shear signal for a given convergence bispectrum model. With the relations we derived, one can instead study the properties of the 3-pt shear signal by constructing models for the κ3PCFs.

The method we used to derive these relations is based on the relation between the 2-pt correlation functions of the lensing deflection potential and the convergence. The same method also allows one to systematically derive the relations between correlations functions of other weak lensing quantities, including the deflection potential ψ, the shear γ, the convergence κ, and the deflection angle α. We present the forms of some 2- and 3-pt relations in Appendix. 4.9. Some of them are potentially of interest for studies of galaxy-galaxy(-galaxy) lensing and lensing of the Cosmic Microwave Background.

Since the relations we obtained have complex kernels with non-trivial spin number and singularities, special care is needed when they are used numerically. We demonstrated how the numerical evaluation can be done, in particular the design of the sampling grid. Examples of numerical evaluation were shown for both 2- and 3-pt relations using toy models for the convergence correlation function. Their results match very well with the analytical expectations.

Separating E- and B-modes from measurements of the γ3PCFs is particularly important since the systematic effects at the 3-pt level are less understood. So far the only 3-pt statistics allowing for an E/B-mode decomposition is the aperture mass statistics (Jarvis et al. 2004; Schneider et al. 2005) which is plagued with the same problem as the 2-pt aperture statistics pointed out by Kilbinger et al. (2006). To amend this problem, one needs to construct the 3-pt correspondence of the newly developed 2-pt statistics (Schneider & Kilbinger 2007; Eifler et al. 2010; Fu & Kilbinger 2010; Schneider et al. 2010) which allows for an E/B-mode decomposition on a finite interval. As a direct theoretical application of the 3-pt relations derived in this study, we used them to formulate the conditions for E/B-mode decomposition of lensing 3-pt statistics, in analogy to the 2-pt condition given by Schneider & Kilbinger (2007). These conditions are the basis of formulating additional constraints which lead to E/B-mode decomposition over a finite region, therefore they provide a starting point for future works on constructing better 3-pt shear statistics.

Our work was done for the case of weak lensing, but since it has only used the mathematical

structure of the shear and the convergence, it applies also to other 2D polarization fields such as the polarization fluctuations of the Cosmic Microwave Background.

4.8 Appendix: Fourier transform of the F and G_0 kernels

The F and G_0 kernels, as defined in (4.7) and (4.4) relate the 2- and 3-pt shear and convergence correlation functions, respectively. In Sect. 2 we have derived their explicit forms, see (4.19) and (4.22). According to the Fourier space relations of the shear and convergence statistics (4.30) and (4.31), one expects that F/π^2 and $\mathrm{e}^{4\mathrm{i}\beta}$ are Fourier pairs, as well as $-G_0/\pi^3$ and $\mathrm{e}^{2\mathrm{i}(\beta_1+\beta_2+\beta_3)}$. Here we perform the Fourier transforms of F/π^2 and $-G_0/\pi^3$, with F and G_0 given in (4.19) and (4.22).

The Fourier transformation of the kernel F is a 2D integral of the form

$$\frac{\tilde{F}(\boldsymbol{\ell})}{\pi^2} = \int \mathrm{d}^2a\ \mathrm{e}^{-\mathrm{i}\boldsymbol{\ell}\cdot\boldsymbol{a}}\ \frac{F(\boldsymbol{a})}{\pi^2} = \frac{2}{\pi}\int \mathrm{d}^2a\ \mathrm{e}^{-\mathrm{i}\boldsymbol{\ell}\cdot\boldsymbol{a}}\ \frac{\boldsymbol{a}}{\boldsymbol{a}^{*3}}\,. \tag{4.83}$$

The Fourier transformation of the kernel G_0 can be greatly simplified by performing translational shifts to the integration variables. It turns out that the full transformation is composed of 2D integrals similar to that in (4.83),

$$\begin{aligned}
-\frac{\tilde{G}_0(\boldsymbol{\ell_1},\boldsymbol{\ell_2})}{\pi^3} &= \int \mathrm{d}^2a\ \int \mathrm{d}^2b\ \mathrm{e}^{-\mathrm{i}(\boldsymbol{\ell_1}\cdot\boldsymbol{a}+\boldsymbol{\ell_2}\cdot\boldsymbol{b})}\left(-\frac{G_0(\boldsymbol{a},\boldsymbol{b})}{\pi^3}\right)\\
&= -\frac{2}{\pi^2}\int \mathrm{d}^2a\ \mathrm{e}^{-\mathrm{i}\boldsymbol{\ell_1}\cdot\boldsymbol{a}}\ \int \mathrm{d}^2b\ \mathrm{e}^{-\mathrm{i}\boldsymbol{\ell_2}\cdot\boldsymbol{b}}\left[\frac{1}{(\boldsymbol{a}^*-\boldsymbol{b}^*)^2}\left(\frac{\boldsymbol{a}}{\boldsymbol{a}^{*3}}+\frac{\boldsymbol{b}}{\boldsymbol{b}^{*3}}\right)+\frac{1}{(\boldsymbol{a}^*-\boldsymbol{b}^*)^3}\left(\frac{\boldsymbol{a}}{\boldsymbol{a}^{*2}}-\frac{\boldsymbol{b}}{\boldsymbol{b}^{*2}}\right)\right]\\
&= -\frac{2}{\pi^2}\int \mathrm{d}^2a\ \mathrm{e}^{-\mathrm{i}\boldsymbol{\ell_1}\cdot\boldsymbol{a}}\frac{\boldsymbol{a}}{\boldsymbol{a}^{*3}}\ \int \mathrm{d}^2b\ \mathrm{e}^{-\mathrm{i}\boldsymbol{\ell_2}\cdot\boldsymbol{b}}\frac{1}{(\boldsymbol{a}^*-\boldsymbol{b}^*)^2}-\frac{2}{\pi^2}\int \mathrm{d}^2b\ \mathrm{e}^{-\mathrm{i}\boldsymbol{\ell_2}\cdot\boldsymbol{b}}\frac{\boldsymbol{b}}{\boldsymbol{b}^{*3}}\ \int \mathrm{d}^2a\ \mathrm{e}^{-\mathrm{i}\boldsymbol{\ell_1}\cdot\boldsymbol{a}}\frac{1}{(\boldsymbol{a}^*-\boldsymbol{b}^*)^2}\\
&\quad -\frac{2}{\pi^2}\int \mathrm{d}^2a\ \mathrm{e}^{-\mathrm{i}\boldsymbol{\ell_1}\cdot\boldsymbol{a}}\frac{\boldsymbol{a}}{\boldsymbol{a}^{*2}}\ \int \mathrm{d}^2b\ \mathrm{e}^{-\mathrm{i}\boldsymbol{\ell_2}\cdot\boldsymbol{b}}\frac{1}{(\boldsymbol{a}^*-\boldsymbol{b}^*)^3}+\frac{2}{\pi^2}\int \mathrm{d}^2b\ \mathrm{e}^{-\mathrm{i}\boldsymbol{\ell_2}\cdot\boldsymbol{b}}\frac{\boldsymbol{b}}{\boldsymbol{b}^{*2}}\ \int \mathrm{d}^2a\ \mathrm{e}^{-\mathrm{i}\boldsymbol{\ell_1}\cdot\boldsymbol{a}}\frac{1}{(\boldsymbol{a}^*-\boldsymbol{b}^*)^3}\\
&= -\frac{2}{\pi^2}\int \mathrm{d}^2a\ \mathrm{e}^{-\mathrm{i}(\boldsymbol{\ell_1}+\boldsymbol{\ell_2})\cdot\boldsymbol{a}}\frac{\boldsymbol{a}}{\boldsymbol{a}^{*3}}\ \int \mathrm{d}^2b\ \mathrm{e}^{-\mathrm{i}\boldsymbol{\ell_2}\cdot\boldsymbol{b}}\frac{1}{\boldsymbol{b}^{*2}}-\frac{2}{\pi^2}\int \mathrm{d}^2a\ \mathrm{e}^{-\mathrm{i}(\boldsymbol{\ell_1}+\boldsymbol{\ell_2})\cdot\boldsymbol{a}}\frac{\boldsymbol{a}}{\boldsymbol{a}^{*3}}\ \int \mathrm{d}^2b\ \mathrm{e}^{-\mathrm{i}\boldsymbol{\ell_1}\cdot\boldsymbol{b}}\frac{1}{\boldsymbol{b}^{*2}}\\
&\quad +\frac{2}{\pi^2}\int \mathrm{d}^2a\ \mathrm{e}^{-\mathrm{i}(\boldsymbol{\ell_1}+\boldsymbol{\ell_2})\cdot\boldsymbol{a}}\frac{\boldsymbol{a}}{\boldsymbol{a}^{*2}}\ \int \mathrm{d}^2b\ \mathrm{e}^{-\mathrm{i}\boldsymbol{\ell_2}\cdot\boldsymbol{b}}\frac{1}{\boldsymbol{b}^{*3}}+\frac{2}{\pi^2}\int \mathrm{d}^2a\ \mathrm{e}^{-\mathrm{i}(\boldsymbol{\ell_1}+\boldsymbol{\ell_2})\cdot\boldsymbol{a}}\frac{\boldsymbol{a}}{\boldsymbol{a}^{*2}}\ \int \mathrm{d}^2b\ \mathrm{e}^{-\mathrm{i}\boldsymbol{\ell_1}\cdot\boldsymbol{b}}\frac{1}{\boldsymbol{b}^{*3}}\,.
\end{aligned} \tag{4.84}$$

Performing the 2D integrals in polar coordinates results in

$$\int \mathrm{d}^2a\ \mathrm{e}^{-\mathrm{i}\boldsymbol{\ell}\cdot\boldsymbol{a}}\frac{\boldsymbol{a}}{\boldsymbol{a}^{*3}} = \int_0^\infty \frac{\mathrm{d}a}{a}\ \int_0^{2\pi}\mathrm{d}\phi_a\ \mathrm{e}^{4\mathrm{i}\phi_a}\ \mathrm{e}^{-\mathrm{i}\ell a\cos(\phi_a-\beta)} = 2\pi\ \mathrm{e}^{4\mathrm{i}\beta}\int_0^\infty\frac{\mathrm{d}a}{a}J_4(\ell a) = \frac{\pi}{2}\ \mathrm{e}^{4\mathrm{i}\beta}\,, \tag{4.85}$$

$$\int \mathrm{d}^2a\ \mathrm{e}^{-\mathrm{i}\boldsymbol{\ell}\cdot\boldsymbol{a}}\frac{\boldsymbol{a}}{\boldsymbol{a}^{*2}} = \int_0^\infty \mathrm{d}a\ \int_0^{2\pi}\mathrm{d}\phi_a\ \mathrm{e}^{3\mathrm{i}\phi_a}\ \mathrm{e}^{-\mathrm{i}\ell a\cos(\phi_a-\beta)} = 2\pi\mathrm{i}\ \mathrm{e}^{3\mathrm{i}\beta}\int_0^\infty \mathrm{d}a\ J_3(\ell a) = 2\pi\mathrm{i}\frac{\mathrm{e}^{3\mathrm{i}\beta}}{\ell}\,, \tag{4.86}$$

$$\int \mathrm{d}^2a\ \mathrm{e}^{-\mathrm{i}\boldsymbol{\ell}\cdot\boldsymbol{a}}\frac{1}{\boldsymbol{a}^{*3}} = \int_0^\infty \frac{\mathrm{d}a}{a^2}\ \int_0^{2\pi}\mathrm{d}\phi_a\ \mathrm{e}^{3\mathrm{i}\phi_a}\ \mathrm{e}^{-\mathrm{i}\ell a\cos(\phi_a-\beta)} = 2\pi\mathrm{i}\ \mathrm{e}^{3\mathrm{i}\beta}\int_0^\infty\frac{\mathrm{d}a}{a^2}J_3(\ell a) = \frac{\mathrm{i}\pi\ell}{4}\ \mathrm{e}^{3\mathrm{i}\beta}\,, \tag{4.87}$$

and

$$\int \mathrm{d}^2a\, \mathrm{e}^{-\mathrm{i}\boldsymbol{\ell}\cdot\boldsymbol{a}} \frac{1}{\boldsymbol{a}^{*2}} = \int_0^\infty \frac{\mathrm{d}a}{a} \int_0^{2\pi} \mathrm{d}\phi_a\, \mathrm{e}^{2\mathrm{i}\phi_a}\, \mathrm{e}^{-\mathrm{i}\ell a\cos(\phi_a-\beta)} = -2\pi\, \mathrm{e}^{2\mathrm{i}\beta} \int_0^\infty \frac{\mathrm{d}a}{a} J_2(\ell a) = -\pi\, \mathrm{e}^{2\mathrm{i}\beta}\,. \tag{4.88}$$

Combining (4.83) and (4.85) yields

$$\frac{\tilde{F}(\boldsymbol{\ell})}{\pi^2} = \frac{2}{\pi}\left(\frac{\pi}{2}\,\mathrm{e}^{4\mathrm{i}\beta}\right) = \mathrm{e}^{4\mathrm{i}\beta}\,, \tag{4.89}$$

which demonstrates that the Fourier transformation of F/π^2 is indeed the phase factor $\mathrm{e}^{4\mathrm{i}\beta}$ in (4.30).

For the G_0 kernel we still need to take account that $\boldsymbol{\ell_3} = \ell_3\,\mathrm{e}^{\mathrm{i}\beta_3} \equiv -\boldsymbol{\ell_1} - \boldsymbol{\ell_2}$, so that

$$\begin{aligned}
-\frac{\tilde{G}_0(\boldsymbol{\ell_1},\boldsymbol{\ell_2})}{\pi^3} =& -\frac{2}{\pi^2}\left(\frac{\pi}{2}\,\mathrm{e}^{4\mathrm{i}\beta_3}\right)\left(-\pi\,\mathrm{e}^{2\mathrm{i}\beta_2}\right) - \frac{2}{\pi^2}\left(2\pi\mathrm{i}\frac{\mathrm{e}^{3\mathrm{i}\beta_3}}{\ell_3}\right)\left(\frac{\mathrm{i}\pi\ell_2}{4}\,\mathrm{e}^{3\mathrm{i}\beta_2}\right)\\
&-\frac{2}{\pi^2}\left(\frac{\pi}{2}\,\mathrm{e}^{4\mathrm{i}\beta_3}\right)\left(-\pi\,\mathrm{e}^{2\mathrm{i}\beta_1}\right) - \frac{2}{\pi^2}\left(2\pi\mathrm{i}\frac{\mathrm{e}^{3\mathrm{i}\beta_3}}{\ell_3}\right)\left(\frac{\mathrm{i}\pi\ell_1}{4}\,\mathrm{e}^{3\mathrm{i}\beta_1}\right)\\
=& \mathrm{e}^{4\mathrm{i}\beta_3+2\mathrm{i}\beta_2} + \frac{\ell_2}{\ell_3}\,\mathrm{e}^{3\mathrm{i}\beta_3+3\mathrm{i}\beta_2} + \mathrm{e}^{2\mathrm{i}\beta_1+4\mathrm{i}\beta_3} + \frac{\ell_1}{\ell_3}\,\mathrm{e}^{3\mathrm{i}\beta_1+3\mathrm{i}\beta_3}\\
=& \frac{\ell_3^2}{\ell_3^{*2}}\frac{\ell_2}{\ell_2^*} + \frac{\ell_2^2}{\ell_2^*}\frac{\ell_3}{\ell_3^{*2}} + \frac{\ell_3^2}{\ell_3^{*2}}\frac{\ell_1}{\ell_1^*} + \frac{\ell_1^2}{\ell_1^*}\frac{\ell_3}{\ell_3^{*2}}\\
=& \frac{\ell_1\ell_3\,(\ell_1+\ell_3)}{\ell_1^*\ell_3^*} + \frac{\ell_2\ell_3\,(\ell_2+\ell_3)}{\ell_2^*\ell_3^*} = -\frac{\ell_1\ell_2\ell_3}{\ell_3^{*2}}\left(\frac{1}{\ell_1^*}+\frac{1}{\ell_2^*}\right)\\
=& \frac{\ell_1\ell_2\ell_3}{\ell_1^*\ell_2^*\ell_3^*} = \mathrm{e}^{2\mathrm{i}(\beta_1+\beta_2+\beta_3)}\,.
\end{aligned} \tag{4.90}$$

Thus the Fourier transformation of $-G_0/\pi^3$ equals the phase factor $\mathrm{e}^{2\mathrm{i}(\beta_1+\beta_2+\beta_3)}$ in (4.31), as expected.

4.9 Appendix: Relations between other correlation functions

In Sect. 4.1.2 we have derived the relations between the shear and the convergence 2- and 3-pt correlation functions. The method we used is based on the relation between the 2-pt correlation functions of the convergence κ and the deflection potential ψ (4.12). Thus the method can easily be generated to derive relations between the correlation function of κ and that of any weak lensing quantity g which can be expressed as derivatives of ψ. We denote $g = \mathrm{D_g}\psi$, and write these 2-pt relations in a general form

$$\langle g(\boldsymbol{x_1})g'(\boldsymbol{x_2})\rangle = \frac{1}{\pi^2}\int \mathrm{d}^2y\ \langle\kappa\kappa\rangle(\boldsymbol{y})\,\mathcal{H}(\boldsymbol{x_1}-\boldsymbol{x_2}-\boldsymbol{y})\,. \tag{4.91}$$

Listed below are some candidates for g and the corresponding operator $\mathrm{D_g}$, where α is the deflection angle defined as $\alpha = \partial\psi$.

g	ψ	κ	γ	α
$\mathrm{D_g}$	1	$\nabla^2/2$	$\partial^2/2$	∂

Table 4.1: Forms of the convolution kernel $\mathcal{H}$ as defined in (4.91) for different weak lensing 2-pt statistics.

$\langle XX\rangle$	$\mathcal{H}(\mathcal{F})$	integration form of $\mathcal{H}$	$\mathcal{H}(\boldsymbol{z})$
$\langle\psi\psi\rangle$	$\mathcal{F}$	$\int \mathrm{d}^2\nu\ \ln\lvert\boldsymbol{\nu}\rvert \ln\lvert\boldsymbol{z}-\boldsymbol{\nu}\rvert$	$(\pi/2)\,\lvert\boldsymbol{z}\rvert^2\,(\ln \boldsymbol{z}-1)$
$\langle\gamma\psi\rangle$	$\frac{1}{2}\partial^2\mathcal{F}$	$-\int \mathrm{d}^2\nu\ \frac{1}{(\boldsymbol{\nu}^*-\boldsymbol{z}^*)^2}\ln\lvert\boldsymbol{\nu}\rvert$	$(\pi/2)\,\boldsymbol{z}/\boldsymbol{z}^*$
$\langle\alpha\alpha\rangle$	$-\partial^2\mathcal{F}$	$\int \mathrm{d}^2\nu\ \frac{1}{\boldsymbol{\nu}^*}\frac{1}{\boldsymbol{\nu}^*-\boldsymbol{z}^*}$	$-\pi\boldsymbol{z}/\boldsymbol{z}^*$
$\langle\kappa\psi\rangle$	$\frac{1}{2}\partial\partial^*\mathcal{F}$	$\pi\int \mathrm{d}^2\nu\ \delta_{\mathrm{D}}^{(2)}(\boldsymbol{z}-\boldsymbol{\nu})\ln\lvert\boldsymbol{\nu}\rvert$	$\pi\ln\boldsymbol{z}$
$\langle\alpha\gamma\rangle$	$\frac{1}{2}\partial^3\mathcal{F}$	$\int \mathrm{d}^2\nu\ \frac{1}{\boldsymbol{\nu}^{*2}}\frac{1}{\boldsymbol{\nu}^*-\boldsymbol{z}^*}$	$-\pi\boldsymbol{z}/\boldsymbol{z}^{*2}$
$\langle\alpha\kappa\rangle$	$\frac{1}{2}\partial^2\partial^*\mathcal{F}$	$\pi\int \mathrm{d}^2\nu\ \frac{1}{\boldsymbol{z}^*-\boldsymbol{\nu}^*}\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\nu})$	$\pi/\boldsymbol{z}^*$
$\langle\alpha^*\gamma\rangle$	$\frac{1}{2}\partial^2\partial^*\mathcal{F}$	$\int \mathrm{d}^2\nu\ \frac{1}{\boldsymbol{\nu}^{*2}}\frac{1}{\boldsymbol{\nu}-\boldsymbol{z}}$	$\pi/\boldsymbol{z}^*$
$\langle\gamma\gamma\rangle$	$\frac{1}{4}\partial^4\mathcal{F}$	$\int \mathrm{d}^2\nu\ \frac{1}{\boldsymbol{\nu}^{*2}}\frac{1}{(\boldsymbol{\nu}^*-\boldsymbol{z}^*)^2}$	$2\pi\boldsymbol{z}/\boldsymbol{z}^{*3}$
$\langle\kappa\gamma\rangle$	$\frac{1}{4}\partial^3\partial^*\mathcal{F}$	$-\pi\int \mathrm{d}^2\nu\ \frac{1}{\boldsymbol{\nu}^{*2}}\delta_{\mathrm{D}}^{(2)}(\boldsymbol{z}-\boldsymbol{\nu})$	$-\pi/\boldsymbol{z}^{*2}$
$\langle\kappa\kappa\rangle$	$\frac{1}{4}\partial^2\partial^{*2}\mathcal{F}$		$\pi^2\delta_{\mathrm{D}}^{(2)}(\boldsymbol{z})$
$\langle\gamma\gamma^*\rangle$	$\frac{1}{4}\partial^2\partial^{*2}\mathcal{F}$	$\int \mathrm{d}^2\nu\ \frac{1}{\boldsymbol{\nu}^2}\frac{1}{(\boldsymbol{\nu}^*-\boldsymbol{z}^*)^2}$	$\pi^2\delta_{\mathrm{D}}^{(2)}(\boldsymbol{z})$

Let $\mathrm{D_g D_{g'}}$ act on both sides of the relation between $\langle\psi\psi\rangle$ and $\langle\kappa\kappa\rangle$ (4.11), and use the statistical homogeneity of the κ field, one can obtain an integration form of the kernel $\mathcal{H}$, in analogy to (4.8). Let the same operator act on both sides of (4.12), one can express $\mathcal{H}$ as derivatives of the convolution kernel $\mathcal{F}$ in (4.12), as $\mathcal{H}(\boldsymbol{x_1}-\boldsymbol{x_2}-\boldsymbol{y}) = \mathrm{D_g D_{g'}}\mathcal{F}(\boldsymbol{x_1}-\boldsymbol{x_2}-\boldsymbol{y})$. Further inserting the explicit form of $\mathcal{F}$ (4.19) allows one to obtain the explicit form of $\mathcal{H}$ as a function of $\boldsymbol{z} = \boldsymbol{x_1}-\boldsymbol{x_2}-\boldsymbol{y}$. We summarize some of the 2-pt relations in Table. 4.1. Note that the form of $\mathcal{H}(\mathcal{F})$ for $\langle\alpha\alpha\rangle$ has a minus sign, which is due to the fact that $\partial_{x_1}\partial_{x_2}\mathcal{F}(\boldsymbol{x_1}-\boldsymbol{x_2}-\boldsymbol{y}) = -\partial^2\mathcal{F}(\boldsymbol{z})$ with $\boldsymbol{z} = \boldsymbol{x_1}-\boldsymbol{x_2}-\boldsymbol{y}$.

We write the relations between $\langle\kappa\kappa\kappa\rangle$ and the 3-pt correlation functions of the g's also in a uniform convolutional form,

$$\langle gg'g''\rangle(\boldsymbol{x_1},\boldsymbol{x_2}) = \frac{1}{\pi^3}\int \mathrm{d}^2y_1\int \mathrm{d}^2y_2\ \langle\kappa\kappa\kappa\rangle(\boldsymbol{y_1},\boldsymbol{y_2})\ I(\boldsymbol{x_1}-\boldsymbol{y_1},\boldsymbol{x_2}-\boldsymbol{y_2})\,. \tag{4.92}$$

To find the explicit form of the convolution kernel I, we first write it into an integral form, in analogy to (4.5), and then split it into terms which can be expressed also as derivatives of the kernel $\mathcal{F}$, like (4.20). Then with the explicit form of $\mathcal{F}$ one can reach the explicit form of I. We list the forms of the convolution kernel I for some 3-pt statistics in Table. 4.2.

Some of these relations, e.g. those for $\langle\kappa\gamma\rangle$, $\langle\gamma\gamma\kappa\rangle$, and $\langle\gamma\kappa\kappa\rangle$, can find their application in galaxy-galaxy(-galaxy) lensing which corresponds to the cross-correlation of shear and galaxy number density. Some other relations, e.g. those for $\langle\alpha\alpha\rangle$, $\langle\alpha\kappa\rangle$, and $\langle\alpha\alpha\kappa\rangle$, are potentially of interest for studies of the lensing effects on the Cosmic Microwave Background and its cross-correlation with galaxy weak-lensing maps (Hu 2000).

Table 4.2: Forms of the convolution kernel I as defined in (4.92) for different weak lensing 3-pt statistics.

$\langle XXX\rangle$	integration form of I	split form of I	$I(\boldsymbol{a},\boldsymbol{b})$
$\langle\alpha\alpha\alpha\rangle$	$-\int \mathrm{d}^2\nu\ \frac{1}{\boldsymbol{\nu}^*}\frac{1}{\boldsymbol{\nu}^*-\boldsymbol{a}^*}\frac{1}{\boldsymbol{\nu}^*-\boldsymbol{b}^*}$	$-\frac{1}{\boldsymbol{a}^*-\boldsymbol{b}^*}\int \mathrm{d}^2\nu\frac{1}{\boldsymbol{\nu}^*}\left(\frac{1}{\boldsymbol{\nu}^*-\boldsymbol{a}^*}-\frac{1}{\boldsymbol{\nu}^*-\boldsymbol{b}^*}\right)$	$\frac{\pi}{\boldsymbol{a}^*-\boldsymbol{b}^*}\left(\frac{\boldsymbol{a}}{\boldsymbol{a}^*}-\frac{\boldsymbol{b}}{\boldsymbol{b}^*}\right)$
$\langle\alpha\alpha\gamma\rangle$	$-\int \mathrm{d}^2\nu\ \frac{1}{\boldsymbol{\nu}^{*2}}\frac{1}{\boldsymbol{\nu}^*-\boldsymbol{a}^*}\frac{1}{\boldsymbol{\nu}^*-\boldsymbol{b}^*}$	$-\frac{1}{\boldsymbol{a}^*-\boldsymbol{b}^*}\int \mathrm{d}^2\nu\frac{1}{\boldsymbol{\nu}^{*2}}\left(\frac{1}{\boldsymbol{\nu}^*-\boldsymbol{a}^*}-\frac{1}{\boldsymbol{\nu}^*-\boldsymbol{b}^*}\right)$	$\frac{\pi}{\boldsymbol{a}^*-\boldsymbol{b}^*}\left(\frac{\boldsymbol{a}}{\boldsymbol{a}^{*2}}-\frac{\boldsymbol{b}}{\boldsymbol{b}^{*2}}\right)$
$\langle\gamma\gamma\alpha\rangle$	$-\int \mathrm{d}^2\nu\ \frac{1}{\boldsymbol{\nu}^*}\frac{1}{(\boldsymbol{\nu}^*-\boldsymbol{a}^*)^2}\frac{1}{(\boldsymbol{\nu}^*-\boldsymbol{b}^*)^2}$	$-\frac{1}{(\boldsymbol{a}^*-\boldsymbol{b}^*)^2}\int \mathrm{d}^2\nu\frac{1}{\boldsymbol{\nu}^*}\left(\frac{1}{(\boldsymbol{\nu}^*-\boldsymbol{a}^*)^2}+\frac{1}{(\boldsymbol{\nu}^*-\boldsymbol{b}^*)^2}\right)$ $+\frac{2}{(\boldsymbol{a}^*-\boldsymbol{b}^*)^3}\int \mathrm{d}^2\nu\frac{1}{\boldsymbol{\nu}^*}\left(\frac{1}{\boldsymbol{\nu}^*-\boldsymbol{a}^*}-\frac{1}{\boldsymbol{\nu}^*-\boldsymbol{b}^*}\right)$	$-\frac{\pi}{(\boldsymbol{a}^*-\boldsymbol{b}^*)^2}\left(\frac{\boldsymbol{a}}{\boldsymbol{a}^{*2}}+\frac{\boldsymbol{b}}{\boldsymbol{b}^{*2}}\right)-\frac{2\pi}{(\boldsymbol{a}^*-\boldsymbol{b}^*)^3}\left(\frac{\boldsymbol{a}}{\boldsymbol{a}^*}-\frac{\boldsymbol{b}}{\boldsymbol{b}^*}\right)$
$\langle\gamma\gamma\kappa\rangle$	$\pi\int \mathrm{d}^2\nu\ \delta_\mathrm{D}^{(2)}(\boldsymbol{\nu})\frac{1}{(\boldsymbol{\nu}^*-\boldsymbol{a}^*)^2}\frac{1}{(\boldsymbol{\nu}^*-\boldsymbol{b}^*)^2}$		$\pi/\left(\boldsymbol{a}^{*2}\boldsymbol{b}^{*2}\right)$
$\langle\gamma\kappa\kappa\rangle$	$\pi^2\int \mathrm{d}^2\nu\ \delta_\mathrm{D}^{(2)}(\boldsymbol{\nu})\ \delta_\mathrm{D}^{(2)}(\boldsymbol{\nu}-\boldsymbol{b})\frac{1}{(\boldsymbol{\nu}^*-\boldsymbol{a}^*)^2}$		$\pi^2\delta_\mathrm{D}^{(2)}(\boldsymbol{b})/\boldsymbol{a}^{*2}$
$\langle\alpha\alpha\kappa\rangle$	$\pi\int \mathrm{d}^2\nu\ \delta_\mathrm{D}^{(2)}(\boldsymbol{\nu})\frac{1}{\boldsymbol{\nu}^*-\boldsymbol{a}^*}\frac{1}{\boldsymbol{\nu}^*-\boldsymbol{b}^*}$		$\pi/\left(\boldsymbol{a}^*\boldsymbol{b}^*\right)$

Chapter 5

Bispectrum covariance in the flat-sky limit

As the matter density field evolves to be more and more non-Gaussian under non-linear gravitational clustering, information which is originally contained exclusively in the 2-pt statistics leaks into higher-order statistics. In cosmic shear studies, several authors have shown that the lowest order of them, i.e. the 3-pt statistics, already adds much information to the 2-pt one; in particular it can break the near degeneracy between the density parameter $\Omega_{\rm m}$ and the power spectrum normalization σ_8 (Bernardeau et al. 1997; Jain & Seljak 1997; van Waerbeke et al. 1999; Hui 1999). More recent studies by Takada & Jain (2004), TJ04 afterwards, and Bergé et al. (2010), showed that including 3-pt statistics can improve parameter constraints significantly, typically by a factor of three.

In order to quantify the information content in lensing 3-pt statistics theoretically, one needs to have an expression for the covariance matrix of the 3-pt statistics. In this chapter we aim at deriving an expression for the bispectrum covariance $\langle B(\ell_1,\ell_2,\ell_3)B(\ell_4,\ell_5,\ell_6)\rangle$ for cosmic shear.

Previous work done within a flat-sky spherical harmonic formalism (Hu 2000) in the context of the CMB has been frequently referred to for such an expression. However several drawbacks exist in this approach. For instance, the expression given by Hu (2000) is valid only for integer arguments and does not allow a free binning choice, whereas it is desirable to evaluate the bispectrum and its covariance at real-valued angular frequencies and use e.g. a logarithmic binning. The other drawbacks are formal ones, e.g. the formula contains the Wigner symbol whose physical meaning within a flat-sky consideration remains obscure; the finite survey size is accounted for only by multiplying a factor, which lacks justification. There is also an unjustified assumption made in the coordinate transformation between the full sky and the 2D plane.

All these drawbacks are associated with the spherical harmonic formalism Hu (2000) adopted. Thus we attempt a pure 2D Fourier-plane approach. We also work in the flat-sky limit since it greatly simplifies the mathematical form. Furthermore, the flat-sky limit is appropriate for practically all applications of weak lensing as the correlation of signals is only measured up to separations of a few degrees.

The major results of this chapter is published in Joachimi et al. (2009).

5.1 Bispectrum estimator

5.1.1 Estimator for $B(\ell_1, \ell_2, \ell_3)$

The first and the most crucial step of deriving an expression for the bispectrum covariance is to find a proper expression for the estimator of the bispectrum $B(\ell_1, \ell_2, \ell_3)$, where the ℓ's are real-valued angular frequencies in our approach. We will use the convergence bispectrum (3.37) instead of the shear bispectrum due to formal simplicity. Since the two differ only by a phase factor, the result can easily be applied to the shear bispectrum.

To define an estimator of the bispectrum is to express it in terms of the convergence κ, i.e. to 'invert' the equation (3.37). There are three points to consider in doing so. First, the argument ℓ's in the bispectrum are the absolute values of the vector $\boldsymbol{\ell}$'s, suggesting that angular averaging is needed. Second, the bispectrum is defined only when the triangle condition is satisfied. If this condition is satisfied, the value of the Dirac delta function is infinity, nevertheless one needs to 'invert' it to obtain an estimator for the bispectrum. This seemingly unsolvable problem vanishes if one considers a finite survey size A. The Dirac Delta function can be expressed as

$$\delta_{\rm D}^{(2)}(\boldsymbol{\ell}) = \frac{1}{(2\pi)^2} \int \mathrm{d}^2 x \;\; \mathrm{e}^{\mathrm{i}\boldsymbol{\ell}\cdot\boldsymbol{x}} \,. \tag{5.1}$$

One can easily verify that, when the integral on the r.h.s. of (5.1) is confined to a region with size A, one has $\delta_{\rm D}^{(2)}(\boldsymbol{\ell} \to \boldsymbol{0}) \to A/(2\pi)^2$ instead of infinity, which means the inversion of the delta function here should simply give a factor $1/A$. Third, one still needs to specify the triangle condition. This can be done by adding a $\delta_{\rm D}^{(2)}(\boldsymbol{\ell_1} + \boldsymbol{\ell_2} + \boldsymbol{\ell_3})$ to the estimator.

Having taken care of all three points, our estimator of the bispectrum reads

$$\hat{B}(\ell_1, \ell_2, \ell_3) = \frac{1}{\Gamma \cdot A} \int_0^{2\pi} \frac{\mathrm{d}\phi_{\ell_1}}{2\pi} \int_0^{2\pi} \frac{\mathrm{d}\phi_{\ell_2}}{2\pi} \int_0^{2\pi} \frac{\mathrm{d}\phi_{\ell_3}}{2\pi} \; \tilde{\kappa}(\boldsymbol{\ell_1}) \, \tilde{\kappa}(\boldsymbol{\ell_2}) \, \tilde{\kappa}(\boldsymbol{\ell_3}) \, \delta_{\rm D}^{(2)}(\boldsymbol{\ell_1} + \boldsymbol{\ell_2} + \boldsymbol{\ell_3}) \,, \tag{5.2}$$

where ϕ_{ℓ_i} is the polar angle of $\boldsymbol{\ell}$, and we have put in a normalization function Γ to keep the estimator unbiased:

$$\Gamma = \int_0^{2\pi} \frac{\mathrm{d}\phi_{\ell_1}}{2\pi} \int_0^{2\pi} \frac{\mathrm{d}\phi_{\ell_2}}{2\pi} \int_0^{2\pi} \frac{\mathrm{d}\phi_{\ell_3}}{2\pi} \; \delta_{\rm D}^{(2)}(\boldsymbol{\ell_1} + \boldsymbol{\ell_2} + \boldsymbol{\ell_3}) \,, \tag{5.3}$$

which is the angular average of the 2D delta function $\delta_{\rm D}^{(2)}(\boldsymbol{\ell_1} + \boldsymbol{\ell_2} + \boldsymbol{\ell_3})$.

The next step is to express Γ in terms of the absolute values of the $\boldsymbol{\ell}$'s. This we achieve by writing the Dirac delta function in its integral form and exchanging the order of the integrals,

$$\begin{aligned} \Gamma &= \int_0^{2\pi} \frac{\mathrm{d}\phi_{\ell_1}}{2\pi} \int_0^{2\pi} \frac{\mathrm{d}\phi_{\ell_2}}{2\pi} \int_0^{2\pi} \frac{\mathrm{d}\phi_{\ell_3}}{2\pi} \int \frac{\mathrm{d}^2\theta}{(2\pi)^2} \, \mathrm{e}^{\mathrm{i}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3)\cdot\boldsymbol{\theta}} \\ &= \int \frac{\mathrm{d}\theta \, \theta}{2\pi} J_0(\ell_1\theta) \, J_0(\ell_2\theta) \, J_0(\ell_3\theta) \\ &= \frac{1}{(2\pi)^2} \, \Lambda(\ell_1, \ell_2, \ell_3) \,. \end{aligned} \tag{5.4}$$

The last expression in (5.4) was given in Gradshteyn et al. (2000), where Λ is defined to be

$$\Lambda(\ell_1, \ell_2, \ell_3) \equiv \begin{cases} \left\{ \frac{1}{4} \sqrt{2\ell_1^2\ell_2^2 + 2\ell_1^2\ell_3^2 + 2\ell_2^2\ell_3^2 - \ell_1^4 - \ell_2^4 - \ell_3^4} \right\}^{-1} & \text{if} \quad |\ell_1 - \ell_2| < \ell_3 < \ell_1 + \ell_2 \\ 0 & \text{else} \end{cases} . \tag{5.5}$$

5.1.2 Geometrical interpretation

It is interesting to note that Λ^{-1} is just the area of the triangle constructed by $\boldsymbol{\ell_1}$,$\boldsymbol{\ell_2}$ and $\boldsymbol{\ell_3}$. This motivated us to find a geometrical interpretation for the angular averaging of $\delta_{\rm D}^{(2)}(\boldsymbol{\ell_1} + \boldsymbol{\ell_2} + \boldsymbol{\ell_3})$ in (5.3). If one fixes the lengths of $\boldsymbol{\ell_1}$,$\boldsymbol{\ell_2}$ and $\boldsymbol{\ell_3}$ and allows their polar angles to vary, in almost all cases the three of them do not form a triangle. Since the delta function specifies the triangle condition of the three vectors, it actually corresponds to the probability of the three vectors forming a triangle when their polar angles can be any value from 0 to 2π. Based on this idea, we consider a fixed vector $\boldsymbol{\ell_1}$, and allow $\boldsymbol{\ell_2}$ and $\boldsymbol{\ell_3}$ to vary within annuli with widths $\Delta\ell_2$ and $\Delta\ell_3$, as sketched in Fig. 5.1.

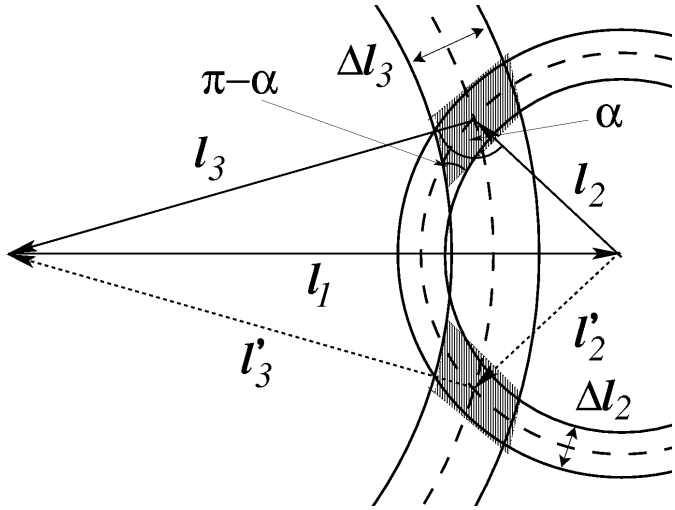


Figure 5.1: Sketch of the annuli and their overlap for fixed $\boldsymbol{\ell_1}$. The region of overlap is approximated by the shaded parallelograms. Figure from Joachimi et al. (2009).

The probability of the three vectors forming a triangle can be represented by the area of the overlap regions of the two annuli $A_{\parallel}$ divided by the areas of the annuli $A_R(\ell_2)$ and $A_R(\ell_3)$, in the limit of $\Delta\ell_2, \Delta\ell_3 \to 0$. Thus

$$\int_0^{2\pi} \frac{{\rm d}\phi_{\ell_1}}{2\pi} \int_0^{2\pi} \frac{{\rm d}\phi_{\ell_2}}{2\pi} \int_0^{2\pi} \frac{{\rm d}\phi_{\ell_3}}{2\pi}\, \delta_{\rm D}^{(2)}(\boldsymbol{\ell}_1 + \boldsymbol{\ell}_2 + \boldsymbol{\ell}_3) = \lim_{\Delta\ell_2,\Delta\ell_3\to 0} \frac{2\,A_{\parallel}}{A_R(\ell_2)\,A_R(\ell_3)}\,. \tag{5.6}$$

With the triangle formed by $\boldsymbol{\ell_1}$, $\boldsymbol{\ell_2}$, and $\boldsymbol{\ell_3}$ being parametrized by ℓ_2, ℓ_3, and α, which is the internal angle opposite ℓ_1 (see Fig. 5.1), Λ as defined in (5.5) can be written as $\Lambda = 2\ell_2^{-1}\ell_3^{-1}/\sin\alpha$. Observing that $A_{\parallel} = \Delta\ell_2\Delta\ell_3/\sin\alpha$, and

$$A_R(\bar{\ell}_i) = 2\pi\bar{\ell}_i\Delta\ell_i \quad \text{when } \Delta\ell_i \to 0 \text{ for } i = 1,2,3\,, \tag{5.7}$$

one reproduces (5.4).

5.1.3 Estimator for bin-averaged bispectrum $B(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3)$

In practice, the bispectrum is estimated not at every angular frequency but in angular frequency bins. Thus we further average (5.2) over the bin-widths to obtain the bin-averaged bispectrum esti-

mator

$$\begin{aligned}\hat{B}(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3) &= \int_{\Delta\ell_1}\frac{\mathrm{d}\ell_1}{\Delta\ell_1}\int_{\Delta\ell_2}\frac{\mathrm{d}\ell_2}{\Delta\ell_2}\int_{\Delta\ell_3}\frac{\mathrm{d}\ell_3}{\Delta\ell_3}\,\hat{B}(\ell_1,\ell_2,\ell_3)\\ &= \frac{(2\pi)^2}{A}\Lambda^{-1}\left(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3\right)\int_{A_R(\bar{\ell}_1)}\frac{\mathrm{d}^2\ell_1}{A_R(\bar{\ell}_1)}\int_{A_R(\bar{\ell}_2)}\frac{\mathrm{d}^2\ell_2}{A_R(\bar{\ell}_2)}\int_{A_R(\bar{\ell}_3)}\frac{\mathrm{d}^2\ell_3}{A_R(\bar{\ell}_3)}\\ &\quad\times\,\kappa(\boldsymbol{\ell_1})\,\kappa(\boldsymbol{\ell_2})\,\kappa(\boldsymbol{\ell_3})\,\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell_1}+\boldsymbol{\ell_2}+\boldsymbol{\ell_3})\,,\end{aligned} \tag{5.8}$$

which takes the average over the angular frequency annuli $A_R(\bar{\ell}_i)$.

We demonstrate that (5.8) is an unbiased estimator by taking the ensemble average of the estimator,

$$\begin{aligned}\left\langle\hat{B}(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3)\right\rangle &= \frac{(2\pi)^2}{A}\Lambda^{-1}\left(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3\right)\int_{A_R(\bar{\ell}_1)}\frac{\mathrm{d}^2\ell_1}{A_R(\bar{\ell}_1)}\int_{A_R(\bar{\ell}_2)}\frac{\mathrm{d}^2\ell_2}{A_R(\bar{\ell}_2)}\int_{A_R(\bar{\ell}_3)}\frac{\mathrm{d}^2\ell_3}{A_R(\bar{\ell}_3)}\\ &\quad\times(2\pi)^2\left(\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3)\right)^2 B(\ell_1,\ell_2,\ell_3)\\ &= (2\pi)^2\Lambda^{-1}\left(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3\right)\int_{A_R(\bar{\ell}_1)}\frac{\mathrm{d}^2\ell_1}{A_R(\bar{\ell}_1)}\int_{A_R(\bar{\ell}_2)}\frac{\mathrm{d}^2\ell_2}{A_R(\bar{\ell}_2)}\int_{A_R(\bar{\ell}_3)}\frac{\mathrm{d}^2\ell_3}{A_R(\bar{\ell}_3)}\\ &\quad\times\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3)\,B(\ell_1,\ell_2,\ell_3)\,.\end{aligned} \tag{5.9}$$

In the first step the definition of the bispectrum (3.37) was inserted, whereas in the second step the identity $\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}\to\mathbf{0})\to A/(2\pi)^2$ has been used.

Further inserting (5.4) into (5.9), one obtains

$$\begin{aligned}\left\langle\hat{B}(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3)\right\rangle &= (2\pi)^3\Lambda^{-1}\left(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3\right)\int_{\Delta\ell_1}\frac{\mathrm{d}\ell_1\ell_1}{A_R(\bar{\ell}_1)}\int_{\Delta\ell_2}\frac{\mathrm{d}\ell_2\ell_2}{A_R(\bar{\ell}_2)}\int_{\Delta\ell_3}\frac{\mathrm{d}\ell_3\ell_3}{A_R(\bar{\ell}_3)}\\ &\quad\times\Lambda\left(\ell_1,\ell_2,\ell_3\right)B(\ell_1,\ell_2,\ell_3)\,.\end{aligned} \tag{5.10}$$

We take the approximation that the annuli are thin enough such that $\Lambda\left(\ell_1,\ell_2,\ell_3\right)$ within the integral can be taken out of the integration and be replaced by $\Lambda\left(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3\right)$. Applying in addition (5.7), one arrives at

$$\left\langle\hat{B}(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3)\right\rangle \approx \int_{\Delta\ell_1}\frac{\mathrm{d}\ell_1\ell_1}{\bar{\ell}_1\Delta\ell_1}\int_{\Delta\ell_2}\frac{\mathrm{d}\ell_2\ell_2}{\bar{\ell}_2\Delta\ell_2}\int_{\Delta\ell_3}\frac{\mathrm{d}\ell_3\ell_3}{\bar{\ell}_3\Delta\ell_3}\,B(\ell_1,\ell_2,\ell_3)\equiv B(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3)\,, \tag{5.11}$$

where in the last step the definition of the bin-averaged bispectrum, which has a similar form as (5.8), was used. Hence, (5.8) defines an unbiased estimator of the bin-averaged bispectrum.

5.2 Bispectrum covariance

The covariance of the bin-averaged bispectrum is defined as

$$\begin{aligned}&\mathrm{Cov}\left(\hat{B}(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3),\ \hat{B}(\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6)\right)\\ &\equiv\left\langle\left(\hat{B}(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3)-\left\langle\hat{B}(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3)\right\rangle\right)\left(\hat{B}(\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6)-\left\langle\hat{B}(\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6)\right\rangle\right)\right\rangle\\ &=\left\langle\hat{B}(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3)\,\hat{B}(\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6)\right\rangle-B(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3)\,B(\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6)\,.\end{aligned} \tag{5.12}$$

With the expression of bispectrum estimator at hand, expanding the r.h.s. of (5.12) is a rather straightforward process, though tedious since it involves expanding $\left\langle \hat{B}(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3)\ \hat{B}(\bar{\ell}_4, \bar{\ell}_5, \bar{\ell}_6) \right\rangle$, a 6-pt correlator into its connected parts. The total bispectrum covariance we obtained reads

$$
\begin{aligned}
&\mathrm{Cov}\left(B(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3),\ B(\bar{\ell}_4, \bar{\ell}_5, \bar{\ell}_6)\right) \\
&= \frac{(2\pi)^3}{A\ \bar{\ell}_1\bar{\ell}_2\bar{\ell}_3\Delta\ell_1\Delta\ell_2\Delta\ell_3}\ \Lambda^{-1}\left(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3\right)\ D_{\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3,\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6}\ P(\bar{\ell}_1)P(\bar{\ell}_2)P(\bar{\ell}_3) \\
&\quad + \frac{C}{A}\ \delta_{\bar{\ell}_3\bar{\ell}_4} \int_1\int_2\int_3\int_5\int_6 \delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_1 + \boldsymbol{\ell}_2 + \boldsymbol{\ell}_3)\ \delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_3 + \boldsymbol{\ell}_5 + \boldsymbol{\ell}_6)\ B(\ell_1, \ell_2, \ell_3)\ B(\ell_3, \ell_5, \ell_6) + (8\ \text{perm.}) \\
&\quad + \frac{C}{A}\ \delta_{\bar{\ell}_3\bar{\ell}_6} \int_1\int_2\int_3\int_4\int_5 \delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_1 + \boldsymbol{\ell}_2 + \boldsymbol{\ell}_3)\ \delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_4 + \boldsymbol{\ell}_5 - \boldsymbol{\ell}_3)\ P_4(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_4, \boldsymbol{\ell}_5)\ P(\ell_3) + (8\ \text{perm.}) \\
&\quad + \frac{C}{A} \int_1\int_2\int_3\int_4\int_5\int_6 \delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_1 + \boldsymbol{\ell}_2 + \boldsymbol{\ell}_3)\ \delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_4 + \boldsymbol{\ell}_5 + \boldsymbol{\ell}_6)\ P_6(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3, \boldsymbol{\ell}_4, \boldsymbol{\ell}_5, \boldsymbol{\ell}_6)\ ,
\end{aligned}
\tag{5.13}
$$

where the prefactor reads $C \equiv (2\pi)^6\ \Lambda^{-1}\left(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3\right)\ \Lambda^{-1}\left(\bar{\ell}_4, \bar{\ell}_5, \bar{\ell}_6\right)$, P_4 and P_6 stand for the trispectrum and the pentaspectrum respectively, and shorthand notations are defined:

$$
\int_{A_R(\bar{\ell}_i)} \mathrm{d}^2\ell_i / A_R(\bar{\ell}_i) \equiv \int_i ,
\tag{5.14}
$$

and

$$
\begin{aligned}
D_{\ell_1,\ell_2,\ell_3,\ell_4,\ell_5,\ell_6} \equiv &\delta_{\ell_1\ell_4}\ \delta_{\ell_2\ell_5}\ \delta_{\ell_3\ell_6} + \delta_{\ell_1\ell_5}\ \delta_{\ell_2\ell_4}\ \delta_{\ell_3\ell_6} + \delta_{\ell_1\ell_4}\ \delta_{\ell_2\ell_6}\ \delta_{\ell_3\ell_5} \\
&+ \delta_{\ell_1\ell_5}\ \delta_{\ell_2\ell_6}\ \delta_{\ell_3\ell_4} + \delta_{\ell_1\ell_6}\ \delta_{\ell_2\ell_4}\ \delta_{\ell_3\ell_5} + \delta_{\ell_1\ell_6}\ \delta_{\ell_2\ell_5}\ \delta_{\ell_3\ell_4} .
\end{aligned}
\tag{5.15}
$$

In the linear and slightly non-linear regime, the first term of (5.13) dominates. Up to now it has been a common practice to use the first term of (5.13) to approximate the total bispectrum covariance. This approximation, dubbed the *Gaussian approximation* for the bispectrum covariance, is actually not well-justified for applying to the actual convergence field. However, the use the full covariance is heavily constrained by its computational load. Therefore it is necessary to investigate how well the Gaussian approximation holds in the non-Gaussian regime. A recent study (Martin 2011) found that for the matter density field in the local Universe it is marginally justified to use the Gaussian approximation.

5.3 Comparison with the spherical harmonic approach

On the celestial sphere one can decompose the random field κ into spherical harmonics, which yields a set of coefficients $\kappa_{\ell m}$ with integer ℓ and m satisfying $\ell > 0$ and $-\ell \leq m \leq \ell$. Hu (2000) defined a bispectrum estimator for CMB observables in terms of spherical harmonics coefficients. We reproduce it for the convergence field κ as

$$
\hat{B}_{\ell_1,\ell_2,\ell_3} = \sum_{m_1,m_2,m_3} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \kappa_{\ell_1 m_1}\ \kappa_{\ell_2 m_2}\ \kappa_{\ell_3 m_3} ,
\tag{5.16}
$$

where the symbol with the parenthesis is the Wigner-3j symbol. It obeys the triangle condition, i.e. it is non-zero only for $|\ell_1 - \ell_2| \leq \ell_3 \leq \ell_1 + \ell_2$ and permutations thereof. In addition, the Wigner

symbol with $m_1 = m_2 = m_3 = 0$ vanishes for odd $\ell_1 + \ell_2 + \ell_3$. The Gaussian approximation of the covariance of $\hat{B}_{\ell_1,\ell_2,\ell_3}$ is given as (Hu 2000)

$$\mathrm{Cov}\left(B_{\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3},\ B_{\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6}\right) = \frac{4\pi}{A} D_{\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3,\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6}\ P_{\bar{\ell}_1} P_{\bar{\ell}_2} P_{\bar{\ell}_3}\ , \tag{5.17}$$

where P_ℓ denotes the full-sky power spectrum. An *ad hoc* factor of $f_{\rm sky}^{-1} = 4\pi/A$ has been added to account for finite sky coverage of the survey.

Approximate relations between the spherical harmonic and Fourier-plane power spectra and bispectra are given in the same paper. They are

$$P_\ell \approx P(\ell)\ ; \qquad B_{\ell_1,\ell_2,\ell_3} \approx \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix} \sqrt{\frac{(2\ell_1+1)(2\ell_2+1)(2\ell_3+1)}{4\pi}}\ B(\ell_1,\ell_2,\ell_3)\ , \tag{5.18}$$

where the approximations hold well for $\ell_1, \ell_2, \ell_3 \gg 1$.

With (5.16), (5.17), and (5.18) one can derive a flat-sky spherical harmonic covariance as (Hu 2000; Takada & Jain 2004)

$$\left\langle \hat{B}(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3)\ \hat{B}(\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6)\right\rangle_{\rm sp} \approx \frac{(4\pi)^2\ D_{\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3,\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6}}{A\,(2\bar{\ell}_1+1)\,(2\bar{\ell}_2+1)\,(2\bar{\ell}_3+1)} \begin{pmatrix} \bar{\ell}_1 & \bar{\ell}_2 & \bar{\ell}_3 \\ 0 & 0 & 0 \end{pmatrix}^{-2} P(\bar{\ell}_1)\ P(\bar{\ell}_2)\ P(\bar{\ell}_3)\ , \tag{5.19}$$

where angular frequencies are required to be integer, and $\ell_1 + \ell_2 + \ell_3$ even. We have put a subscript 'sp' to distinguish it from the covariance matrix obtained through the Fourier plane approach. As is true for the Fourier plane approach, (5.19) holds for $\ell \gg 1$ only.

To compare the widely used formula (5.19) to our results, a relation between the Wigner symbol and Λ has to be found. We refer to Borodin et al. (1978) and the references therein for an approximation formula for the Wigner 3-j symbol,

$$\begin{aligned} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix}^2 \approx \frac{2}{\pi} \Bigg\{ & 2\left(\ell_1+\frac{1}{2}\right)^2 \left(\ell_2+\frac{1}{2}\right)^2 + 2\left(\ell_2+\frac{1}{2}\right)^2 \left(\ell_3+\frac{1}{2}\right)^2 \\ & + 2\left(\ell_3+\frac{1}{2}\right)^2 \left(\ell_1+\frac{1}{2}\right)^2 - \left(\ell_1+\frac{1}{2}\right)^4 - \left(\ell_2+\frac{1}{2}\right)^4 - \left(\ell_3+\frac{1}{2}\right)^4 \Bigg\}^{-1/2}\ , \end{aligned} \tag{5.20}$$

which is a very good approximation for $\ell_1, \ell_2, \ell_3 \gg 1$. We further make the approximation that $\ell_i + 1/2 \approx \ell_i$, which allow us to find from (5.5) and (5.20)

$$\begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix}^2 \approx \frac{\Lambda(\ell_1,\ell_2,\ell_3)}{2\pi}\ . \tag{5.21}$$

Using the Gaussian approximation, i.e. keeping the first term of (5.13), inserting (5.21), and specifying that $\Delta\ell_1 = \Delta\ell_2 = \Delta\ell_3 = 1$, one can see that the resulting relation is equivalent to (5.19) if one takes the limit $2\ell + 1 \approx 2\ell$ for $\ell \gg 1$, except for (5.19) being a factor of 2 smaller.

This factor of 2 discrepancy is simply related to the fact that (5.13) is defined for all ℓ values whereas (5.19) is defined only for $\ell_1 + \ell_2 + \ell_3$ being even. If one calculates the Fisher information using both approaches, the number of bispectrum entries in the Fourier plane approach is twice as large as that in the spherical harmonic approach. Therefore the factor of 2 difference in their covariance matrices is actually required to guarantee the agreement of the two approaches in terms of their results on the information content. We calculate the Fisher information contributed by lensing bispectra with $100 \le \ell_1 \le \ell_2 \le \ell_3 \le \ell_{\rm max}$ and $\ell_{\rm max}$ ranging from 100 and 150 with both approaches. The result is shown in Fig. 5.2, which shows good agreement between the two approaches (for more details see Joachimi et al. 2009).

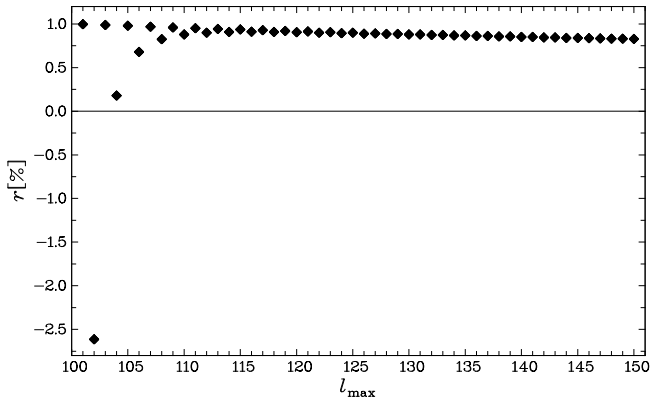


Figure 5.2: Comparison of the Fisher information as obtained by spherical harmonics and Fourier-plane approach. Given is the relative deviation r as a function of the maximum angular frequency $\ell_{\max}$. Figure taken from Joachimi et al. (2009).

5.4 Conclusion

In this study we derived the form of the bispectrum covariance in the flat-sky approximation. We defined an unbiased bispectrum estimator for 2D Fourier modes, averaged it over angular frequency bins to mimic the measuring process in reality, and computed the covariance of the averaged bispectrum estimator. During this process a geometrical interpretation of the integral

$$\int_0^{2\pi} \frac{\mathrm{d}\phi_{\ell_1}}{2\pi} \int_0^{2\pi} \frac{\mathrm{d}\phi_{\ell_2}}{2\pi} \int_0^{2\pi} \frac{\mathrm{d}\phi_{\ell_3}}{2\pi}\, \delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_1 + \boldsymbol{\ell}_2 + \boldsymbol{\ell}_3)\,,$$

which is needed in the averaging step, was found.

We showed the equivalence between the covariance matrix we derived and that given in Hu (2000) in terms of the Fisher information content they lead to. Moreover, we showed that our approach does not suffer from the drawbacks of the spherical harmonic approach used in Hu (2000). In Hu (2000) the covariance matrix formula contains the Wigner symbol whose physical meaning within a flat-sky consideration remains obscure; the finite survey size is accounted for only by multiplying a factor; and an unjustified assumption is made in the coordinate transformation between full sky and 2D plane. In addition to these formal drawbacks, the covariance matrix in Hu (2000) is valid only for integer arguments and does not allow a free binning choice due to the limitation of the spherical harmonic approach, whereas ours can evaluate the bispectrum and its covariance at real-valued angular frequencies and use e.g. a logarithmic binning.

Chapter 6

Controlling intrinsic-shear alignment in three-point weak lensing statistics

3-pt weak lensing statistics provide cosmic information that complements 2-pt statistics. However, both statistics suffer from systematic errors. The ultimate performance of these future lensing surveys largely depends on the how well the systematic errors can be controlled (e.g. Huterer et al. 2006). In this study we focus on a particularly worrisome systematic error in cosmic shear studies: the intrinsic-shear alignment, and demonstrate a way to control it for shear 3-pt statistics. Specifically, we generalize the nulling technique, which is a model-independent method developed to eliminate intrinsic-shear alignment at the 2-pt level, to the 3-pt level, and thereby controlling the corresponding 3-pt systematics. The content in this chapter is published in Shi et al. (2010).

6.1 Intrinsic-shear alignment

In the weak lensing limit the observed ellipticity of a galaxy $\epsilon_{\rm obs}$ can be written as the sum of the intrinsic ellipticity $\epsilon_{\rm I}$ of the galaxy, and the shear γ which is caused by gravitational lensing of the foreground matter distribution,

$$\epsilon_{\rm obs} = \epsilon_{\rm I} + \gamma\,. \tag{6.1}$$

Here $\epsilon_{\rm obs}$, $\epsilon_{\rm I}$ and γ are complex quantities. Intrinsic-shear alignment is defined in 2-pt cosmic shear statistics as the correlation between the intrinsic ellipticity of one galaxy and the shear of another galaxy (the GI term, Hirata & Seljak 2004). The 3-pt statistics $\left\langle \epsilon^i_{\rm obs}\epsilon^j_{\rm obs}\epsilon^k_{\rm obs}\right\rangle$, a correlator of ellipticities of three galaxy images i, j and k, can also be expanded into lensing (GGG), intrinsic-shear (GGI and GII), and intrinsic (III) terms:

$$\left\langle \epsilon^i_{\rm obs}\epsilon^j_{\rm obs}\epsilon^k_{\rm obs}\right\rangle = {\rm GGG} + {\rm GGI} + {\rm GII} + {\rm III}\,, {\rm with} \tag{6.2}$$

$${\rm GGG} = \left\langle \gamma^i\gamma^j\gamma^k\right\rangle\,, \tag{6.3}$$

$${\rm GGI} = \left\langle \epsilon^i_{\rm I}\gamma^j\gamma^k\right\rangle + \left\langle \epsilon^j_{\rm I}\gamma^k\gamma^i\right\rangle + \left\langle \epsilon^k_{\rm I}\gamma^i\gamma^j\right\rangle\,, \tag{6.4}$$

$${\rm GII} = \left\langle \epsilon^i_{\rm I}\epsilon^j_{\rm I}\gamma^k\right\rangle + \left\langle \epsilon^j_{\rm I}\epsilon^k_{\rm I}\gamma^i\right\rangle + \left\langle \epsilon^k_{\rm I}\epsilon^i_{\rm I}\gamma^j\right\rangle\,, \tag{6.5}$$

$$\mathrm{III} = \left\langle \epsilon_\mathrm{I}^i \epsilon_\mathrm{I}^j \epsilon_\mathrm{I}^k \right\rangle \ . \tag{6.6}$$

Physically, if one assumes that galaxies are randomly oriented on the sky, only the desired GGG term remains on the right-hand side of (6.2). However, when these galaxies are subject to the tidal gravitational force of the same matter structure (e.g. they formed under the influence of the same massive dark matter halo), their shapes can intrinsically align and become correlated, giving rise to a nonvanishing III term. Furthermore, GGI and GII terms can be generated when a matter structure tidally influences close-by galaxies and at the same time contributes to the shear signal of background objects, leading to correlations among them.

In 2-pt statistics, the corresponding intrinsic (II) and intrinsic-shear (GI) terms have been subject to detailed studies both theoretically (e.g. Catelan et al. 2001; Croft & Metzler 2000; Heavens et al. 2000; Hui & Zhang 2002; Mackey et al. 2002; Jing 2002; Hirata & Seljak 2004; Heymans et al. 2006; Bridle & Abdalla 2007; Schneider & Bridle 2010) and observationally (Brown et al. 2002; Heymans et al. 2004; Mandelbaum et al. 2006, 2011; Hirata et al. 2007; Fu et al. 2008; Brainerd et al. 2009; Okumura et al. 2009; Okumura & Jing 2009; Joachimi et al. 2011). Although the results of these studies show large variations, most of them are consistent with a 10 % contamination by both II and GI correlations for future surveys with photometric redshift information. Especially, neglecting these correlations can bias the dark energy equation of state parameter w_0 by as much as 50 % (Bridle & King 2007) for a "shallow" survey as described in Amara & Réfrégier (2007). For 3-pt shear statistics, there have been few measurements up to now (Bernardeau et al. 2002b; Pen et al. 2003; Jarvis et al. 2004). However the potential systematics level in these studies is found to be high. A recent numerical study by Semboloni et al. (2008) showed that intrinsic alignments affect 3-pt weak lensing statistics more strongly than at the 2-pt level for a given survey depth. In particular, neglecting GGI and GII systematics would lead to an underestimation of the GGG signal by $5-10$ % for a moderately deep survey like the CFHTLS Wide. Therefore, to match the statistical power expected for cosmic shear in the future surveys, it is essential to control these systematics.

The intrinsic alignment, II (III) in the two- (three-) point case, is relatively straightforward to eliminate, since it requires that the galaxies in consideration are physically close to each other, i.e. have very similar redshifts and small angular separation (King & Schneider 2002, 2003; Heymans & Heavens 2003; Takada & White 2004). The control of intrinsic-shear systematics, GI for the 2-pt case and GGI in the 3-pt case (GII also requires that two of the three galaxies are physically close and thus can be eliminated in the same way as II and III), turns out to be a much greater challenge. However, as already pointed out by HS04, the characteristic dependence on galaxy redshifts is a valuable piece of information that helps to control the intrinsic-shear alignments.

Several methods for this have already been constructed in the context of 2-pt statistics. They can be roughly classified into three categories: modeling (King 2005; Bridle & King 2007; Bernstein 2009), nulling (Joachimi & Schneider 2008, JS08 hereafter; Joachimi & Schneider 2009) and self-calibration (Zhang 2010; Joachimi & Bridle 2010). Modeling separates cosmic shear from the intrinsic-shear alignment effect by constructing template functions for the latter. It suffers from uncertainties of the model due to the lack of knowledge of the angular scale and redshift dependence of the intrinsic-shear signal. The nulling technique employs the characteristic redshift dependence of the intrinsic-shear signal to "null it out". It is a purely geometrical method and is model-independent, but suffers from a significant information loss. Self-calibration intends to solve the problem of information loss by using additional information from the galaxy distribution to "calibrate" the signal. The original form of self-calibration, proposed by Zhang (2010), is model-independent but strong assumptions have been made. Joachimi & Bridle (2010) then develop it into a modeling method, by treating intrinsic alignments and galaxy biasing as free functions of scale and redshift.

All these methods have the potential of being generalized to 3-pt statistics. Here we focus on the nulling technique, and establish it as a method to reduce the 3-pt intrinsic-shear alignments GGI and GII. Since GII can be removed by discarding close pairs of galaxies as in the case of II controlling (e.g. Heymans & Heavens 2003), we focus on the control of GGI systematics.

We will work in the context of a spatially flat CDM cosmology with a variable dark energy whose equation of state is parameterized as $w_{\rm eos} = w_0 + w_a(1-a)$, with a being the cosmic scale factor. The adopted fiducial values for cosmological parameters are $\Omega_{\rm m} = 0.3$, $\Omega_{\rm b} = 0.045$, $\Omega_{\rm de} = 0.7$, $w_0 = -0.95$, $w_a = 0.0$, $h = 0.7$, $n_s = 1.0$, and $\sigma_8 = 0.8$. Here, h is the dimensionless Hubble parameter defined by $H_0 = 100\,h\,$km/s/Mpc.

6.2 The nulling technique applied to three-point shear tomography

6.2.1 Principle of the nulling technique

The shear on the image of a distant galaxy is a result of gravitational distortion of light caused by the inhomogeneous 3D matter distribution in the foreground of that galaxy. For notational simplicity, we will use the dimensionless surface mass density (the convergence) κ instead of the shear γ as a measure for the lensing signal throughout the chapter, although in reality the signal is based on the measurement of the shear. This will not affect our results since κ and γ are linearly related on each redshift plane while our method is dealing with the redshift dependence of them (the same reason justifies the turning to the Fourier domain in the next subsection).

When one measures the shear γ, the direct observable is the galaxy ellipticity $\epsilon_{\rm obs} = \epsilon_{\rm I} + \gamma$. The shear γ is a signal caused by gravitational distortion which is a deterministic process, where the intrinsic ellipticity $\epsilon_{\rm I}$ can be further written as the sum of a deterministic part $\epsilon_{\rm I}^{\rm det}$ which is caused by intrinsic alignment, and a stochastic part $\epsilon_{\rm I}^{\rm ran}$ which does not correlate with any other quantity. There is no correlation between $\epsilon_{\rm I}^{\rm ran}$ of different galaxies either.

We define $\kappa_{\rm obs}$ and $\kappa_{\rm I}$ which are the correspondences of $\epsilon_{\rm I}^{\rm det}+\gamma$ and $\epsilon_{\rm I}^{\rm det}$. We remove the stochastic part since κ is deterministic. Note that $\kappa_{\rm obs}$ and $\kappa_{\rm I}$ are analogs of the dimensionless surface mass density κ but do not have any direct physical meaning as κ does. They are complex quantities in general and can lead to a B-mode signal. To better distinguish the real measurable κ from them, we denote it as $\kappa_{\rm G}$ in the rest of this chapter since it is the physical quantity which is related to the gravitational lensing signal. Keeping the dominating linear term, the convergence $\kappa_{\rm G}$ can be written as (details see e.g. Schneider 2006):

$$\kappa_{\rm G}(\boldsymbol{\theta},\chi_{\rm s}) = \frac{3\,\Omega_{\rm m}H_0^2}{2c^2}\int_0^{\chi_{\rm s}} d\chi\,\frac{\chi(\chi_{\rm s}-\chi)}{\chi_{\rm s}}\,\frac{\delta(\chi\boldsymbol{\theta},\chi)}{a(\chi)}\,, \tag{6.7}$$

where δ is the 3D matter density contrast, $\chi_{\rm s}$ is the comoving distance of the background galaxy which is acting as a source, and $a(\chi)$ is the cosmic scale factor at the comoving distance χ of δ which is acting as a lens.

Equation (6.7) clearly shows that the contribution of the matter inhomogeneity δ at comoving distance χ_i to the cosmic shear signal of background galaxies can be considered as a function of the source distance $\chi_{\rm s}$, and this function is proportional to $1-\chi_i/\chi_{\rm s}$. The nulling technique takes advantage of this characteristic dependence on source distance $\chi_{\rm s}$ by constructing a weight function $T(\chi_i,\chi_{\rm s})$ such that the product of $T(\chi_i,\chi_{\rm s})$ and $1-\chi_i/\chi_{\rm s}$ has an average of zero on the range between χ_i and the comoving distance to the horizon $\chi_{\rm hor}$:

$$\int_{\chi_i}^{\chi_{\rm hor}} d\chi_{\rm s}\; T(\chi_i,\chi_{\rm s}) \left(1-\frac{\chi_i}{\chi_{\rm s}}\right) = 0\;. \tag{6.8}$$

One then uses this weight function as a weight for integrating over the source distance:

$$\hat{\kappa}_{\rm G}(\chi_i,\boldsymbol{\theta}) := \int_{\chi_i}^{\chi_{\rm hor}} d\chi_{\rm s}\; T(\chi_i,\chi_{\rm s})\,\kappa_{\rm G}(\boldsymbol{\theta},\chi_{\rm s})\;. \tag{6.9}$$

The resulting new measure of shear signal $\hat{\kappa}_{\rm G}(\chi_i,\boldsymbol{\theta})$ is then free of contributions from the matter inhomogeneity at distance χ_i. Note that although the weight function T has two arguments χ_i and $\chi_{\rm s}$ here, we consider it as a function of $\chi_{\rm s}$ for a particular χ_i.

Consider a correlator $\left\langle\kappa^i_{\rm obs}\kappa^j_{\rm obs}\right\rangle$ with comoving distances $\chi_i<\chi_j$. With a similar decomposition as (6.2), it is straightforward to see that the GI term in it is $\left\langle\kappa^i_{\rm I}\kappa^j_{\rm G}\right\rangle$. The term $\left\langle\kappa^i_{\rm G}\kappa^j_{\rm I}\right\rangle$ vanishes since the lensing signal at χ_i is correlated only with matter with $\chi\le\chi_i$, whereas $\kappa^j_{\rm I}$ originates solely from physical processes happening at χ_j. If we integrate $\left\langle\kappa^i_{\rm I}\kappa^j_{\rm G}\right\rangle$ over χ_j with a weight function that eliminates the contributions to $\kappa^j_{\rm G}$ by the matter inhomogeneity at distance χ_i, this correlator will also vanish,

$$\int_{\chi_i}^{\chi_{\rm hor}} d\chi_j\; T(\chi_i,\chi_j)\left\langle\kappa^i_{\rm I}\kappa^j_{\rm G}\right\rangle = 0, \tag{6.10}$$

since it is just the matter inhomogeneity at distance χ_i that gives rise to the correlation between $\kappa^i_{\rm I}$ and $\kappa^j_{\rm G}$. Thus, when we integrate over $\left\langle\kappa^i_{\rm obs}\kappa^j_{\rm obs}\right\rangle$ with the same weight function, the GI contamination in it will be "nulled out". Equation (6.8) is the condition that the weight function T should satisfy in order to "null" the intrinsic-shear alignment terms, so we call it "the nulling condition".

The same applies to 3-pt statistics. Consider a correlator $\left\langle\kappa^i_{\rm obs}\kappa^j_{\rm obs}\kappa^k_{\rm obs}\right\rangle$ with χ_i being the smallest comoving distance of the three. Both GII and GGI systematics contained in it also originate from the matter inhomogeneity at distance χ_i. Typically, the generation of GII systematics requires that $\chi_i\approx\chi_j<\chi_k$, while the generation of GGI requires $\chi_i<\chi_j$ and $\chi_i<\chi_k$. For both cases, the dependence of GII or GGI systematics on χ_k is also just $1-\chi_i/\chi_k$. So new measures built as $\int_{\chi_i}^{\chi_{\rm hor}} d\chi_k\; T(\chi_i,\chi_j,\chi_k)\left\langle\kappa^i_{\rm obs}\kappa^j_{\rm obs}\kappa^k_{\rm obs}\right\rangle$ with T satisfying the nulling condition for 3-pt statistics

$$\int_{\chi_i}^{\chi_{\rm hor}} d\chi_k\; T(\chi_i,\chi_j,\chi_k)\left(1-\frac{\chi_i}{\chi_k}\right) = 0 \tag{6.11}$$

will be free of both GII and GGI contamination. Again, $T(\chi_i,\chi_j,\chi_k)$ here should be seen as a function of χ_k whose form depends on χ_i and χ_j.

Note that this method only depends on the characteristic redshift dependencies of the lensing signal and intrinsic-alignment signals, and is not limited to E-mode fields. This is a reassuring feature since while the $\kappa_{\rm G}$ field is a pure E-mode field to first order, the $\kappa_{\rm I}$ field can have a B-mode component. However, if parity-invariance is assumed, any correlation function which contains an odd number of B-mode shear components vanishes (Schneider 2003), thus there should be no B-mode component in the GGI signal.

6.2.2 Nulling formalism for lensing bispectrum tomography

Since the nulling technique relies on the distinct redshift dependence of the intrinsic-alignment signal, redshift information is crucial for it. With the help of near-infrared bands, forthcoming multi-color imaging surveys can provide rather accurate photometric redshift information for the galaxies

(e.g. Abdalla et al. 2008; Bordoloi et al. 2010), allowing tomographic studies of cosmic shear statistics. We base our study on cosmic shear bispectrum tomography, and outline the corresponding formalism of the nulling technique in the following.

Given the galaxy redshift probability distribution of redshift bin i which we denote as $p_{\rm s}^{(i)}(z) = p_{\rm s}^{(i)}(\chi_{\rm s})\, d\chi_{\rm s}/dz$, one can define the average convergence field in redshift bin i by integrating $\kappa(\boldsymbol{\theta},\chi_{\rm s})$ in (6.7) over $p_{\rm s}^{(i)}(\chi_{\rm s})$. We turn to angular frequency space now and define

$$\tilde{\kappa}_{\rm G}^{(i)}(\boldsymbol{\ell}) := \int_0^{\chi_{\rm hor}} d\chi_{\rm s}\, p_{\rm s}^{(i)}(\chi_{\rm s})\, \tilde{\kappa}_{\rm G}(\boldsymbol{\ell},\chi_{\rm s})\,, \tag{6.12}$$

where $\tilde{\kappa}_{\rm G}(\boldsymbol{\ell},\chi_{\rm s})$ is the Fourier transform of $\kappa_{\rm G}(\boldsymbol{\theta},\chi_{\rm s})$. To better show the relation between $\tilde{\kappa}_{\rm G}^{(i)}(\boldsymbol{\ell})$ and 3D matter inhomogeneity in Fourier space $\tilde{\delta}\,(k,\chi)$, one can combine (6.7) and (6.12) and write

$$\tilde{\kappa}_{\rm G}^{(i)}(\boldsymbol{\ell}) = \int_0^{\chi_{\rm hor}} d\chi\; W^{(i)}(\chi)\, \tilde{\delta}\,(\boldsymbol{\ell}/\chi,\chi)\;\;, \tag{6.13}$$

by defining a lensing weight function $W^{(i)}(\chi)$ as

$$W^{(i)}(\chi) := \frac{3\,\Omega_{\rm m} H_0^2\,\chi}{2\,a(\chi)\,c^2} \int_\chi^{\chi_{\rm hor}} d\chi_{\rm s}\; p_{\rm s}^{(i)}(\chi_{\rm s}) \frac{\chi_{\rm s}-\chi}{\chi_{\rm s}}\;. \tag{6.14}$$

The tomographic lensing bispectrum is defined via

$$\left\langle \tilde{\kappa}_{\rm G}^{(i)}(\boldsymbol{\ell}_1)\tilde{\kappa}_{\rm G}^{(j)}(\boldsymbol{\ell}_2)\tilde{\kappa}_{\rm G}^{(k)}(\boldsymbol{\ell}_3)\right\rangle = (2\pi)^2 B_{\rm GGG}^{(ijk)}(\ell_1,\ell_2,\ell_3)\;\delta_D(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3)\,, \tag{6.15}$$

where the Dirac delta function ensures that the bispectrum is defined only when $\boldsymbol{\ell}_1$, $\boldsymbol{\ell}_2$, and $\boldsymbol{\ell}_3$ form a triangle. This fact arises from statistical homogeneity, while that the bispectrum can be defined as a function independent of the directions of the angular frequency vectors arises from statistical isotropy.

In a survey, the convergence field $\tilde{\kappa}_{\rm obs}$ is determined from the observed galaxy ellipticities, and the corresponding bispectrum $B_{\rm obs}$ suffers from intrinsic-shear alignments. As we did with the 3-pt correlator in Sect. 6.1, we separate the observed lensing bispectrum into the four terms:

$$B_{\rm obs} = B_{\rm GGG} + B_{\rm GGI} + B_{\rm GII} + B_{\rm III}\;. \tag{6.16}$$

Among them, $B_{\rm GGI}$, $B_{\rm GII}$ and $B_{\rm III}$ can be linked to the convergence in a similar way as (6.15), for example

$$\left\langle \tilde{\kappa}_{\rm I}^{(i)}(\boldsymbol{\ell}_1)\tilde{\kappa}_{\rm G}^{(j)}(\boldsymbol{\ell}_2)\tilde{\kappa}_{\rm G}^{(k)}(\boldsymbol{\ell}_3)\right\rangle = (2\pi)^2 B_{\rm GGI}^{(ijk)}(\ell_1,\ell_2,\ell_3)\;\delta_D(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3)\;. \tag{6.17}$$

Here we assume disjunct redshift bins and let i to be the redshift bin with the lowest redshift, so $\left\langle \tilde{\kappa}_{\rm G}^{(i)}(\boldsymbol{\ell}_1)\tilde{\kappa}_{\rm I}^{(j)}(\boldsymbol{\ell}_2)\tilde{\kappa}_{\rm G}^{(k)}(\boldsymbol{\ell}_3)\right\rangle$ and $\left\langle \tilde{\kappa}_{\rm G}^{(i)}(\boldsymbol{\ell}_1)\tilde{\kappa}_{\rm G}^{(j)}(\boldsymbol{\ell}_2)\tilde{\kappa}_{\rm I}^{(k)}(\boldsymbol{\ell}_3)\right\rangle$ both vanish due to the same reason as explained in Sect. 6.2.1 for the 2-pt statistics.

The purpose of the nulling technique is to filter $B_{\rm obs}$ in such a way that the GGI term is strongly suppressed in comparison with the GGG term. The GII and III terms can be removed by ignoring the signal coming from bispectrum $B_{\rm obs}^{(ijk)}(\ell_1,\ell_2,\ell_3)$ with two or three equal redshift bins.

To fulfill this purpose, we construct our new measures as

$$Y^{(ij)}(\ell_1,\ell_2,\ell_3) := \sum_{k=i+1}^{N_z} T^{(ij)}(\chi_k)\; B_{\rm obs}^{(ijk)}(\ell_1,\ell_2,\ell_3)\;\chi_k'\;\Delta z_k\;, \tag{6.18}$$

where N_z is the total number of redshift bins, χ'_k is the derivative of comoving distance with respect to redshift, and Δz_k is the width of redshift bin k. The weight function is written now as $T^{(ij)}(\chi_k)$ since i and j indicate two redshift bins, i.e. two populations of galaxies, rather than two comoving distances as in the previous subsection. The weight $T^{(ij)}$ is required to satisfy the nulling condition (6.8) in its discretized form,

$$O^{(ij)} := \sum_{k=i+1}^{N_z} T^{(ij)}(\chi_k) \left(1 - \frac{\chi_i}{\chi_k}\right) \chi'_k \, \Delta z_k = 0 \,, \tag{6.19}$$

for all $j > i$. Here, χ_i and χ_k should be chosen such that they represent well the distance to redshift bins i and k. In this study we choose them to be the distances corresponding to the median redshift of the bin. The summation over index k runs from $i+1$ rather than i since we consider only bispectrum measures with $j > i$ and $k > i$ to avoid III and GII systematics. In this case $B^{(ijk)}_{\rm obs}$ in (6.18) can be written as a sum of $B^{(ijk)}_{\rm GGG}$ and $B^{(ijk)}_{\rm GGI}$, and $Y^{(ij)}$ can be expressed as

$$Y^{(ij)}(\ell_1, \ell_2, \ell_3) = \sum_{k=i+1}^{N_z} T^{(ij)}(\chi_k) \, B^{(ijk)}_{\rm GGG}(\ell_1, \ell_2, \ell_3) \, \chi'_k \, \Delta z_k + \sum_{k=i+1}^{N_z} T^{(ij)}(\chi_k) \, B^{(ijk)}_{\rm GGI}(\ell_1, \ell_2, \ell_3) \, \chi'_k \, \Delta z_k \,. \tag{6.20}$$

Suppose one has infinitely many redshift bins, then the lensing signal in bin k caused by the matter inhomogeneity in bin i is exactly proportional to $1 - \chi_i/\chi_k$, which means $B^{(ijk)}_{\rm GGI}(\ell_1, \ell_2, \ell_3)$ can be written as a product of $1 - \chi_i/\chi_k$ and some function of the parameters other than χ_k:

$$B^{(ijk)}_{\rm GGI}(\ell_1, \ell_2, \ell_3) = \mathcal{F}(\chi_i, \chi_j, \ell_1, \ell_2, \ell_3) \left(1 - \frac{\chi_i}{\chi_k}\right) \,. \tag{6.21}$$

Then we have

$$\sum_{k=i+1}^{N_z} T^{(ij)}(\chi_k) \, B^{(ijk)}_{\rm GGI}(\ell_1, \ell_2, \ell_3) \, \chi'_k \, \Delta z_k = \mathcal{F}(\chi_i, \chi_j, \ell_1, \ell_2, \ell_3) \sum_{k=i+1}^{N_z} T^{(ij)}(\chi_k) \left(1 - \frac{\chi_i}{\chi_k}\right) \chi'_k \, \Delta z_k = 0 \,. \tag{6.22}$$

This suggests that only the GGG contribution is left in the nulled measure $Y^{(ij)}$, the GGI contribution has been "nulled out" due to the nulling condition. If only a limited number of redshift bins is available, (6.21) holds only approximately, leading to a residual in (6.22).

Since the nulling condition is the only condition that the weight $T^{(ij)}$ must satisfy in order to "null", there is much freedom in choosing the form of it. We would like to further specify its form such that it preserves as much Fisher information in $Y^{(ij)}$ as possible. The method we have adopted for the nulling weight construction will be detailed in Sect. 6.4.

Note that for each (i, j) combination, one can in principle apply more than one nulling weight to the original bispectrum, and obtain more nulled measures. If one retains the condition of maximizing the Fisher information and demands that all the weight functions built for one (i, j) combination are orthogonal to each other, one arrives at higher-order modes that have the second-most, third-most, etc., information content (higher-order weights, see JS08). The total number of such linearly independent nulled measures for a certain (i, j) equals the possible values of $k \geq i + 1$. In this schematic study we will only use the optimum, i.e. the first-order nulling weights. We will assess the information loss due to this limitation in Sect. 6.5.4.

6.3 Modeling

6.3.1 Survey characteristics

We set up a fictitious survey with a survey size of $A = 4000$ deg^2 which is similar to the survey size of DES. This can be easily scaled to any survey size using the proportionality of statistical errors to $A^{-1/2}$. We assume a galaxy intrinsic ellipticity dispersion $\sigma_\epsilon = \sigma(\epsilon_{\rm I}^{\rm ran}) = 0.35$. As galaxy redshift probability distribution we adopt the frequently used parameterization (Smail et al. 1994),

$$p_{\rm s}(z) \propto \left(\frac{z}{z_0}\right)^\alpha \exp\left\{-\left(\frac{z}{z_0}\right)^\beta\right\}, \tag{6.23}$$

and use $z_0 = 0.64$, $\alpha = 2$, $\beta = 1.5$. The distribution is cut at $z_{\rm max} = 3$ and normalized to 1. The corresponding median redshift of this fictitious survey is $z_{\rm m} = 0.9$, which is compatible to a survey like EUCLID. We adopt an average galaxy number density $\bar{n}_{\rm g} = 40$ arcmin^{-2} which is again EUCLID-like.

Disjunct redshift bins without photo-z error are assumed, which means that the galaxy redshift probability distribution in redshift bin i takes the form $p_{\rm s}^{(i)}(z) \propto p_{\rm s}(z)$ if and only if the redshift that corresponds to comoving distance $\chi_{\rm s}$ is within the boundaries of redshift bin i. A number of 10 redshift bins is used by default. The boundaries of the redshift bins are set such that each bin contains the same number of galaxies.

We adopt 20 angular frequency bins spaced logarithmically between $\ell_{\rm min} = 50$ and $\ell_{\rm max} = 3000$, and denote the characteristic angular frequency of a bin as $\bar{\ell}$. Within this range the noise properties of the cosmic shear field are still not too far in the non-Gaussian regime, allowing a more realistic theoretical estimation of the bispectrum and its covariance. Whether this number of angular frequency bins can reconstruct the angular frequency dependence of the bispectrum is tested, and 20 bins are found to be sufficient for our requirements on precision. This is also expected since the bispectrum is rather featureless as a function of angular frequency.

6.3.2 Bispectrum and its covariance

We show the modeling of $B_{\rm GGG}$ and its covariance in this section. We will only consider the tomographic bispectrum at redshift bins satisfying $z_i < z_j$ and $z_i < z_k$, which already ensures an elimination of $B_{\rm III}$ and $B_{\rm GII}$ systematics in our case.

Applying Limber's equation, it can be shown that the tomographic convergence bispectrum can be written as a projection of the 3D matter bispectrum $B_\delta(k_1, k_2, k_3; \chi)$ (see e.g. TJ04):

$$B_{\rm GGG}^{(ijk)}(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3) = \int_0^{\chi_{\rm hor}} d\chi \, \frac{W^{(i)}(\chi) W^{(j)}(\chi) W^{(k)}(\chi)}{\chi^4} \, B_\delta\left(\frac{\bar{\ell}_1}{\chi}, \frac{\bar{\ell}_2}{\chi}, \frac{\bar{\ell}_3}{\chi}; \chi\right) . \tag{6.24}$$

To compute B_δ, we employ the fitting formula by Scoccimarro & Couchman (2001). A comparison of this formula with the halo model results can be found in Takada & Jain (2003a,b).

We use the expression (5.13) for the bispectrum covariance and keep only the first term of it. This Gaussian approximation is justified in this case since we constrain ourselves by using angular scales with $\ell \leq 3000$. As argued in TJ04, the first term in (5.13) still dominates for ℓ values in this range. Keeping only this term and adapt (5.13) to tomographic lensing bispectrum, the bispectrum covariance reads

$$\mathrm{Cov}\left(B_{\mathrm{GGG}}^{(ijk)}(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3),\ B_{\mathrm{GGG}}^{(lmn)}(\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6)\right) = \frac{(2\pi)^3}{A\,\bar{\ell}_1\bar{\ell}_2\bar{\ell}_3\,\Delta\bar{\ell}_1\Delta\bar{\ell}_2\Delta\bar{\ell}_3}\,\Lambda^{-1}\left(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3\right) \times\left(\bar{P}^{(il)}(\bar{\ell}_1)\bar{P}^{(jm)}(\bar{\ell}_2)\bar{P}^{(kn)}(\bar{\ell}_3)\,\delta_{\bar{\ell}_1\bar{\ell}_4}\,\delta_{\bar{\ell}_2\bar{\ell}_5}\,\delta_{\bar{\ell}_3\bar{\ell}_6}+5\,\text{perms.}\right), \tag{6.25}$$

in which The term $\Lambda(\ell_1,\ell_2,\ell_3)$ is defined in (5.5), and $\bar{P}^{(ij)}(\bar{\ell})$ is the observed tomographic power spectrum which contains the intrinsic ellipticity noise (e.g. Kaiser 1992; Hu 1999; Joachimi et al. 2008):

$$\bar{P}^{(ij)}(\bar{\ell}) = P^{(ij)}(\bar{\ell}) + \delta_{ij}\frac{\sigma_\epsilon^2}{2\bar{n}_i}\,, \tag{6.26}$$

where $\bar{n}_i$ is the galaxy number density in redshift bin i. We use the Eisenstein & Hu (1998) transfer function to evaluate the linear 3D matter power spectrum, and the Smith et al. (2003) fitting function for the nonlinear power spectrum.

6.3.3 Toy intrinsic-shear alignment model

In this section we present a toy model for generating GGI systematics. Since the physical generation of intrinsic-shear alignments concerns nonlinear growth of structure and complex astrophysical processes which are not easy to quantify, a realistic model is not yet available. Current simulations involving baryonic matter also have some way to go before they can simulate the generation of the GGI systematics reliably.

Up to now there has not been any attempt to measure GGI and GII in galaxy surveys. Semboloni et al. (2008) studied these systematics using ray-tracing simulations. They provided fits in real space to projected GII and GGI signals, but the results are still too crude to lead to sufficient constraints on an intrinsic-shear alignment model.

This situation emphasizes the importance of a method intended to control intrinsic-shear alignment to be model-independent, especially at the 3-pt level. Since this is the case for the nulling technique, for this work we only require a simple model for $B_{\mathrm{GGI}}^{(ijk)}$ which satisfies the characteristic redshift dependence and leads to a reasonable bias.

Based on the observation that the lensing bispectrum expression (6.24) comes directly from (6.13) and the definition of the tomography bispectrum (6.15), we link $B_{\mathrm{GGI}}^{(ijk)}$ also to a 3D bispectrum $B_{\delta_I\delta\delta}$ via

$$B_{\mathrm{GGI}}^{(ijk)}(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3) = \int_0^{\chi_{\mathrm{hor}}} d\chi\ \frac{p_{\mathrm{s}}^{(i)}(\chi)W^{(j)}(\chi)W^{(k)}(\chi)}{\chi^4}\,B_{\delta_I\delta\delta}\left(\frac{\bar{\ell}_1}{\chi},\frac{\bar{\ell}_2}{\chi},\frac{\bar{\ell}_3}{\chi};\chi\right). \tag{6.27}$$

Similar to $B_\delta(k_1,k_2,k_3)$ which is given by

$$\left\langle\tilde{\delta}(\boldsymbol{k_1},\chi)\tilde{\delta}(\boldsymbol{k_2},\chi)\tilde{\delta}(\boldsymbol{k_3},\chi)\right\rangle = (2\pi)^3\,\delta_D(\boldsymbol{k_1}+\boldsymbol{k_2}+\boldsymbol{k_3})\,B_\delta(k_1,k_2,k_3;\chi), \tag{6.28}$$

$B_{\delta_I\delta\delta}$ is defined via

$$\left\langle\tilde{\delta}_I(\boldsymbol{k_1},\chi)\tilde{\delta}(\boldsymbol{k_2},\chi)\tilde{\delta}(\boldsymbol{k_3},\chi)\right\rangle = (2\pi)^3\,\delta_D(\boldsymbol{k_1}+\boldsymbol{k_2}+\boldsymbol{k_3})\,B_{\delta_I\delta\delta}(k_1,k_2,k_3;\chi), \tag{6.29}$$

where $\tilde{\delta}_{\mathrm{I}}(\boldsymbol{k})$ is the 3D density field which is responsible for the intrinsic alignment, and it satisfies

$$\tilde{\kappa}_{\mathrm{I}}^{(i)}(\boldsymbol{\ell}) = \int_0^{\chi_{\mathrm{hor}}} d\chi\ p_{\mathrm{s}}^{(i)}(\chi)\,\tilde{\delta}_{\mathrm{I}}\left(\frac{\boldsymbol{\ell}}{\chi},\chi\right)\,, \tag{6.30}$$

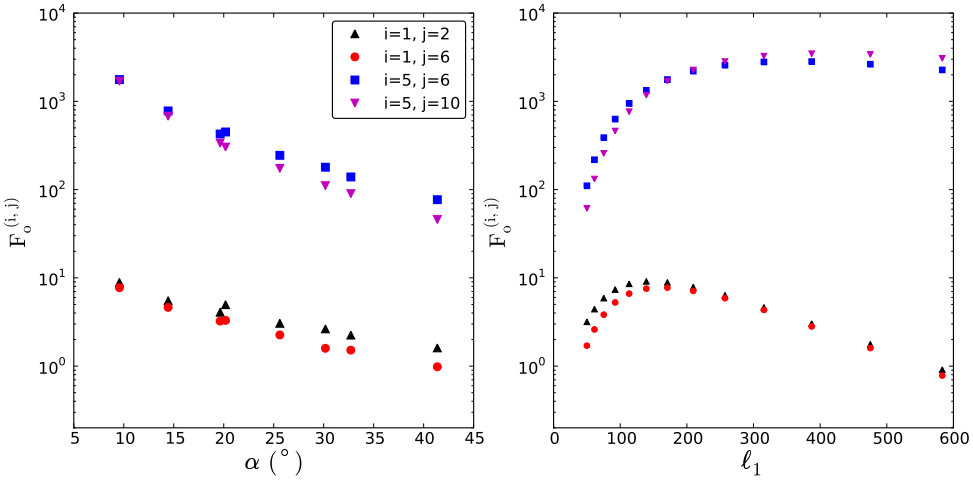


Figure 6.1: Distribution of the nulled Fisher information as defined in (6.40) per $(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3)$ bin and per redshift bin combination among different angular frequency triangle shapes and sizes. Results for four redshift bin combinations (i, j) are presented. *Left panel*: Distribution of the nulled Fisher information among different triangle configurations. We consider triangles with the common shortest side length $\bar{\ell}_1 = 171$ which corresponds to the 7th angular frequency bin. Due to our logarithmic binning and the constraint that the three side lengths must be able to form a triangle, only 8 such triangle configurations exist. Plotted is the nulled Fisher information contained in these 8 triangles against α, which is the angle opposite to the shortest side length in that triangle. Smaller α correspond to more elongated triangles, and larger α correspond to almost equilateral triangles. *Right panel*: Distribution of the nulled Fisher information contributed by each $(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3)$ bin over different triangle sizes. A fixed triangle shape with $\bar{\ell}_1 : \bar{\ell}_2 : \bar{\ell}_3 = 1 : 3.64 : 4.52$ (corresponds to the leftmost points in the left panel) is chosen. The nulled Fisher information contained in one $(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3)$ bin is plotted against the shortest side length $\bar{\ell}_1$ of each triangle.

The definition of both $\tilde{\kappa}_{\rm I}^{(i)}$ and $\tilde{\delta}_{\rm I}$ originates from the deterministic part of galaxy intrinsic ellipticity $\epsilon_{\rm I}^{\rm det}$. We have assumed the existence of these underlying smooth fields. Similar quantities have been defined in Joachimi & Bridle (2010), see also Hirata & Seljak (2004) and Schneider & Bridle (2010). We would like to point out again that, although we introduce these quantities for the clarity of our model, we do not need them for the main purpose of this study. What we need to model is the projected GGI bispectrum $B_{\rm GGI}^{(ijk)}$.

Note that in (6.27), the weight for the lowest redshift bin i is the source redshift distribution function $p_{\rm s}^{(i)}$ which is zero outside redshift bin i, rather than the lensing weight $W^{(i)}$ which is a much broader function. Since $\tilde{\kappa}_{\rm I}^{(i)}$ depends only on physical processes at redshift bin i and is inferred from ellipticity measurements in this bin, and $\tilde{\kappa}_{\rm G}^{(j)}$ is linked to the 3D matter density through the lensing weight $W^{(j)}$, this assignment of weight functions will ensure the correct redshift dependence of $B_{\rm GGI}^{(ijk)}$.

When the redshift bins are not disjunct, however, the intrinsic alignment signal can no longer be associated with bin i. There will be two permutations in both the left-hand side of (6.17) and the right-hand side of (6.27), similar to the 2-pt case, e.g. Eq. 11 in Hirata & Seljak (2004).

The modeling of $B_{\delta_I\delta\delta}$ is then a pure matter of choice. We build a simple 3D GGI bispectrum

with power-law dependence on both redshift z and spatial frequency k:

$$B_{\delta_I\delta\delta}(k_1,k_2,k_3;\chi) := -\mathcal{A}\,B_\delta(k_{\rm ref},k_{\rm ref},k_{\rm ref};\chi(z_{\rm med}))\left(\frac{1+z}{1+z_{\rm med}}\right)^{r-2} \times\left\{\left(\frac{k_1}{k_{\rm ref}}\right)^{2(s-2)}+\left(\frac{k_2}{k_{\rm ref}}\right)^{2(s-2)}+\left(\frac{k_3}{k_{\rm ref}}\right)^{2(s-2)}\right\}, \tag{6.31}$$

where $z_{\rm med}$ is the median redshift of the whole survey, and $\mathcal{A}$, $k_{\rm ref}$, r, s are free parameters. Among them the parameter $k_{\rm ref}$ is designed to be a characteristic wave number, whose value we set to be a weakly nonlinear scale of 10 hMpc^{-1} here. The minus sign ensures that the contamination of GGI systematics leads to an underestimation of the GGG signal, as found by Semboloni et al. (2008).

Little is known about the redshift and angular scale dependence of $B_{\delta_I\delta\delta}$. However one can roughly estimate how it compares to the $B_{\delta\delta\delta}$ signal. A linear alignment model suggests $\delta_I \propto \delta_{\rm lin}\overline{\rho}(z)/[(1+z)\,D_+(z)]$ (see e.g. Hirata & Seljak 2004), in which $\overline{\rho}(z)$ is the mean density of the Universe, $D_+(z)$ is the growth factor, and $\delta_{\rm lin}$ is the linear matter density contrast. Thus we have, very roughly, $\delta_I \propto (1+z)^3\,\delta_{\rm lin}$ which suggests $B_{\delta_I\delta\delta} \propto (1+z)^3\,B_{\delta\delta\delta}$. The linear alignment model assumes that the intrinsic alignment is linearly related to the local tidal gravitational field (e.g. Catelan et al. 2001; Hirata & Seljak 2004). If this holds true, we also expect $B_{\delta_I\delta\delta}$ to have a stronger angular scale dependence than $B_{\delta\delta\delta}$ since tidal gravitational interaction follows the inverse cube law rather than the inverse square law which gravity itself follows. For a ΛCDM model, in the weakly nonlinear regime where perturbation theory holds, the dependence of $B_{\delta\delta\delta}$ on $(1+z)$ has a negative power shallower than -4, and the dependence on k has a power of around -2. In this study we choose $r=0$, $s=1$ as default. We also study the cases of $r=-2$, $r=2$, and $s=0$ whose results will be shown in Fig. 6.4 below.

As for the amplitude $\mathcal{A}$ of the GGI signal, the only direct study up to now is Semboloni et al. (2008), which suggests an overall GGI/GGG ratio of 10 % for a $z_{\rm m}=0.7$ survey for elliptical galaxies and few percent for a mixed sample of elliptical and spiral galaxies. In this study we adjust $\mathcal{A}$ such that the amplitude of the tomographic GGI bispectrum is limited to be within 10 % of the amplitude of the lensing GGG signal, i.e. GGI/GGG$\lesssim$ 10 % at redshift bin combinations with $z_i \ll z_j$ and $z_i \ll z_k$ where the GGI signal is expected to be most significant. This leads to a relatively modest overall GGI/GGG ratio at percent level. We will show examples of the generated GGI and GGG signals in Fig. 6.3. As an order-of-magnitude estimate, one can also relate the GGI/GGG ratio to that of GI/GG by expanding 3-pt signals to couples of 2-pt signals using perturbation theory, in analogy to the Scoccimarro & Couchman (2001) fitting formula. For the case of $z_i \ll z_j \approx z_k$, the leading order terms would give that the GGI/GGG ratio approximates that of GI/GG evaluated at redshifts z_i and z_j. This suggests that our adopted GGI/GGG ratio is also consistent with available observational studies of the GI signal (Mandelbaum et al. 2006, 2011; Hirata et al. 2007; Fu et al. 2008; Okumura et al. 2009; Okumura & Jing 2009), although the results of these studies vary a lot according to different median redshift, color and luminosity of the selected galaxy sample.

6.4 Construction of nulling weights

As mentioned in Sect. 6.2.2, we would like to construct a single first-order weight function $T^{(ij)}(\chi)$ for each (i,j) combination which preserves the maximum of information. This can be seen as a constrained optimization problem. The constraining condition here is the nulling condition and the quantity to be optimized is the Fisher information after nulling. In JS08, several practical methods

were developed to solve this optimization problem at the 2-pt level, and very good agreement was found among the different methods.

We adopt the simplified analytical approach as described in JS08, and reformulate it for 3-pt statistics here. For convenience we introduce the following notations:
the bispectrum covariance matrix **C**ovB, whose elements are

$$\mathrm{CovB}(^{ijk}_{\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3};^{lmn}_{\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6}) := \mathrm{Cov}\left(B^{(ijk)}_{\mathrm{GGG}}(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3),\ B^{(lmn)}_{\mathrm{GGG}}(\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6)\right)\ ; \tag{6.32}$$

the covariance matrix **C**ovY of the nulled bispectra Y, whose elements are

$$\begin{aligned}\mathrm{CovY}(^{ij}_{\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3};^{lm}_{\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6}) &:= \mathrm{Cov}\left(Y^{(ij)}(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3), Y^{(lm)}(\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6)\right)\\ &= \sum_{\substack{k=\\ i+1}}^{N_z}\sum_{\substack{n=\\ l+1}}^{N_z} \mathrm{Cov}\left(B^{(ijk)}_{\mathrm{GGG}}(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3), B^{(lmn)}_{\mathrm{GGG}}(\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6)\right) T^{(ij)}(\chi_k)\, T^{(lm)}(\chi_n)\chi'_k\,\chi'_n\,\Delta z_k\,\Delta z_n\ ;\end{aligned} \tag{6.33}$$

a vector $\mathbf{B}_{,\mu}$ whose elements are partial derivatives of the bispectrum with respect to the cosmological parameter p_μ

$$B_{,\mu}(^{ijk}_{\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3}) := \frac{\partial B^{(ijk)}_{\mathrm{GGG}}(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3)}{\partial p_\mu}\ ; \tag{6.34}$$

and a corresponding vector $\mathbf{Y}_{,\mu}$ for nulled bispectra Y, whose elements are

$$Y_{,\mu}(^{ij}_{\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3}) := \frac{\partial Y^{(ij)}(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3)}{\partial p_\mu}\ . \tag{6.35}$$

Then the Fisher information matrix from the original bispectra can be written as (following TJ04)

$$\mathrm{F}^{\mathrm{i}}_{\mu\nu} = \mathbf{B}_{,\mu}\,\mathbf{C}\mathrm{ovB}^{-1}\,\mathbf{B}_{,\nu}\ , \tag{6.36}$$

and that from the nulled bispectra can be written as

$$\mathrm{F}^{\mathrm{f}}_{\mu\nu} = \mathbf{Y}_{,\mu}\,\mathbf{C}\mathrm{ovY}^{-1}\,\mathbf{Y}_{,\nu}\ . \tag{6.37}$$

Here the matrix multiplication is a summation of possible angular frequency combinations $(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3)$ and redshift bin combinations, (ijk) for the original bispectra and (ij) for the nulled bispectra. In (6.36) and (6.37), $\mathbf{C}\mathrm{ovB}^{-1}$ and $\mathbf{C}\mathrm{ovY}^{-1}$ indicate the inverse of the covariance matrix. When the covariance is approximated by triples of power spectra, the covariance between two different angular frequency combinations $(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3) \neq (\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6)$ is zero, see (6.25), which means that the covariance matrix is block diagonal. In this case the matrix inversion can be done separately for each block specified by an angular frequency combination $(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3)$.

According to the idea of the simplified analytical approach, we consider the Fisher information on one cosmological parameter contained in bispectrum measures $B^{(ijk)}_{\mathrm{GGG}}(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3)$ with a single $(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3)$ combination and with redshift bin (i,j,k) combinations having common (i,j) indices. For every (i,j) combination we build nulling weights $T^{(ij)}$ which maximizes the nulled Fisher matrix using the method of Lagrange multipliers. Since here the nulled Fisher matrix receives contribution only from certain angular frequency and redshift combinations, we denote it as $F^{(ij)}_{\mathrm{o}}$ to avoid ambiguity. $F^{(ij)}_{\mathrm{o}}$ has only one component since only one cosmological parameter is taken into consideration. As only a single $(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3)$ combination is involved, we will omit the $\bar{\ell}$-dependence in all variables in the rest of this subsection to keep a compact form.

Again for notational simplicity, we follow JS08 and introduce a vector notation as follows. For each (i, j) in consideration, let the values of the weights $T^{(ij)}(\chi_k)$ form a vector $\boldsymbol{T} = T_k$, and define another vector $\boldsymbol{\rho}$ and a matrix $\bar{\mathbf{C}}$ with elements

$$\rho_k := B_{,\mu}(^{ijk})\,\chi_k'\,\Delta z_k\,, \tag{6.38}$$

$$\bar{C}_{kn} := \mathrm{CovB}(^{ijk};{}^{ijn})\,\chi_k'\,\chi_n'\,\Delta z_k\,\Delta z_n\,. \tag{6.39}$$

Thus $F_{\mathrm{o}}^{(ij)}$ can be expressed, according to (6.37), as

$$F_{\mathrm{o}}^{(ij)} := \mathbf{Y}_{,\mu}(^{ij})\ \mathbf{C}\mathrm{ovY}^{-1}(^{ij};{}^{ij})\ \mathbf{Y}_{,\mu}(^{ij}) = \frac{(\boldsymbol{T}\cdot\boldsymbol{\rho})^2}{\boldsymbol{T}^\tau\bar{\mathbf{C}}\boldsymbol{T}}. \tag{6.40}$$

We further define a vector $\boldsymbol{f}$ with elements

$$f_k = \left(1 - \frac{\chi_i}{\chi_k}\right)\chi_k'\,\Delta z_k \tag{6.41}$$

to write the nulling condition (6.19) as

$$O^{(ij)} = \boldsymbol{T}\cdot\boldsymbol{f} = 0\,. \tag{6.42}$$

The problem of finding nulling weights $\boldsymbol{T}$ which maximize $F_{\mathrm{o}}^{(ij)}$ under the constraint given by the nulling condition can be solved with the method of Lagrange multipliers by defining a function

$$G := F_{\mathrm{o}}^{(ij)} + \lambda O^{(ij)} = \frac{(\boldsymbol{T}\cdot\boldsymbol{\rho})^2}{\boldsymbol{T}^\tau\bar{\mathbf{C}}\boldsymbol{T}} + \lambda\boldsymbol{T}\cdot\boldsymbol{f} \tag{6.43}$$

with λ being the Lagrange multiplier, and setting the gradient of G with respect to $\boldsymbol{T}$ to zero,

$$\nabla_T\, G = 2\boldsymbol{\rho}\,\frac{(\boldsymbol{T}\cdot\boldsymbol{\rho})}{\boldsymbol{T}^\tau\bar{\mathbf{C}}\boldsymbol{T}} - 2\,\bar{\mathbf{C}}\boldsymbol{T}\left(\frac{(\boldsymbol{T}\cdot\boldsymbol{\rho})}{\boldsymbol{T}^\tau\bar{\mathbf{C}}\boldsymbol{T}}\right)^2 + \lambda\boldsymbol{f} = 0\,. \tag{6.44}$$

The solution to this equation is (for more details see JS08)

$$\boldsymbol{T} = \mathcal{N}\left\{\bar{\mathbf{C}}^{-1}\boldsymbol{\rho} - \frac{\boldsymbol{f}^\tau\bar{\mathbf{C}}^{-1}\boldsymbol{\rho}}{\boldsymbol{f}^\tau\bar{\mathbf{C}}^{-1}\boldsymbol{f}}\,\bar{\mathbf{C}}^{-1}\boldsymbol{f}\right\}\,, \tag{6.45}$$

with the normalization $\mathcal{N}$ adjusted to give $|T|^2 = 1$.

Apparently the thus constructed nulling weights depend on which $(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3)$ combination is considered and with respect to which cosmological parameter we optimize the information content. In this study the default cosmological parameter to optimize is Ω_{m}, and we choose for each (i, j) combination the $(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3)$ combination which maximizes $F_{\mathrm{o}}^{(ij)}$. However one needs to be aware that this serves only as a clear choice of a $(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3)$ combination and is not necessarily the best in terms of information preservation considering all angular frequency bins and all cosmological parameters.

To show which triangle shapes and sizes contain more information, we plot $F_{\mathrm{o}}^{(ij)}$ against the $(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3)$ triangle shape and size for four typical (i, j) combinations in Fig. 6.1. In the left panel, the nulled information $F_{\mathrm{o}}^{(ij)}$ contained in different triangles with a common shortest side length $\bar{\ell}_1 = 171$ is plotted against α, which is the angle opposite to $\bar{\ell}_1$. Due to our logarithmic binning in angular frequency, only eight $(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3)$ combinations with $\bar{\ell}_1 = 171$ can form triangles. One sees that the more elongated triangles (small α) contain much more Fisher information than the almost equilateral

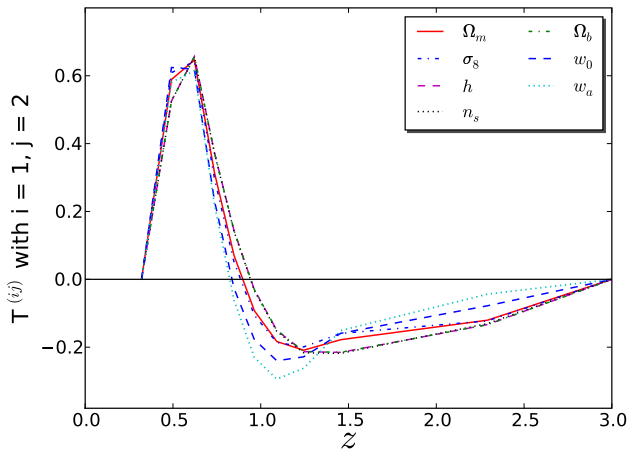


Figure 6.2: Nulling weights $T^{(ij)}$ for redshift bins $i = 1$, $j = 2$ are plotted against the redshift value of the third redshift index k. Remarkable consistency is found between nulling weights optimized on different parameters, shown with different line styles.

triangles (large α). The small separation between the 3rd and the 4th points from the left is caused by the degeneracy of different triangle shapes with respect to α, e.g. two equal and very long side lengths can result in the same value of α as two shorter side lengths with a length difference close to the length of the shortest side length. The right panel shows the distribution of the Fisher information contained in one $(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3)$ bin over the triangle size. When the redshift in consideration is higher, the peak of the information distribution moves to higher angular frequencies. The figure suggests that most information comes from high redshifts and small angular scales.

To explore the sensitivity of nulling weights on the choice of the cosmological parameter, we construct seven sets of weight functions, each optimizing the information content in terms of one parameter. For all (i, j) combinations we find that the nulling weights are not very sensitive to the choice of parameter. As an example, the weights for $(i, j) = (1, 2)$ are shown in Fig. 6.2. This result is rather surprising at first sight, since for different parameters the distribution of information (contained in the bispectrum) over redshift bins is quite different. However, such insensitivity suggests that the shapes of nulling weights are already strongly constrained under our construction scheme. One constraint is, evidently, the nulling condition. Moreover, considering the fact that we optimize the nulling weights for each (i, j) combination with respect to the information content they preserve, we have already required the shapes of these first order nulling weights to be as smooth as possible.

The fact that these two conditions have already imposed strong constraints on the nulling weights also suggests that nulling weights can be robustly and efficiently constructed, i.e. it is not critical to construct the "best" nulling weights.

6.5 Performance of the nulling technique

6.5.1 GGI/GGG ratio

What the nulling technique "nulls" is the GGI signal $B_{\rm GGI}$, so the GGI/GGG ratio is the most direct quantification of its performance. We plot the modeled GGI and GGG bispectra before and after nulling in Fig. 6.3. The original GGI signal is shown in the left panels by dashed lines. For

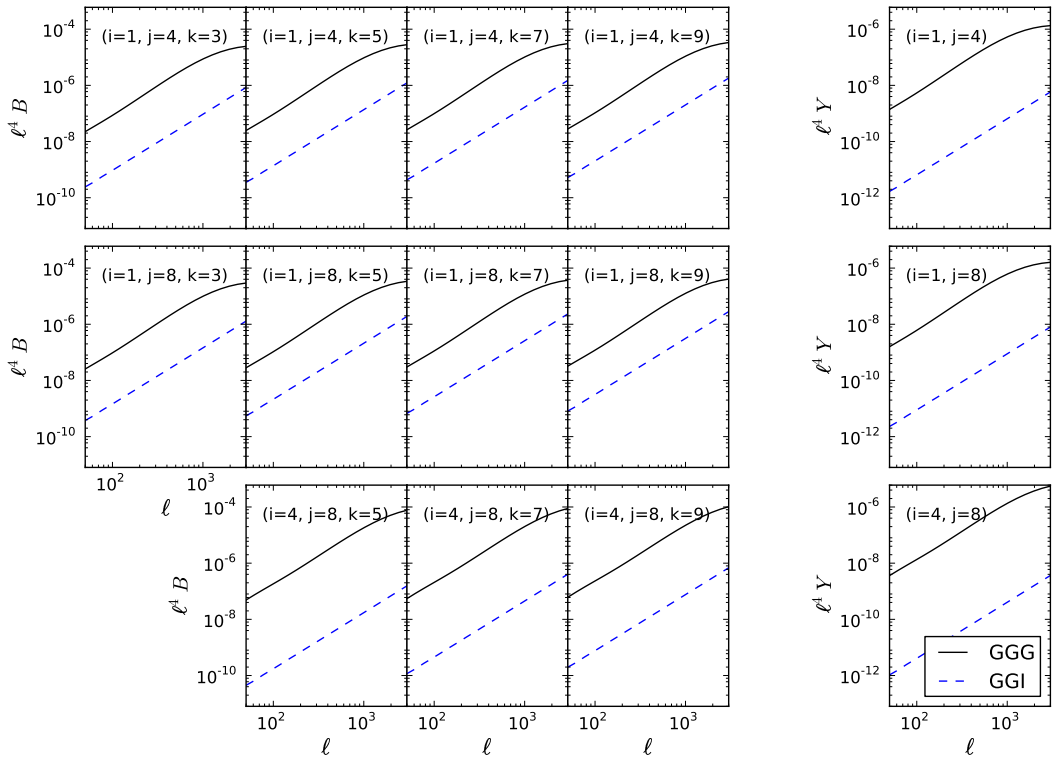


Figure 6.3: Tomographic convergence bispectrum (GGG, solid curves) and intrinsic-shear alignment (GGI, dashed curves) for equilateral triangles are plotted against triangle side length. Measures both before (left panel) and after (right panel) applying the nulling technique (B and Y respectively) are shown for three typical redshift bin (i, j) combinations in the three rows.

comparison the GGG signals are shown as solid curves. The results are shown for equilateral triangle configurations for the convenience of presenting. One sees that when the redshift bin number j and/or k increase, the changes in GGG and GGI signals are different, which shows the expected different redshift dependence. For all redshift bin combinations the GGI signal is modeled to be subdominant to the GGG signal. In the nulled measures shown in the right panels, the GGI/GGG ratio is suppressed by a factor of 10 over all angular scales, which reflects the success of the nulling technique.

6.5.2 Information loss and downweighting of systematics

We further evaluate the performance of the nulling technique by looking at the constraining power of cosmic shear bispectrum tomography on cosmological parameters, as well as the biases caused by the GGI systematics before and after nulling.

The full characterization of the bispectrum involves three angular frequency vectors which form a triangle. In some works concerning 3-pt statistics, only equilateral triangle configurations i.e. $\ell_1 = \ell_2 = \ell_3 = \ell$ are used for simplicity reasons (e.g. Pires et al. 2009). But as several authors have pointed out (e.g. Kilbinger & Schneider 2005; Bergé et al. 2010), only a low percentage of information is contained in equilateral triangles. Thus, to calculate the full information content,

we use general triangle configurations but limit our calculation to triangles with three different side lengths, again for reasons of simplicity (for details see Sect. 6.7).

We will use the figure of merit (FoM, Albrecht et al. 2006) to quantify the goodness of parameter constraints. Here the FoM for constraints in the parameter plane $p_\alpha - p_\beta$ is defined to be proportional to the inverse of the area of the parameter constraint ellipses:

$$\mathrm{FoM}(\mathrm{p}_\alpha, \mathrm{p}_\beta) \equiv \left((\mathbf{F}^{-1})_{\alpha\alpha}(\mathbf{F}^{-1})_{\beta\beta} - (\mathbf{F}^{-1})^2_{\alpha\beta}\right)^{-\frac{1}{2}} . \tag{6.46}$$

To compute biases, we adopt a method based on a simple extension of the Fisher matrix formalism (e.g. Huterer et al. 2006; Amara & Réfrégier 2008). Then one needs to define a bias vector $\mathbb{B}^{\mathrm{GGI}}$ which in our case reads:

$$\mathbb{B}^{\mathrm{GGI}}_{\nu,\mathrm{i}} = \mathbf{B}_{\mathrm{GGI}}\, \mathbf{C}\mathrm{ovB}^{-1}\, \mathbf{B}_{,\nu} \ , \tag{6.47}$$

$$\mathbb{B}^{\mathrm{GGI}}_{\nu,\mathrm{f}} = \mathbf{Y}_{\mathrm{GGI}}\, \mathbf{C}\mathrm{ovY}^{-1}\, \mathbf{Y}_{,\nu} \ , \tag{6.48}$$

with

$$\mathbf{B}_{\mathrm{GGI}}(^{ijk}_{\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3}) := B^{(ijk)}_{\mathrm{GGI}}(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3) \ , \tag{6.49}$$

$$\mathbf{Y}_{\mathrm{GGI}}(^{ij}_{\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3}) := Y^{(ij)}_{\mathrm{GGI}}(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3) \ . \tag{6.50}$$

The bias of the parameter estimator $\hat{p}_\mu$ is given by the difference between its ensemble average and the fiducial value of the parameter p^{fid}_μ:

$$b_\mu = \langle \hat{p}_\mu \rangle - p^{\mathrm{fid}}_\mu = \sum_\nu \left(\mathbf{F}^{-1}\right)_{\mu\nu} \mathbb{B}^{\mathrm{GGI}}_\nu \ . \tag{6.51}$$

The information content before and after nulling can be seen in Fig. 6.6. On the cost of increasing the error on each parameter to about twice its original value, GGI systematics are reduced to be within the original statistical error. The relative information loss in terms of FoM can be found in Table 6.1. The constraints shown in Fig. 6.6 do not represent the best constraints obtainable from a cosmic shear bispectrum analysis since we consider only the triangles with angular scale $\bar{\ell}_1 \neq \bar{\ell}_2 \neq \bar{\ell}_3$. Also note that the nulling technique can in principle remove the GGI systematics completely. But as shown in Fig. 6.6, the systematics still cause some residual biases on cosmological parameters after nulling, due to the finite number of redshift bins. The GGI systematics will be reduced to a lower level when more redshift bins are available. We will discuss this further in the following subsection.

6.5.3 How many redshift bins are needed?

Analyzing the cosmic shear signal in a tomographic way was originally meant to maximize the information. For this purpose alone, a crude redshift binning will suffice (Hu 1999). However, to control intrinsic-shear alignment, which is a redshift-dependent effect, much more detailed redshift information is required (e.g. King & Schneider 2002; Bridle & King 2007; Joachimi & Schneider 2008). Thus, for a method intended to eliminate intrinsic-shear alignment, it is necessary to show its requirement on the redshift precision. In the case of nulling, detailed redshift information is not only needed for the method to be able to eliminate the bias, but also for the preservation of a reasonably

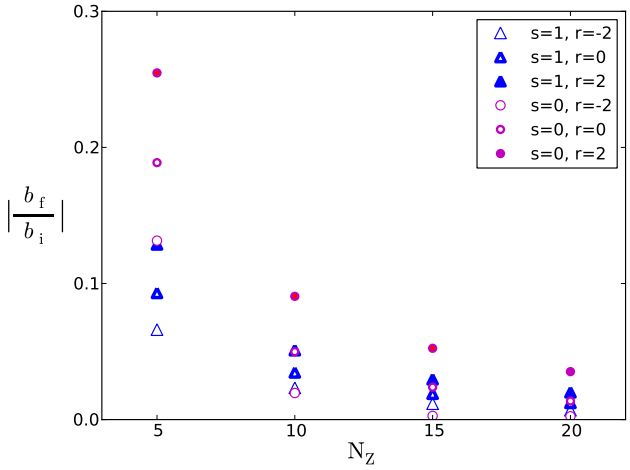


Figure 6.4: Ratio of the nulled and the original biases for cosmological parameter Ω_m as a function of number of redshift bins N_z. Results for different GGI models are shown. Parameter s and r are the slopes of angular frequency and redshift dependence of our power-law model (6.31).

large amount of information through the nulling process. JS08 examined the number of redshift bins required for the nulling technique in the 2-pt case, and showed that 10 redshift bins already ensure that parameters are still well-constrained after nulling.

To re-assess this problem at the 3-pt level, we consider two different situations to address the requirements coming from control of the intrinsic-shear alignment and preservation of the information content separately. In both cases we split the redshift range between $z = 0$ and $z = 3$ into 5, 10, 15, and 20 (only in the first situation) redshift bins, with the redshift bins split in a way that there is an equal number of galaxies in each bin.

First we consider a single cosmological parameter, Ω_m, to be free and study the biases introduced by the GGI signal on Ω_m both before and after nulling. We use only equilateral triangle configurations to reduce the amount of calculation. The results are shown in Fig. 6.4. Within the range of consideration, the ratio of the nulled and the original biases drops quickly with the increase of the number of redshift bins for all GGI models. For most of the models, 5 redshift bins seem to be not sufficient for the nulling technique to control the bias induced by GGI down to a percent level. Going from 5 redshift bins to 10 redshift bins is very rewarding in terms of bias reduction. However, we note that a decrease $|b_f/b_i|$ doesn't necessarily indicate a better performance of the nulling method, or generally speaking, of any method intended to control the intrinsic-shear alignment. One can see the reason for this by noticing that, it is the original unbinned GGI/GGG signal that is directly controlled by any of these methods. Between $|b_f/b_i|$ and the original unbinned GGI/GGG signal lies the binning process as well as the summation over angular frequency bins and redshift bins. Since the signs of the biases contributed by different angular frequencies and redshifts can be different, there can be bias cancellation during these processes. In another word, $|b_f/b_i|$ can depend on binning choices.

We then vary two cosmological parameters (Ω_m and σ_8) and investigate how the original and the nulled parameter constraints change with respect to the number of redshift bins available. For this case we use all triangle shapes to enable a comparison with results for 2-pt statistics.

Our result (Fig. 6.5) shows that a further increase of the number of redshift bins beyond 10 is not

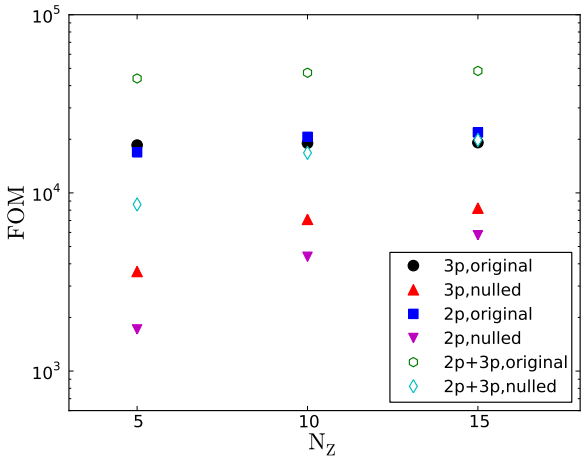


Figure 6.5: Figure of merit (FoM) as defined in (6.46) in the Ω_m - σ_8 plane as a function of number of redshift bins N_z. FoM from 2-pt (2p) measures, 3-pt measures (3p) and combined (2p+3p) are shown both before nulling (original) and after nulling (nulled).

very rewarding in terms of information preservation as characterized by the FoM, in either 2p, 3p, or 2p+3p cases. This suggests, when the possibility of more redshift bins exists, the choice of redshift bin number should be based mainly on the requirement of bias reduction level in case of negligible photometric errors. When there are non-negligible photometric errors, however, the information loss will probably be more severe, as found by Joachimi & Schneider (2009) for the 2-pt case.

6.5.4 The nulling technique as a conditioned compression of data

The necessity of carrying out data compression in cosmology has long been recognized (e.g. Tegmark et al. 1997) and has been ever increasing due to the increasing size of the data sets. In cosmic shear studies the survey area of next generation multicolor imaging surveys will be an order of magnitude larger than the current ones. The study of 3-pt statistics also implies a huge increase in the amount of data directly entering the Fisher-matrix/likelihood analysis, compared to the 2-pt case.

The basic principle of data compression is to reduce the amount of data while preserving most of the information. This is already naturally encoded in the nulling technique. If one keeps only the first-order weights for nulling, as we do in this study, the nulling procedure reduces the number of data entries in each angular frequency bin from the number of redshift bin (i, j, k) combinations, to the number of (i, j) combinations, which means roughly from ${N_z}^3$ to ${N_z}^2$. The nulling transformation is linear since the resulting nulled entry is a linear combination of k original entries weighted by the nulling weight (6.18). In the sense that an "optimum" set of nulling weights is constructed, the nulling technique also intends to preserve as much information as possible. But there is yet another additional constraining condition in the nulling procedure: the nulling condition (6.8), which largely confines the shape of the nulling weights by requiring the existence of at least one zero-crossing (see Fig. 6.2). In short, the nulling technique can be seen as a conditioned linear compression of data.

It is then interesting to know how much of the information loss during the nulling process actually comes from the nulling condition, and how much just comes from the fact that a data compression process is naturally involved in nulling. To explore this, we perform an unconditioned linear

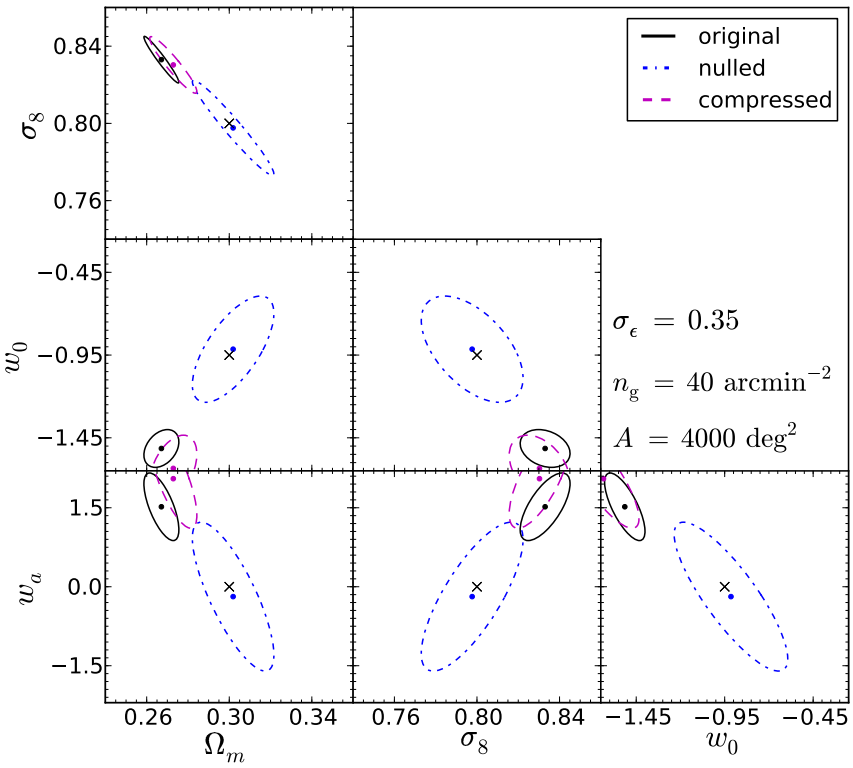


Figure 6.6: Projected 1-sigma (68 % CL) parameter constraints from cosmic shear bispectrum tomography. Hidden parameters are marginalized over. The black solid and blue dash-dotted ellipses correspond to the original constraints and those after nulling, respectively. The black cross in the center of each panel represents the fiducial values adopted for the parameters, and the distance from the center of one ellipse to the black cross reflects the bias caused by intrinsic-alignment GGI systematics on the corresponding parameter. As nulling can be seen as a linear data compression under the constraint of the nulling condition, we also plot the constraints and biases after an unconditioned linear data compression as magenta dashed ellipses for comparison (see Sect. 6.5.4).

data compression, by simply ignoring the nulling condition in the whole nulling procedure i.e. dropping the Lagrange multiplier term in (6.43), but otherwise keeping the simplifications inherent to the analytical approach. The results are shown in Fig. 6.6. A summary of the FoM from the original and the nulled bispectrum measures as well as the compressed measures is shown in Table 6.1.

In contrast to nulling, an unconditioned linear compression does not eliminate the parameter bias, but increases or reduces some of them marginally. Regarding the parameter constraints, although the increase in the size of the ellipses is much less than in the case of nulling, around one third of the information in terms of FoM is lost through compression, which means that the amount of degradation in parameter constraints after compression is not negligible. This suggests that keeping only the first-order terms contributes to non-negligible information loss. To regain part of this information, one could add higher-order weights to the nulling procedure. But the difference between the nulled and the compressed FoM serves as an indication for the inevitable information loss through the nulling process, which is imposed by the nulling condition.

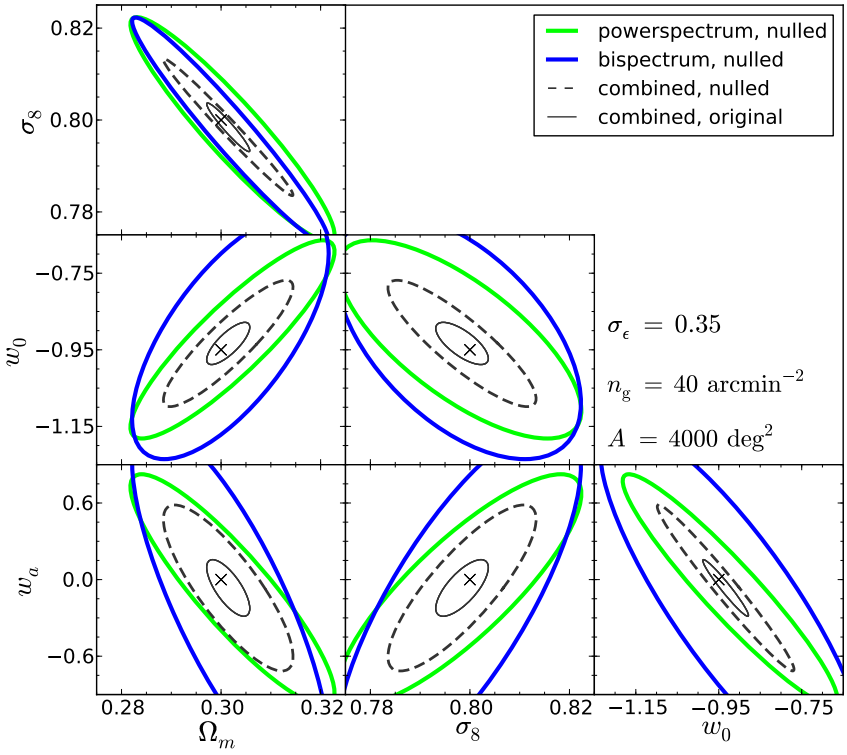


Figure 6.7: The thick green (gray) solid, thick blue (black) solid and thin black dashed ellipses indicate 1-sigma (68 % CL) parameter constraints from the nulled power spectrum measures, bispectrum measures, and combined. Hidden parameters are marginalized over. The distance from the center of an ellipse to the black cross reflects the nulled bias on the corresponding parameter. The original biases from bispectrum measures can be seen in Fig.6.6. The thin black solid ellipses over-plotted on to the centers of the nulled combined constraint ellipses indicate the statistical power (68% CL) of combined constraints before nulling. Note the different ranges of parameters compared to Fig.6.6.

6.5.5 Two-point and three-pt constraints combined

Besides constraining cosmological parameters using 3-pt cosmic shear alone, we investigate the combined constraints from both 2-pt and 3-pt cosmic shear measures. The performance of the nulling technique on cosmic shear power spectrum tomography alone and the resulting constraints on cosmological parameters were presented in JS08. For consistency, we use the same setting for the cosmic shear power spectrum as described for the bispectrum in Sect. 6.3. In particular, we neglect photometric redshift errors, use only a limited range and number of ℓ-bins, and adopt a power-law intrinsic-shear alignment model with a form described by (36) in JS08 and a slope of 0.4. We have confirmed the consistency between our power spectrum and bispectrum codes with those used in Bergé et al. (2010). Our power spectrum code agrees also with iCosmo (Refregier et al. 2011).

Figure 6.7 shows the resulting constraint ellipses after nulling from the cosmic shear power spectrum analysis, the bispectrum analysis, and the two combined. To show how much information is lost during the nulling process, we overplot the original 2- and 3-pt combined constraints on top of the nulled constraint ellipses in Fig. 6.7, but center them on the corresponding nulled constraints by subtracting the bias difference before and after nulling. The information content in terms of FoM

Table 6.1: Change of cosmic shear bispectrum statistical power after nulling (null) and linear data compression (compress). Presented are FoM on 2D parameter planes between cosmological parameters Ω_m, σ_8, w_0 and w_a. The cosmological parameters h, Ω_b and n_s are marginalized over. The second column is the FoM from the original bispectrum; the third and fifth columns are FoM from the nulled and the compressed measures, respectively; the fourth (sixth) column shows the percentage of the third (fifth) column compared to the first column, which reflects the relative information loss through the nulling (the unconditioned compression) procedure.

	i	null	null/i	compress	compress/i
Ω_m-σ_8	21455	4609	21.5 %	12242	57.1 %
Ω_m-w_0	637	123	19.3 %	428	67.2 %
Ω_m-w_a	145	33	23.0 %	110	75.9 %
σ_8- w_0	434	87	20.0 %	299	68.9 %
σ_8- w_a	101	26	25.4 %	72	71.3 %
w_0- w_a	11.4	2.3	20.2 %	8.0	70.2 %

Table 6.2: FoM before ('i') and after ('f') nulling and their ratio, using the cosmic shear power spectrum (2pt), bispectrum (3pt), and combined (2pt+3pt) analysis.

	2pt, i	3pt, i	2pt+3pt, i	2pt, f	2pt, i/f	3pt, f	3pt, i/f	2pt+3pt, f	2pt+3pt, i/f
Ω_m-σ_8	21774	21455	86851	3297	15.1 %	4609	21.5 %	18555	21.4 %
Ω_m-w_0	1590	637	3806	236	14.8 %	123	19.3 %	600	15.8 %
Ω_m-w_a	517	145	872	69	13.3 %	33	23.0 %	121	13.9 %
σ_8- w_0	864	434	3832	132	15.2 %	87	20.0 %	488	17.2 %
σ_8- w_a	326	101	709	47	14.4 %	26	25.4 %	107	15.1 %
w_0- w_a	45	11	184	7.4	16.4 %	2.3	20.2 %	27	14.5 %

for each parameter pair is presented in Table 6.2.

One sees that the amount of information contained in bispectrum measures and power spectrum measures are indeed comparable. With bispectrum information added, typically three times better constraints in terms of FoM are achieved, both before and after nulling. This factor is smaller than the result in TJ04, although the same angular frequency range and the same set of 7 cosmological parameters are chosen for both studies. However a direct comparison is prohibited by different fiducial values adopted and different survey specification.

Through the nulling procedure, around 15 % of the original information in terms of FoM is preserved in the 2-pt case, and around 20 % in the 3-pt case. It is a bit higher in the 3-pt case, in accordance to the fact that a roughly ${N_z}^3 \rightarrow {N_z}^2$ compression is involved in the 3-pt case and a ${N_z}^2 \rightarrow {N_z}^1$ one in the 2-pt case, while this fact is due to the summation over one redshift bin index during the nulling procedure (the same trend is evident in Fig. 6.5). The information loss is considerable, but it is a price to pay for a model-independent method. As we have discussed in the previous subsection, the difference between the information loss through the nulling and the unconditioned compression procedures represents the inevitable loss of information through nulling. However, this difference is less than 50 % in the considered 3-pt case. The other information loss is due to the simplifications we adopted in this study, including using only the first-order weights, and discarding the measures with two or three equal redshift bins. A further detailed consideration of these aspects can regain part of the lost information. Another simplification we have made in the 3-pt case is to use only triangles with three different angular frequencies. This reduces both the

original and the nulled information contained in the 3-pt measures. However, this simplification can be easily removed with a careful distinction of all cases.

Also notice that, the dependence of number of possible bispectrum modes, i.e. triangles, on the maximum angular frequency $\ell_{\rm max}$ is roughly $\ell^3_{\rm max}$, while that of power spectrum modes is roughly $\ell^1_{\rm max}$. For this study $\ell_{\rm max} = 3000$ is chosen. If reliable information on smaller angular scales can be obtained, the 3-pt statistics will possibly give us more information than the 2-pt statistics.

6.6 Conclusion

In this study we developed a method to control the intrinsic-shear alignment in 3-pt cosmic shear statistics by generalizing the nulling technique. We showed that the generalization of the nulling technique to 3-pt statistics is quite natural, providing a model-independent method to reduce the intrinsic-shear alignment signals (GGI and GII) in comparison to the lensing GGG signal.

To test the performance of the nulling technique, we assumed a fictitious survey with a setup typical of future multicolor imaging surveys, and applied the nulling technique to the modeled bispectra with intrinsic-shear alignment contamination. The lensing bispectra (GGG) was computed based on perturbation theory, while the GGI signal was modeled by a simple power-law toy model. We focused on the reduction of the GGI contaminant, since GII can be removed simply by not considering tomographic bispectra with two or three equal redshift bins.

The reduction of the intrinsic-shear alignment contamination at the 3-pt level by the nulling technique was demonstrated both in terms of the GGI/GGG ratio, and in terms of biases on cosmological parameters in the context of an extended Fisher matrix study. In terms of the GGI/GGG ratio, a factor of 10 suppression is achieved after nulling over all angular scales. Correspondingly, the biases on cosmological parameters are reduced to be less than or comparable to the original statistical errors. We studied the performance of the nulling technique when 5, 10, 15, or 20 redshift bins are available, and found that the performance on bias reduction, rather than how much information is preserved during the nulling procedure, depends more significantly on the number of redshift bins. In case one requires better control of intrinsic-shear alignment, more detailed redshift information allowing more redshift bins is the most direct way to go.

When dealing with real data, there is one further source of complication which we did not consider in this study, that is the photometric redshift uncertainty. The photometric redshift uncertainty can be characterized by a redshift-dependent photometric redshift scatter and catastrophic outliers. Joachimi & Schneider (2009) studied the influence of photometric redshift uncertainty on the performance of the nulling technique at the 2-pt level. They found that the photometric scatter places strong bounds on the remaining power to constrain cosmological parameters after nulling. The existence of catastrophic outliers, on the other hand, can lead to an incomplete removal of the intrinsic (II, III) alignments as well as the intrinsic-shear alignments (GI, GII, GGI). However, methods to control the photometric redshift uncertainty have been proposed. For example, recent studies concerning the problem of catastrophic outliers point to the solutions of either limiting the lensing analysis to $z < 2.5$ or by conducting an additional small-scale spectroscopic survey (Sun et al. 2009; Bernstein & Huterer 2009; Bordoloi et al. 2010).

As already demonstrated by JS08 in the 2-pt case, some information loss is inherent to the nulling procedure. For the setup of this study we found that, in terms of FoM about 20 % of the original information is preserved through the nulling procedure in the 3-pt case, and 15 % in the 2-pt case. We further studied the source of such information loss by comparing the nulling technique to an unconditioned linear compression of the data, since the nulling procedure can be seen as a linear

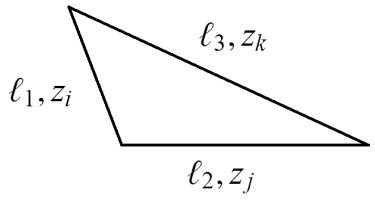
compression of data under the constraint of the nulling condition (6.8). We found that around one third of the original information is lost through an unconditioned compression of the data, suggesting that this situation can be improved by considering higher-order terms in the nulling and compression processes.

Results on parameter constraints from the 2- and 3-pt cosmic shear statistics combined are also presented. The amount of nulled information contained in bispectrum measures and power spectrum measures are comparable. With bispectrum information added, typically three-times better constraints are achieved both before and after nulling, in terms of FoM.

Again, due to the large amount of information existing in the 3-pt cosmic shear field, one would certainly like to exploit it in the future. The nulling method we developed in this work solves a potentially severe problem hampering the use of 3-pt information, namely the intrinsic-shear alignment systematic. Our method works at the cost of a large information loss, which can hopefully be avoided by a future method of removing the intrinsic-shear alignment contaminants. But as the only completely model-independent method so far, the nulling technique can serve as a working method now and can provide a valuable cross check even with the availability of better methods.

6.7 Appendix: Counting of triangles

A triangle is specified by six indices, i.e. three redshift bin indices $\{i, j, k\}$ and three angular frequency bin indices ℓ_1, ℓ_2, ℓ_3. To ensure that we count each triangle configuration only once, we set the condition that $\ell_1 \leq \ell_2 \leq \ell_3$. Moreover, we would like the first index among $\{i, j, k\}$ in (6.36) to have the lowest redshift, i.e. $z_i < z_j$ and $z_i < z_k$, for the convenience of performing the nulling technique. The possible $\{i, j, k\}$ combinations under these constraints in the case of $N_z = 4$ are listed in Fig. 6.8.



ijk	ijk	ijk
122	233	344
123	234	
124	244	
132		
133		
134		
144		

Figure 6.8: List of possible triangles (redshift bin combinations) with condition $z_i < z_j$ and $z_i < z_k$ when 4 redshift bins are available. An angular frequency combination satisfying $\ell_1 \leq \ell_2 \leq \ell_3$ is chosen. Note that the redshift indices and the angular frequencies are linked in pairs due to the definition of the tomographic bispectrum (6.15). In this study a default of 10 redshift bins is assumed.

However, setting both conditions is problematic. Inspecting the definition of the tomographic bispectrum (6.15), one sees that the redshift indices and the angular frequencies are linked in pairs, e.g. convergence κ in redshift bin i has angular frequency ℓ_1, which is not desirable since the smallest angular scale does not necessarily correspond to the lowest redshift. To solve this problem, we perform nulling three times for each general angular frequency combination with $\ell_1 < \ell_2 < \ell_3$, swapping the redshift-angular scale correspondence in-between, thus allowing each redshift to be able to correspond to any angular frequency.

Note that the situation complicates a bit when two of the angular frequencies are equal, since then the swapping may lead to exactly the same configuration. To avoid this, we will restrict ourselves to three different angular frequencies. This can exclude a high percentage of possible configurations. In our case, i.e. 20 logarithmically spaced bins between $\ell_{\mathrm{min}} = 50$ and $\ell_{\mathrm{max}} = 3000$, 37 % of the angular frequency combinations which can form a triangle have been excluded. However, this is only a technical complication which can be solved with a careful distinction of all cases. Since this study is intended to be a proof of applicability of the nulling technique to three-point statistics, we defer the intricacies of accounting for all triangle configurations to future work.

Chapter 7

Summary & Outlook

7.1 Summary

With the remarkable success of the ΛCDM model in explaining astronomical observations, it is now well-established as the standard model of cosmology. Accordingly, the focus of cosmological studies has shifted to tying up the loose ends of the ΛCDM model– exploring the nature of dark matter and dark energy which are the two exotic components assumed in it. A natural step forward is to observationally constrain their properties, which requires a combination of various observational probes. Weak gravitational lensing has emerged in the last decade as a competitive cosmological probe to this end. Especially, it is considered to be the most powerful probe in constraining the properties of dark energy when the results of forthcoming large-field imaging surveys will become available.

In this thesis we have investigated weak lensing three-point statistics, a statistical tool which will be applied to future surveys to further enhance the power of weak lensing as a cosmological probe. Three aspects of weak lensing three-point statistics have been touched, namely how the observable shear can be related to theoretical predictions of the matter density field, how much information three-point statistics can provide, and how systematical errors can be controlled. We summarize the work presented in Chaps. 4-6 in the following.

7.1.1 Relations between three-point configuration space shear and convergence statistics

In Chap. 4 we have derived some fundamental relations between weak lensing statistics, which make the theoretical framework more complete. In particular, we have related in configuration space the shear three-point functions to three-point convergence statistics which is a line-of-sight projection of the three-point statistics of the matter density field. Thereby one can compare the shear three-point functions measured from data to those constructed using the matter density field predicted by cosmological models through configuration space convergence statistics. This way of confronting observation with theory has an advantage over the current way which is based on theoretical predictions of the Fourier space quantity– the convergence bispectrum. While it is hard to precisely estimate the shear three-point functions using convergence bispectrum models due to numerical difficulties, this is not the case if one uses models of configuration space convergence statistics, as we have demonstrated using toy models.

Another major achievement of this work is the formulation of the condition for E/B-mode separation at the three-point level. A polarization field such as the cosmic shear field can be decomposed

into an E-mode and a B-mode which have distinct mathematical structures. As weak lensing signals contribute only to the E-mode while systematical errors usually do not distinguish between E-mode and B-mode, separating E- and B-modes can allow for a check of possible systematics. This is especially important for cosmic shear studies where the individual signal is small and the systematic level is possibly high. The aperture mass statistics have been used to separate E- and B-modes. However, it was found by Kilbinger & Schneider (2005) that the aperture mass statistics cannot separate them cleanly since they require shear correlation functions at small separations as an input. This motivated recent efforts in finding better statistics which allow for E/B-mode separation using only shear correlations of galaxy pairs whose separation is within a specified finite range, and such statistics have been successfully constructed at the two-point level. At the three-point level, aperture mass statistics is still the only known method to separate E- and B-modes up to now. To make the same improvements at the three-point level, the first step is to formulate the condition for E/B-mode separation, which is what we achieved in this study with the help of the relations we derived. The condition is expressed via constraints on the weight functions of shear three-point functions. Constructing weight functions satisfying these constraints and using them on shear three-point functions can lead to various three-point statistics allowing for E/B-mode separation, including the aperture mass statistics. One can then impose additional requirements on their behavior.

We have also obtained a number of by-products in this study. As an intermediate step, we have derived the relation between the two-point correlation function of the deflection potential and that of the convergence. This relation has enabled us to systematically derive two- and three-point relations between the convergence correlation function and the cross correlation functions of several lensing related quantities, including the deflection angle, the shear, the convergence, and the deflection potential. Some of these relations are applicable to galaxy-galaxy(-galaxy) lensing studies, and some others are of potential interest to studies of the gravitational lensing effect on the Cosmic Microwave Background.

Mathematically speaking, the shear field and the convergence field are the spin-2 and spin-0 second-order derivatives of the same scalar field– the deflection potential field. How the statistical properties of these spin-2 and spin-0 fields are related to each other has been demonstrated in this study with the derived relations. Due to the non-trivial spin number, special care is required in numerical evaluating these relations. We have presented a non-standard way of constructing sampling grids which lead to good numerical precision.

7.1.2 Bispectrum covariance in the flat-sky limit

To quantify the information content in lensing 3-pt statistics, one needs an expression for the covariance matrix of the 3-pt statistics. In Chap. 5 we derived an expression for the bispectrum covariance $\langle B(\ell_1,\ell_2,\ell_3)B(\ell_4,\ell_5,\ell_6)\rangle$ for cosmic shear. Our work has avoided the drawbacks of a previous work (Hu 2000), e.g. the expression given by Hu (2000) is valid only for integer arguments and does not allow a free binning choice, the formula contains the Wigner symbol whose physical meaning within a flat-sky consideration remains obscure, the finite survey size is accounted for only by multiplying a factor, which lacks solid justification, and an unjustified assumption is made in the coordinate transformation between the full sky and the 2D plane.

Since all these drawbacks are associated with the spherical harmonic formalism Hu (2000) adopted, we avoided them by using a pure two-dimensional Fourier-plane approach which has posed different challenges in the analytical derivation process. We defined an unbiased bispectrum estimator for 2D Fourier modes, averaged it over angular frequency bins to mimic the measuring process in reality, and computed the covariance of the averaged bispectrum estimator.

The covariance matrix we derived leads to the same Fisher information content as that given in Hu (2000). Moreover, our approach is mathematically rigorous, and allows one to evaluate the bispectrum and its covariance at real-valued angular frequencies and use e.g. a logarithmic binning.

7.1.3 Controlling intrinsic-shear alignment in three-point weak lensing statistics

In Chap. 6 we dealt with a particularly worrisome systematic error in cosmic shear studies: the intrinsic-shear alignment. We developed a method to control the intrinsic-shear alignment in three-point cosmic shear statistics by generalizing the nulling technique, a model-independent method developed to eliminate intrinsic-shear alignment at the two-point level. The generalization was found to be quite natural, and it resulted in a model-independent method to reduce the intrinsic-shear alignment signals (GGI and GII) in comparison to the lensing GGG signal.

As a test of the performance of the nulling technique at the three-point level, we assumed a fictitious survey with a setup typical of future multicolor imaging surveys, and applied the nulling technique to the modeled bispectra with intrinsic-shear alignment contamination. Since the GII signal can be removed simply by not considering tomographic bispectra with two or three equal redshift bins, we focused on the GGI signal, and quantified the intrinsic-shear alignment contamination with the GGI/GGG ratio. How much the contamination is reduced by the nulling technique was demonstrated both in terms of the GGI/GGG ratio, and in terms of biases on cosmological parameters in the context of an extended Fisher matrix study. In terms of the GGI/GGG ratio, a factor of 10 suppression is achieved after nulling over all angular scales. Correspondingly, the biases on cosmological parameters are reduced to be less than or comparable to the original statistical errors. By studying the performance of the nulling technique when different numbers of redshift bins are available, we found that the performance on bias reduction, rather than how much information is preserved during the nulling procedure, depends more significantly on the number of redshift bins. This suggests the need of more detailed redshift information if better control of intrinsic-shear alignment is required.

One disadvantage of the nulling technique is the unavoidable information loss. For the setup of our study, only about 20 % of the original information in lensing three-point statistics is preserved through the nulling procedure in terms of the Figure of Merit. A similar result (15 %) has been found when applying the nulling technique to lensing two-point statistics. We further studied the source of such information loss by comparing the nulling technique to an unconditioned linear compression of the data, motivated by the observation that the nulling procedure can be regarded as a linear compression of data under the constraint of the nulling condition. We found that around one third of the original information is lost through an unconditioned compression of the data, suggesting that this situation can be improved by considering higher-order terms in the nulling and compression processes.

We have also studied the combination of two- and three-point cosmic shear statistics. We found that the amount of nulled information contained in bispectrum measures alone and power spectrum measures alone are comparable. Adding the Fisher information matrix of the two, typically three-times better constraints are achieved both before and after nulling in terms of the Figure of Merit, than those obtained by each probe alone.

7.2 Outlook

The work presented in this thesis are all progresses towards a common goal: enabling the usage of three-point statistics in future lensing surveys to complement two-point statistics so as to exploit

more information from them. There are still many studies to be done in this respect. We mention some of them in the following.

7.2.1 Comparing observation to theory

When observational data is available, one needs to apply statistics on it, and relate the measured statistics to quantities predicted by theoretical models in order to obtain constraints on cosmological parameters. Since the direct observables of cosmic shear studies are the ellipticities of galaxies, regions around bright stars, satellite tracks, cosmic rays, etc. must all be discarded since precise shape measurement cannot be performed in these regions. The resulting complex survey geometry singles out the shear correlation functions as the statistics to be directly measured on the data. The remaining problems include, what is the statistic to be used to link the shear correlation functions to theoretical models. Precise theoretical models also need to be obtained.

Statistics allowing for a clean E/B-mode separation

Concerning the statistic to be used to link the shear correlation functions to theoretical models, an important requirement on it is the ability to cleanly separate the E- and B-modes. At the three-point level, the aperture mass statistics are the only statistics known up to now that enable an E/B-mode decomposition. The aperture mass statistics require measurements of shear correlation functions down to zero separation length which cannot be achieved in practice. It is found at the two-point level that this leads to a mixing of E- and B-modes. The three-point aperture mass statistics are expected to be plagued by the same E/B-mode mixing, but it is still yet to be shown how severe this mixing is. If the mixing is found to be of tolerable level compared to the statistical error, then the aperture mass statistics would be a convenient choice of the three-point statistics to be applied to future surveys. If it is not the case, then better statistics need to be constructed.

The formulation of the condition for E/B-mode separation at the three-point level (see Chap. 4) marks the first step in constructing better E/B-mode separating statistics. The next step is to formulate the condition for E/B-mode separation over a finite region, where by "over a finite region" we mean only the three-point shear correlation functions evaluated at specified spatial configurations are used as inputs. When this condition is also found, one can then construct a set of statistics satisfying these two conditions, in analogy to the recent works (Eifler et al. 2010; Fu & Kilbinger 2010; Schneider et al. 2010) done for two-point statistics. The resulting statistics will be the statistics to be derived both from theory and from observation, i.e. where the two will be compared.

Theoretical models

The currently available models for lensing three-point statistics are constructed from the matter density bispectrum. The best analytical approximation of the matter density bispectrum up to now is the fitting formula by Scoccimarro & Couchman (2001) which we used in Chap. 6. It fits the measurements in N-body simulations with an error of 15%, which is too high compared to the statistical error expected from future surveys. Thus efforts are required in this direction. Furthermore, as we have stated in Chap. 4, numerical difficulties exist in using bispectrum models to compare with observation. As the shear correlation functions are configuration space statistics while the bispectrum is a Fourier space quantity, highly oscillatory integrals are unavoidable while linking the two. This will probably render the comparison particularly time-consuming, and may further affect the precision of the final constraints on cosmological parameters.

In this situation, it may be worthwhile to investigate configuration space models, i.e. models for the three-point correlation function of the matter density field or that of the convergence field. If precise configuration space models are available, then one can constrain cosmological parameters using observational data with the aid of configuration space statistics only. Staying in configuration space has the potential advantage of better efficiency and precision. Direct theoretical predictions for matter density correlation functions remain at a far less-developed stage than those for their Fourier space counterparts due to mathematical difficulties. Predictions have only been given for two-point statistics at linear scales (Bashinsky & Bertschinger 2001). However, this does not rule out the possibility of finding precise fitting formulae for matter density correlation functions using N-body simulations. Even if configuration space models cannot reach the precision to be used to constrain cosmological parameters with future survey data, they can still play an important role in the pre-study phase. At least for the three-point statistics, the predictions of the shear signal made from configuration space models are much easier to numerically compute, as we have found in the study presented in Chap. 4.

7.2.2 Assessing the information content

The full expression for the covariance matrix given in Chap. 5 involves four-point and six-point statistics which are difficult to calculate analytically and hard to evaluate numerically. A way around is to use only the first term or the first two terms of it which are computationally feasible. It is then essential to study what error this causes. An on-going study by Martin (2011) looks at the error one introduces by keeping only the first term of the full expression. They find the error to be marginally acceptable for the convergence field in the local Universe, but their result is not conclusive yet. Finally a balance between precision and computational load has to be found.

Three-point statistics will be used complementary to two-point statistics to extract information from future surveys. So far it has been assumed that the two statistics are independent and their information content can be directly added. This assumption is expected to hold well at least at linear scales, since the covariance of two- and three-point statistics is a five-point statistic which vanishes completely for Gaussian random fields. However, in order to give correct combined constraints on cosmological parameters, and to give correct errors for these constraints, it is necessary to study the covariance of two- and three-point statistics. The result of such a study will determine the weight to be put on studies of three-point statistics compared to that on the two-point statistics.

Cluster counts have also been considered as a powerful probe of non-Gaussianity, in addition to the three-point statistics. How cluster counts compare to three-point statistics in probing non-Gaussian signals, and how they can be combined, are both questions requiring further study.

7.2.3 Controlling systematical errors

Systematic errors relevant to cosmic shear are usually classified into three groups: errors from the measurement process, those from theoretical modeling, and those due to complications of astrophysical processes. Errors in measuring the shapes and the redshifts of galaxies are the major sources of measurement errors. The requirements on controlling them are basic considerations in the design of lensing surveys. Theoretical modeling errors have been briefly discussed in Sect. 7.2.1. We shall now focus on errors of astrophysical origin, among which the intrinsic alignments are the most worrisome.

After the work presented in Chap. 6 was performed, there have been some new studies on intrinsic alignments (Joachimi & Schneider 2010; Joachimi et al. 2011), but we are still far from a

final answer to how severe intrinsic alignments are (especially at the three-point level), and how they can be best controlled. Improvements on the understanding of the intrinsic-alignment effect are needed. This requires efforts in theoretically studying the related astrophysical processes, e.g. simulating the process of galaxy formation in dark matter halos, and identifying the physical origins of the intrinsic-alignment effect. Observational measurements of the intrinsic-alignment signal are also necessary. Ideally the measurements should be detailed enough to distinguish the dependencies on galaxy type, luminosity, environment, etc. This also demands high-precision shape and redshift measurements in a survey.

If the interplay between observational and theoretical approaches enables one to precisely model the intrinsic alignments, then one can directly subtract the modeled effect from the lensing signal. However, the origin of the intrinsic-alignment effect may involve, or even is dominated by, stochastic processes, in which case precise modelling would be impossible. If this turns out to be the case, then one has to use model-independent methods to control the intrinsic alignments. Currently, the only completely model-independent method available is the nulling technique (see Chap. 6), which works at the cost of a large information loss. One possible way to model-independently control the intrinsic alignments without information loss is to 'calibrate' the shear signal using the cross-correlation of shear and galaxy number density (galaxy-galaxy lensing) and galaxy number density correlations in addition to shear correlations (Zhang 2010; Joachimi & Bridle 2010). The galaxy number density signal comes for free from a lensing survey, but a bias parameter is involved in its relation to the dark matter density field which gives rise to the shear signal. Thus one crucial pre-condition of this method being useful is the bias parameter being well-constrained by other observations and/or theoretical arguments. Hence, which method to be used to control the intrinsic alignments in future surveys depends on the development of many related fields. A combination of the ideas in the current methods may be needed.

7.2.4 General remarks

The application of three-point statistics in astronomy is not limited to cosmic shear studies. As a probe for non-Gaussianity, three-point statistics have been studied in the context of the Cosmic Microwave Background, the galaxy distribution, primordial curvature perturbations, etc. A common difficulty to all these studies is the huge number of configurations (triangles) that need to be examined, which poses hard problems in the measurement process, as well as in visualization and interpretation of the results. More studies are also needed for these aspects.

With the increasing interest in three-point statistics and the efforts put into it, it is promising that the existing problems related to cosmic shear three-point statistics can be solved or controlled to a negligible degree for future surveys e.g. the Euclid mission. In that case, cosmic shear three-point statistics may play an important role in determining the cosmological parameters, especially those related to the dark energy, and putting constraints on cosmological models. Furthermore, since cosmic shear three-point statistics can reflect the non-linear growth of structure under gravity, it can in principle provide a test on structure formation theory. By comparing the results from cosmic shear studies to those from other cosmological probes e.g. the Baryon Acoustic Oscillations and the Cosmic Microwave Background which constrain the geometry of the Universe, the theoretical foundation of modern cosmology– the General Relativity theory, can be put to a test at cosmological scales.

Acknowledgement

I owe many thanks to many people for their help throughout my PhD.

First of all, I want to thank my supervisor Peter Schneider for his guidance, encouragement, and support. I benefited a great deal from the many discussions we had, his remarks in the seminars and his wonderful lectures. During them he shared his knowledge as well as his way of thinking, showed a strong dedication to science and education as well as great enjoyment of them. As supervisor, he gave me a lot of freedom, and yet was himself ready for answering questions, proof-reading papers and this thesis, and helping with other scientific and bureaucratic problems. All of these I appreciate deeply.

Then I would like to thank Benjamin Joachimi who guided me as a senior student. Benjamin always amazed me with his maturity as a young scientist. I have benefited a lot from, and hopefully learnt from his diligence, patience and sense of responsibility, which are key factors to our fruitful collaboration. I am grateful also for his reading this thesis and making corrections and suggestions.

I would also like to show my gratitude to many other members of the institute, especially Jan Hartlap, Tim Eifler, Stefan Hilbert, and Cristiano Porciani for the insights they shared.

Special thanks to the DUEL network. It was great experience to get involved in the inter-institution communications enabled by DUEL. Thanks to all the people who devoted their time and energy to organizing the network, especially Andy Taylor and Nathalie Dupin.

It is pleasant to be a member of the International Max-Planck Research School for Astronomy and Astrophysics. I'm grateful to the coordinators Eduardo Ros and Emmanouil Angelakis, and the secretaries Gabriele Breuer and Simnone Pott for their advices, help and their efforts in organizing a variety of student activities.

Participation in the designing and tutoring of a photometry lab course is another important aspect of my PhD life. Big thanks to my co-workers Andreas Küpper, Marcel Pawlowski and Bharadwaj Vijaysarathy for making the process so enjoyable. I am happy to acknowledge Klaus Reif for helping with the telescope, Mischa Shirmer and Thomas Erben for helping with the data reduction procedure, and Maria Massi, Iskren Georgiev, Mischa Shirmer, Oliver-Mark Cordes and Michael Geffert who kindly contributed in the idea-collecting phase. I would also like to thank Peter for giving me the opportunity to work on this lab, and the people involved in the HOLIGRAIL project, especially Benjamin and Mischa, with whom I got trained in optical observation.

Many thanks to my colleagues at AIFA and fellow students in the IMPRS for the cordial atmosphere they created. My gratitudes to Kathy Schrüfer and Christina Stein-Schmitz for the help with administrative matters, and the computer group for the quick responses to computer-related problems.

Finally I want to thank my family members for their constant love and support.

Bibliography

Abdalla, F. B., Amara, A., Capak, P., et al. 2008, MNRAS, 387, 969

Albrecht, A., Bernstein, G., Cahn, R., et al. 2006, arXiv:astro-ph/0609591

Alpher, R. A., Bethe, H., & Gamow, G. 1948, Physical Review, 73, 803

Alpher, R. A. & Herman, R. 1948, Nature, 162, 774

Alpher, R. A. & Herman, R. C. 1950, Reviews of Modern Physics, 22, 153

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

Amara, A. & Réfrégier, A. 2008, MNRAS, 391, 228

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

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

Bartelmann, M. & Schneider, P. 2001, Phys. Rep., 340, 291

Bashinsky, S. & Bertschinger, E. 2001, Physical Review Letters, 87, 081301

Bergé, J., Amara, A., & Réfrégier, A. 2010, ApJ, 712, 992

Bernardeau, F., Colombi, S., Gaztañaga, E., & Scoccimarro, R. 2002a, Phys. Rep., 367, 1

Bernardeau, F., Mellier, Y., & van Waerbeke, L. 2002b, A&A, 389, L28

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

Bernstein, G. & Huterer, D. 2009, MNRAS, 1648

Bernstein, G. M. 2009, ApJ, 695, 652

Bordoloi, R., Lilly, S. J., & Amara, A. 2010, MNRAS, 783

Borodin, K. S., Kroshilin, A. E., & Tolmachev, V. V. 1978, Theoretical and Mathematical Physics, 34, 69

Brainerd, T. G., Agustsson, I., Madsen, C. A., & Edmonds, J. A. 2009, arXiv:0904.3095, submitted to ApJ

Bridle, S. & Abdalla, F. B. 2007, ApJ, 655, L1

Bridle, S. & King, L. 2007, New Journal of Physics, 9, 444

Brown, M. L., Taylor, A. N., Hambly, N. C., & Dye, S. 2002, MNRAS, 333, 501

Bunn, E. F. 2010, arXiv:1008.0827

Bunn, E. F., Zaldarriaga, M., Tegmark, M., & de Oliveira-Costa, A. 2003, Phys. Rev. D, 67, 023501

Catelan, P., Kamionkowski, M., & Blandford, R. D. 2001, MNRAS, 320, L7

Cooray, A. & Hu, W. 2002, ApJ, 574, 19

Crittenden, R. G., Natarajan, P., Pen, U., & Theuns, T. 2002, ApJ, 568, 20

Crocce, M. & Scoccimarro, R. 2006, Phys. Rev. D, 73, 063519

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

Davis, M., Huchra, J., Latham, D. W., & Tonry, J. 1982, ApJ, 253, 423

Dodelson, S. 2003, Modern cosmology (Academic Press)

Dyson, F. W., Eddington, A. S., & Davidson, C. 1920, Royal Society of London Philosophical Transactions Series A, 220, 291

Eifler, T., Schneider, P., & Krause, E. 2010, A&A, 510, A7

Einstein, A. 1917, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), Seite 142-152., 142

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

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

Erben, T., Hildebrandt, H., Lerchster, M., et al. 2009, A&A, 493, 1197

Esposito-Farèse, G. & Polarski, D. 2001, Phys. Rev. D, 63, 063504

Friedman, A. 1922, Zeitschrift fur Physik, 10, 377

Fry, J. N. 1984, ApJ, 279, 499

Fu, L. & Kilbinger, M. 2010, MNRAS, 401, 1264

Fu, L., Semboloni, E., Hoekstra, H., et al. 2008, A&A, 479, 9

Gamow, G. 1946, Physical Review, 70, 572

Goroff, M. H., Grinstein, B., Rey, S., & Wise, M. B. 1986, ApJ, 311, 6

Gradshteyn, I. S., Ryzhik, I. M., Jeffrey, A., & Zwillinger, D. 2000, Table of Integrals, Series, and Products (Academic Press)

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

Heymans, C., Brown, M., Heavens, A., et al. 2004, MNRAS, 347, 895

Heymans, C. & Heavens, A. 2003, MNRAS, 339, 711

Heymans, C., White, M., Heavens, A., Vale, C., & van Waerbeke, L. 2006, MNRAS, 371, 750

Hilbert, S., Hartlap, J., White, S. D. M., & Schneider, P. 2009, A&A, 499, 31

Hirata, C. M., Mandelbaum, R., Ishak, M., et al. 2007, MNRAS, 381, 1197

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

Hogg, D. W. 1999, arXiv:astro-ph/9905116v4

Hu, W. 1999, ApJ, 522, L21

Hu, W. 2000, Phys. Rev. D, 62, 043007

Hubble, E. 1929, Proceedings of the National Academy of Science, 15, 168

Hui, L. 1999, ApJ, 519, L9

Hui, L. & Zhang, J. 2002, arXiv:astro-ph/0205512, submitted to ApJ

Huterer, D., Takada, M., Bernstein, G., & Jain, B. 2006, MNRAS, 366, 101

Jain, B. & Bertschinger, E. 1994, ApJ, 431, 495

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

Jarvis, M., Bernstein, G., & Jain, B. 2004, MNRAS, 352, 338

Jing, Y. P. 2002, MNRAS, 335, L89

Joachimi, B. & Bridle, S. L. 2010, A&A, 523, A1

Joachimi, B., Mandelbaum, R., Abdalla, F. B., & Bridle, S. L. 2011, A&A, 527, A26

Joachimi, B. & Schneider, P. 2008, A&A, 488, 829

Joachimi, B. & Schneider, P. 2009, A&A, 507, 105

Joachimi, B. & Schneider, P. 2010, A&A, 517, A4

Joachimi, B., Schneider, P., & Eifler, T. 2008, A&A, 477, 43

Joachimi, B., Shi, X., & Schneider, P. 2009, A&A, 508, 1193

Kaiser, N. 1992, ApJ, 388, 272

Kaiser, N. 1995, ApJ, 439, L1

Kaiser, N. & Squires, G. 1993, ApJ, 404, 441

Kaiser, N., Wilson, G., & Luppino, G. A. 2000, arXiv:astro-ph/0003338, submitted to ApJ Letters

Kayo, I., Taruya, A., & Suto, Y. 2001, ApJ, 561, 22

Kilbinger, M. & Schneider, P. 2005, A&A, 442, 69

Kilbinger, M., Schneider, P., & Eifler, T. 2006, A&A, 457, 15

King, L. & Schneider, P. 2002, A&A, 396, 411

King, L. J. 2005, A&A, 441, 47

King, L. J. & Schneider, P. 2003, A&A, 398, 23

Lemaître, G. 1927, Annales de la Societe Scietifique de Bruxelles, 47, 49

Linder, E. V. 2003, Phys. Rev. Lett., 90, 091301

Longair, M. 2006, The cosmic century: a history of astrophysics and cosmology (Cambridge University Press)

Mackey, J., White, M., & Kamionkowski, M. 2002, MNRAS, 332, 788

Mandelbaum, R., Blake, C., Bridle, S., et al. 2011, MNRAS, 410, 844

Mandelbaum, R., Hirata, C. M., Ishak, M., Seljak, U., & Brinkmann, J. 2006, MNRAS, 367, 611

Martin, S. 2011, Diplomarbeit, 67, 1

Munshi, D., Valageas, P., van Waerbeke, L., & Heavens, A. 2008, Phys. Rep., 462, 67

Okumura, T. & Jing, Y. P. 2009, ApJ, 694, L83

Okumura, T., Jing, Y. P., & Li, C. 2009, ApJ, 694, 214

Ostriker, J. P., Nagamine, K., Cen, R., & Fukugita, M. 2003, ApJ, 597, 1

Peacock, J. A. 1999, Cosmological Physics (Cambridge University Press)

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

Peacock, J. A., Schneider, P., Efstathiou, G., et al. 2006, ESA-ESO Working Group on "Fundamental Cosmology", Tech. rep.

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

Pen, U.-L., Zhang, T., van Waerbeke, L., et al. 2003, ApJ, 592, 664

Penzias, A. A. & Wilson, R. W. 1965, ApJ, 142, 419

Pires, S., Starck, J., Amara, A., Réfrégier, A., & Teyssier, R. 2009, A&A, 505, 969

Refregier, A., Amara, A., Kitching, T. D., & Rassat, A. 2011, A&A, 528, A33

Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, AJ, 116, 1009

Robertson, H. P. 1935, ApJ, 82, 284

Rubin, V. C., Ford, W. K. J., & . Thonnard, N. 1980, ApJ, 238, 471

Sánchez, A. G., Crocce, M., Cabré, A., Baugh, C. M., & Gaztañaga, E. 2009, MNRAS, 400, 1643

Schneider, M. D. & Bridle, S. 2010, MNRAS, 402, 2127

Schneider, P. 1996, MNRAS, 283, 837

Schneider, P. 2003, A&A, 408, 829

Schneider, P. 2006, in Saas-Fee Advanced Course 33: Gravitational Lensing: Strong, Weak and Micro, ed. G. Meylan, P. Jetzer, P. North, P. Schneider, C. S. Kochanek, & J. Wambsganss, 269–451

Schneider, P. & Bartelmann, M. 1997, MNRAS, 286, 696

Schneider, P., Ehlers, J., & Falco, E. E. 1992, Gravitational Lenses (Springer-Verlag)

Schneider, P., Eifler, T., & Krause, E. 2010, A&A, 520, A116

Schneider, P. & Kilbinger, M. 2007, A&A, 462, 841

Schneider, P., Kilbinger, M., & Lombardi, M. 2005, A&A, 431, 9

Schneider, P. & Lombardi, M. 2003, A&A, 397, 809

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

Schneider, P., van Waerbeke, L., & Mellier, Y. 2002, A&A, 389, 729

Scoccimarro, R. & Couchman, H. M. P. 2001, MNRAS, 325, 1312

Scoccimarro, R. & Frieman, J. A. 1999, ApJ, 520, 35

Seitz, S. & Schneider, P. 1996, A&A, 305, 383

Seljak, U., Makarov, A., McDonald, P., et al. 2005, Phys. Rev. D, 71, 103515

Semboloni, E., Heymans, C., van Waerbeke, L., & Schneider, P. 2008, MNRAS, 388, 991

Shapiro, C. & Cooray, A. 2006, J. Cosmology Astropart. Phys., 3, 7

Shi, X., Joachimi, B., & Schneider, P. 2010, A&A, 523, A60

Shi, X., Schneider, P., & Joachimi, B. 2011, arxiv:1105.2309, submitted to Astronomy and Astrophysics

Smail, I., Ellis, R. S., & Fitchett, M. J. 1994, MNRAS, 270, 245

Smith, R. E., Peacock, J. A., Jenkins, A., et al. 2003, MNRAS, 341, 1311

Spergel, D. N., Bean, R., Doré, O., et al. 2007, ApJS, 170, 377

Squires, G. & Kaiser, N. 1996, ApJ, 473, 65

Sun, L., Fan, Z., Tao, C., et al. 2009, ApJ, 699, 958

Szapudi, I. & Kaiser, N. 2003, ApJ, 583, L1

Takada, M. & Jain, B. 2003a, ApJ, 583, L49

Takada, M. & Jain, B. 2003b, MNRAS, 344, 857

Takada, M. & Jain, B. 2004, MNRAS, 348, 897

Takada, M. & White, M. 2004, ApJ, 601, L1

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

van Waerbeke, L., Bernardeau, F., & Mellier, Y. 1999, A&A, 342, 15

Van Waerbeke, L. & Mellier, Y. 2003, arXiv:astro-ph/0305089

Van Waerbeke, L., Mellier, Y., Erben, T., et al. 2000, A&A, 358, 30

Walker, A. G. 1937, Proc. London Math. Soc., 42, 90

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

Zel'Dovich, Y. B. 1970, A&A, 5, 84

Zhang, P. 2010, ApJ, 720, 1090

Zwicky, F. 1937, ApJ, 86, 217