Cosmological Parameters from Second- and Third-Order Cosmic Shear Statistics

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Martin Kilbinger

aus

Ulm

Bonn 2006

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

1. Referent: Prof. Dr. Peter Schneider
2. Referent: Prof. Dr. Johannes Schmid-Burgk

Tag der Promotion: 9.12.2005

Contents

Introduction and overview

In recent years our knowledge about the Universe appears to have consolidated along with all observational data that seem to converge consistently into a standard model of cosmology. The satellite-based Cosmic Microwave Anisotropy Probe WMAP confirmed previous measurements of cosmological parameters with flying colors and to an unprecedented accuracy. Since then, experiments such as large galaxy surveys, Lyman-α forest and cosmic shear measurements have contributed to further tighten the commonly accepted paradigm of a flat, accelerating Universe whose large-scale structures are dominated by cold dark matter.

Despite all of the above, we still lack a fundamental understanding of the physics of the Universe. The nature of both dark matter and "dark energy" which together comprise 96% of the content of the Universe is still a mystery. Moreover, even though inflationary models provide the necessary special initial conditions in the early Universe and explain in a consistent way our observations, their predictions have yet to go through the mandatory path of scrutiny and confirmation. On the other hand, structure formation in the present Universe on scales smaller than a couple of Mpc, where non-linear physics plays an essential role, is still not understood in very much detail.

In order to probe the large-scale structure in the Universe and measure its distribution and properties, weak gravitational lensing is an excellent tool. The light of distant galaxies is deflected coherently by the tidal gravitational field of the intervening matter along the line of sight to the observer. The resulting distortions, or *cosmic shear*, can be measured statistically, revealing valuable information about the properties of the large-scale structure and, consequently, about cosmology and cosmological parameters. Since gravitational light deflection is independent of the nature and the state of matter, it is a unique tool to study the dark matter in the Universe. Future cosmic shear surveys will play a major role in cosmology. The sensitivity of weak lensing observations will be high enough to allow for precision measurements of cosmological parameters. These measurements will provide constraints on cosmology that are independent and complementary to those obtained from other experiments such as CMB anisotropy probes, supernovae of type Ia and galaxy surveys.

Cosmic shear is sensitive to matter on non-linear scales. It probes the non-linear evolution of structure and objects in the Universe and can be used to measure the non-Gaussianity of the large-scale structure arising from gravitational collapse. Up to now, cosmic shear observations have focused mainly on second-order statistics which only probe the Gaussian part of the matter distribution. Higher-order statistics that are able to determine non-Gaussian aspects of the large-scale structure will become more and more important and observational feasible with larger and deeper surveys, higher sensitivity and better understanding of systematics in the measurement and data analysis. Moreover, since the dependence on cosmology is different for shear statistics of second- and third order, the combination of both will improve the determination of cosmological parameters and reduce the degeneracies between them.

This thesis addresses the problem of constraining cosmological parameters from measurements of second- and third-order cosmic shear statistics. It is organized as follows:

- Chapter 1 sets the foundation needed for the understanding of cosmic shear which is the cosmological model together with structure formation and those aspects of gravitational lensing that are relevant for weak lensing by the large-scale structure.
- In Chapter 2, second- and third-order statistics of shear are defined along with estimators and their covariance, being essential for the analyses in the chapters that follow. Parts of this chapter are based on the publications Schneider, van Waerbeke, Kilbinger, & Mellier (2002); Schneider, Kilbinger, & Lombardi (2005) and Kilbinger & Schneider (2005).
- In Chapter 3, the effect of the survey geometry on cosmological parameter constraints is studied using analysis techniques involving likelihood, Fisher information matrix and Karhunen-Loève eigenmodes. The results of this work are published in Kilbinger & Schneider (2004) and Kilbinger & Munshi (2005).
- Chapter 4 compares third-order shear statistics from theoretical, non-linear models with ray-tracing simulations and portrays a visualization of the shear three-point correlation function.
- Chapter 5 presents predictions of the improvement of constraints on cosmological parameters by combining second- and third-order aperture mass statistics of cosmic shear. This part of the thesis is published in Kilbinger & Schneider (2005).

The major results of this work are summarized at the end of the corresponding Chapters 3, 4 and 5. This thesis concludes with a summary.

Chapter 1

Cosmology and the basics of gravitational lensing

Cosmology is the science of the physical Universe as a whole. It attempts to understand the properties and the dynamics of space-time itself, in which all objects we observe including ourselves are embedded. A cosmological model has to explain the overall shape and structure of the Universe, its time-evolution, and in a statistical sense describe how objects in the Universe form and evolve. The standard or "concordance" model of cosmology, which is outlined in this chapter, is based on the assumption of large-scale homogeneity and isotropy and quantifies the contents, the expansion and the age of the Universe with great success. On top of this homogeneous background, linear cosmological perturbation theory explains the growth of matter inhomogeneities from small initial fluctuation seeds, which leads to the formation of structures on all scales in the Universe. Finally, non-linear approaches based on high-resolution N-body simulations try to model collapsed and virialized objects like galaxy clusters which only formed in the recent past.

The most successful model to explain the nowadays numerous and consistent cosmological observations is based on cold dark matter (CDM, see Sect. 1.1.9) with a non-vanishing cosmological constant Λ (Sects. 1.1.5, 1.1.10) and consequently is abbreviated by the term ΛCDM model.

The main goal of this thesis concerns predictions from observations of the weak gravitational lensing effect by large-scale matter inhomogeneities in the Universe. More precisely, I examine the ability of these observations to constrain parameters of the ΛCDM model. In this introductory chapter, the standard model of cosmology with its parameters and the theory of structure formation in a CDM Universe are reviewed (Sect.1.1). Linear perturbation theory and some non-linear extensions are covered in Sect. 1.2 before the basic definitions, notations and properties of those aspects of gravitational lensing are given in Sect. 1.3 which are relevant for the later analysis of weak lensing statistics.

1.1 The homogeneous and isotropic Universe

1.1.1 The field equations

The dominant force in the Universe is gravitation. Not only is the formation of all major objects, starting from planets to stars, galaxies up to clusters of galaxies, governed by gravity, also the dynamics and evolution of the Universe as a whole is determined by this force. The fundamental theory of gravity is the General Theory of Relativity (Einstein 1916), which interprets gravity as a property of space-time. Space-time is described as a Riemannian manifold, which is "deformed" in the presence of masses. All particles travel on geodesics of

this manifold, and the effect of gravity on particle trajectories manifests itself in an alteration of the geodesics with respect to empty space. In classical (Newtonian) theory, this corresponds to the deviation of a particle trajectory from a straight line in the presence of a force field. The Riemann tensor $R^{\mu}{}_{\nu\lambda\rho}$ describes the curvature of the manifold, and depends on the metric $g_{\mu\nu}$ and its first and second derivatives.

The relation between space-time geometry and matter is expressed by Einstein's covariant field equations,

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \frac{8\pi G}{c^2} T_{\mu\nu} - \Lambda g_{\mu\nu}. \tag{1.1}$$

The Ricci tensor $R_{\mu\nu} = R^{\alpha}{}_{\mu\nu\alpha}$ and the Ricci scalar $R = R^{\alpha}{}_{\alpha}$ are contractions of the Riemann tensor. The matter and energy-momentum content of the Universe is described by the tensor $T_{\mu\nu}$. $G = 6.673 \times 10^{-11}\,\mathrm{m^3\,kg^{-1}s^{-2}}$ is Newton's gravitational constant and c the speed of light. A non-vanishing "cosmological constant" Λ is allowed in the field equations since it preserves the covariance of the equations and can be considered as a free parameter of the field equations. Initially, it was introduced by Einstein in order to allow for static solutions of (1.1). Later, it was realized that it can be interpreted as vacuum energy (Sect. 1.1.10) and only very recent observations strongly suggest that we live in a Universe with a non-vanishing Λ. However, there is still no convincing physical interpretation for this mathematical term in the equation.

1.1.2 The Robertson-Walker metric

The field equations (1.1) are non-linear and to solve them is far from trivial. Only in a few special cases for the form of the energy-momentum tensor $T_{\mu\nu}$ and the metric tensor $g_{\mu\nu}$ solutions can be found. A simple case, the one on which the concordance cosmology model is based, is a Universe which is homogeneous and isotropic. These two assumptions are the pillars of the *cosmological principle*. Observationally, isotropy around us is verified on scales larger than about 100 Mpc from the isotropic distribution of galaxies and distant radio sources, and from the isotropy of the CMB (COBE[1], Smoot et al. 1991). Combining these observations with the assumption that the Earth is not a special place in the Universe, isotropy around any point and therefore homogeneity of the Universe follows. Moreover, recent spectroscopic galaxy surveys confirm statistical homogeneity on scales larger than 100 Mpc (Hogg et al. 2005).

A metric which describes a homogeneous and isotropic Universe is the *Robertson-Walker* metric, independently found by Robertson (1935) and Walker (1936),

$$\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^{\mu}\mathrm{d}x^{\nu} = c^2\mathrm{d}t^2 - a^2(t)\left[\mathrm{d}w^2 + f_K^2(w)\left(\mathrm{d}\vartheta^2 + \sin^2\vartheta\,\mathrm{d}\varphi^2\right)\right]. \tag{1.2}$$

Here, $a(t)$ is the scale factor which describes the global expansion or contraction of the Universe. A remarkable consequence of this metric is that there exists a common universal time for each observer with zero peculiar velocity, which is the residual velocity component after subtracting the global expansion. Each of these so-called *fundamental observers* therefore experiences the same history of the Universe.

The spatial coordinates in (1.2) are given by the radial part or *comoving coordinate* $\mathrm{d}w$ and the angular element $\mathrm{d}\vartheta^2 + \sin^2\vartheta\,\mathrm{d}\varphi^2$. Note that the coordinate origin can be chosen arbitrarily. The radial function f_K is the comoving angular distance (see also Sect. 1.1.7) for

[1] COsmic Background Explorer

one of the three distinct 3-surfaces of constant curvature K: The sphere ($K > 0$), the flat Euclidean space ($K = 0$) and the hyperboloid ($K < 0$). It is of the form

$$f_K(w) = \int_0^w \frac{\mathrm{d}w'}{\sqrt{1 - K \cdot (w')^2}} = \begin{cases} K^{-1/2} \sin\left(K^{1/2} w\right) & \text{for} \quad K > 0 \\ w & \text{for} \quad K = 0 \\ (-K)^{-1/2} \sinh\left((-K)^{1/2} w\right) & \text{for} \quad K < 0\,. \end{cases} \tag{1.3}$$

1.1.3 The Friedmann equations

In order to solve Einstein's field equations (1.1), one has to make a particular choice for the energy-momentum tensor. Under the assumption of the cosmological principle, it has to take the form of a homogeneous and isotropic perfect fluid, $T_{\mu\nu} = (p/c^2 + \rho) U_\mu U_\nu - p\, g_{\mu\nu}$, where U_μ is the 4-velocity of the fluid. The density ρ and the pressure p are functions of time or of the scale factor a only.

Together with the Robertson-Walker metric, the Einstein equations can be simplified to the *Friedmann* equations (Friedmann 1922, 1924)

$$\begin{aligned} 3\left(\frac{\dot{a}^2}{a^2} + \frac{Kc^2}{a^2}\right) &= 8\pi G\rho + \Lambda, \\ \ddot{a} &= -\frac{4\pi G}{3}\left(\frac{3p}{c^2} + \rho\right) a + \frac{\Lambda}{3} a, \end{aligned} \tag{1.4}$$

where the dot represents the time-derivative. Usually, the normalization $a(t_0) = a_0 = 1$ is made for t_0 being the present epoch. These equations and some special solutions have independently been found by Lemaître (1927, 1931); a model which satisfies (1.2) and (1.4) is therefore called *Friedmann-Lemaître-Robertson-Walker* cosmology.

The two Friedmann equations (1.4) can be combined to yield the *adiabatic equation*,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\rho a^3 c^2\right) + p\,\frac{\mathrm{d}}{\mathrm{d}t} a^3 = 0, \tag{1.5}$$

which states that the energy change in a fixed comoving volume is compensated by the pressure times the volume change, in other words, it is the first law of thermodynamics in cosmological context.

1.1.4 Redshift

In a non-static Universe, photons are affected by the dynamics of space-time while they are propagating. Let a light beam of wavelength λ_e be emitted by a comoving source at time t_e. This signal as received at a later time t_0 by a comoving observer may have a different wavelength λ_0. The comoving distance w between source and observer is constant. Because light travels on null geodesics, $\mathrm{d}w = -c\mathrm{d}t/a$, and thus

$$w = \int_{t_\mathrm{e}}^{t_0} \frac{c\mathrm{d}t}{a} = \text{const.} \tag{1.6}$$

Differentiating with respect to t_e yields the time-dilation $\mathrm{d}t_0/\mathrm{d}t_\mathrm{e} = a(t_0)/a(t_\mathrm{e})$ which also applies to frequency and wavelength,

$$\frac{\mathrm{d}t_0}{\mathrm{d}t_\mathrm{e}} = \frac{a(t_0)}{a(t_\mathrm{e})} = \frac{\nu_\mathrm{e}}{\nu_0} = \frac{\lambda_0}{\lambda_\mathrm{e}} = 1 + z, \tag{1.7}$$

where in the last step the relative change of wavelength $1 + z$ was defined. In an expanding Universe, the scale factor a is increasing with time and photons are shifted towards longer wavelength. Therefore, z is called *redshift*.

Already in the 1920s, shortly after Hubble discovered his famous law according to which the redshift of galaxies is proportional to their distance from Earth, this redshift of galaxy spectra was interpreted as relative velocity and the expansion of the Universe was concluded.

1.1.5 Constituents of the Universe and density parameters

The energy-momentum content of the Universe can be separated into several species, which are characterized by their corresponding equation of state relating the density ρ to the pressure p of the species via

$$p = wc^2\rho, \tag{1.8}$$

where w is called the equation-of-state parameter. The Friedmann equations (1.4) can be solved for a single species with a constant w. From the adiabatic equation (1.5), one finds the following solution for the density as a function of a,

$$\rho \propto a^{-3(1+w)}. \tag{1.9}$$

Primarily, there exist three kinds of species that are relevant for the evolution of the Universe. The first kind consists of non-relativistic matter ρ_m with vanishing pressure ($w = 0$), also called "dust". This includes non-relativistic baryonic ρ_b and cold dark matter ρ_cdm ($\rho_\mathrm{m} = \rho_\mathrm{b} + \rho_\mathrm{cdm}$), see also Sect. 1.1.9. The second kind is radiation ρ_r which consists of relativistic particles with $w = 1/3$; its main ingredients are the CMB photons (Sect. 1.1.8) and relic relativistic neutrinos from the early Universe. The third kind is the cosmological constant which can be assigned the energy density ρ_Λ. Because of isotropy the Λ-term in (1.1) must be the same for all fundamental observers. Chosing appropriate coordinates, this term is proportional to the Minkowsk metric $\eta_{\mu\nu} = \mathrm{diag}(1, -1, -1, -1)$ and therefore, one can attribute an equation-of-state parameter $w = -1$ to the cosmological constant, which has the same form than a quantum-mechanical vacuum filled with virtual particles, see also Sect. 1.1.10.

A useful parametrization of the expansion of the Universe is done by introducing the *Hubble parameter* $H = \dot{a}/a$ as the relative expansion rate. Its present value, $H_0 = H(t_0)$ is the Hubble constant and is often written as $H_0 = 100\,h\,\mathrm{km\ s^{-1}Mpc^{-1}}$ with the observational uncertainty hidden in h. Most recent measurements from the combination of WMAP[2] and the Hubble key project yield $h = 0.71\,^{+0.04}_{-0.03}$ (Spergel et al. 2003).

The total density ($\rho_\mathrm{m} + \rho_\mathrm{r} + \rho_\Lambda$) of the Universe can take a special value, called the critical density ρ_c, for which the Universe is flat, or $K = 0$. Its present-day value is

$$\rho_{\mathrm{c},0} = \frac{3H_0^2}{8\pi G} = 1.879 \times 10^{-29}\,h^2\,\mathrm{g\,cm^{-3}}. \tag{1.10}$$

Therefore for an under-critical Universe ($\rho < \rho_\mathrm{c}$, $K < 0$), the geometry is open and for a super-critical one ($\rho > \rho_\mathrm{c}$, $K > 0$), it is closed. Note that if $\Lambda \neq 0$, a closed Universe can expand forever and does not necessarily recollapse.

The densities of the various species can be expressed as fractions of the critical density, e.g. $\Omega_\mathrm{m} = \rho_{\mathrm{m},0}/\rho_{\mathrm{c},0}$ for non-relativistic matter. Similarly, the density parameter for baryons is Ω_b

[2]Wilkinson Microwave Anisotropy Probe, `http://map.gsfc.nasa.gov`

and for relativistic particles is Ω_r. The density parameter corresponding to the cosmological constant is $\Omega_\Lambda = \Lambda/(3H_0^2)$. Recent measured values of the density parameters from a combination of CMB experiments (WMAP, CBI[3], ACBAR[4]), the 2dF galaxy survey and Lyman-α forest data are $\Omega_\mathrm{m} = 0.27 \pm 0.04$, $\Omega_\mathrm{b} = 0.044 \pm 0.004$ and $\Omega_\Lambda = 0.73 \pm 0.04$ (Spergel et al. 2003).

An upper boundary for the cosmic neutrino density is $\Omega_\nu < 0.015$ (Spergel et al. 2003), which seems non-negligible e.g. in comparison to the baryon density. However, to get a large neutrino density, the neutrino masses have to be high, up to 0.2 eV, which makes them behave as non-relativistic particles for most of the cosmic time (the present thermal energy of the cosmic background neutrinos is of the order of 10^{-4} eV, thus their rest mass is much larger than their kinetic energy) and their contribution to Ω_m is within the measurement errors.

In terms of the density parameters defined above, the first of the Friedmann equations can be written as

$$\left(\frac{H}{H_0}\right)^2 = \frac{\Omega_\mathrm{r}}{a^4} + \frac{\Omega_\mathrm{m}}{a^3} + \frac{1 - \Omega_\mathrm{m} - \Omega_\Lambda - \Omega_\mathrm{r}}{a^2} + \Omega_\Lambda, \tag{1.11}$$

and for the curvature one gets

$$K = \left(\frac{H_0}{c}\right)^2 (\Omega_\mathrm{m} + \Omega_\Lambda + \Omega_\mathrm{r} - 1)\,. \tag{1.12}$$

Because of its small contribution to the total present density, Ω_r can be neglected for late times.

1.1.6 Hot Big Bang and expansion history

From the observation that all objects whose peculiar velocity is small with respect to the Hubble flow are receding from us, we know that the Universe is expanding today, $\dot{a}(t_0) > 0$. If the right-hand side of the Friedmann equation (1.11) is not zero for some $a < 1$, one finds that $\dot{a}$ was always larger than zero in the past and the Universe expanded from an initial singularity[5] $a = 0$ which is assigned the time $t = 0$. The time that elapsed since then is called the world age t_0 and is of the order of the Hubble time H_0^{-1}. Recent estimates constrain the world age to $t_0 = 13.7 \pm 0.2$ Gyr (Spergel et al. 2003). The finite distance a photon can travel since the Big Bang is called the horizon which is $d_\mathrm{h}(t) = \int_0^t c\,\mathrm{d}t'/a(t')$ in comoving coordinates.

Inserting the three species $\rho_\mathrm{m}, \rho_\mathrm{r}$ and ρ_Λ individually into (1.9), one sees that for dust $\rho_\mathrm{m} \propto a^{-3}$, for radiation $\rho_\mathrm{r} \propto a^{-4}$ and for the cosmological constant $\rho_\Lambda = \mathrm{const}$. These functional dependencies also reveal themselves in the parametrized Friedmann equation 1.11. At very early times, the Universe was radiation-dominated and expanded with $a \propto t^{1/2}$. Using the Stefan-Boltzmann law, $\rho \propto T^4$, one gets for the temperature $T \propto a^{-1}$, thus the Universe was very hot in this era and cooled in the course of the expansion. The CMB (Sect. 1.1.8) is an observed relic of this hot epoch. At a redshift of $z_\mathrm{eq} \approx 23\,900\,\Omega_\mathrm{m} h^2$, the matter-dominated era started and the expansion slowed down to $a \propto t^{2/3}$. This lasted until a redshift of about $z_\mathrm{acc} = 0.4$ when the vacuum energy or cosmological constant took over. Since then the

[3] Cosmic Background Imager

[4] Arcminute Cosmology Bolometer Array Receiver

[5] For special initial values of the densities, i.e. large ρ_Λ and small ρ_m, a singularity could be avoided with a finite maximum redshift. However, observations exclude these "loitering" models.

Universe is considered to be in a phase of accelerated expansion, which eventually will be exponential with $a \propto \exp(\sqrt{\Lambda/3}\,t)$.

1.1.7 Distances

In a curved space-time, the concept of "distance" is not unambiguous. Distance is closely related to the measurement method used to obtain it and in contrast to the Euclidian case, different methods give different answers in a curved and non-static space-time.

A small *comoving distance* $\mathrm{d}D_{\mathrm{com}}$ between two nearby objects is the distance which remains constant in time if these two objects follow the Hubble flow, and which therefore is the comoving coordinate distance, $\mathrm{d}D_{\mathrm{com}} = \mathrm{d}w$. For null-geodesics, $\mathrm{d}s$ vanishes and $c\,\mathrm{d}t = -a\,\mathrm{d}w$ holds for the distance measured "backwards in time", i.e. towards smaller cosmic time. Thus, $\mathrm{d}D_{\mathrm{com}} = -c\,\mathrm{d}a/(Ha^2)$, and the total line-of-sight comoving distance is the integral over many infinitesimal contributions between the two redshifts z_1 and z_2 of the objects,

$$D_{\mathrm{com}}(z_1, z_2) = w(z_1, z_2) = \frac{c}{H_0} \int_{a(z_2)}^{a(z_1)} \mathrm{d}a \left[a\Omega_{\mathrm{m}} + a^2(1 - \Omega_{\mathrm{m}} - \Omega_\Lambda) + a^4\Omega_\Lambda\right]^{-1/2}. \qquad (1.13)$$

The *angular diameter distance* D_{ang} is given by the ratio of the physical size δL of an object at redshift z_2 to its apparent angular diameter $\delta\vartheta$ as seen by an observer at z_1. In analogy to the Euclidean case, one defines $D_{\mathrm{ang}} = \delta L/\delta\vartheta$. The coordinates of two edges of the object can be set to (w, ϑ, φ) and $(w, \vartheta + \delta\vartheta, \varphi)$, respectively. The spatial part of the line element in (1.2) then becomes $a^2(z_2)\, f_K^2(w(z_1, z_2))\, \delta\vartheta^2 = \delta L^2$ and thus,

$$D_{\mathrm{ang}}(z_1, z_2) = a(z_2)\, f_K(w(z_1, z_2)). \qquad (1.14)$$

The angular distance is the most prevalent distance used in gravitational lensing, because of its simple geometrical interpretation, see the lens equation (1.59). Note, that in general $D_{\mathrm{ang}}(z_1, z_2) \neq D_{\mathrm{ang}}(z_2, z_1)$.

1.1.8 The cosmic microwave background

At the redshift of $z_* \approx 1090$, 380 000 years after the Big Bang, the Universe cooled down to about $T \approx 3\,000$ K and photons and electrons which hitherto formed an ionized plasma decoupled. The electrons combined with protons to hydrogen and the Universe became neutral and transparent for photons. In the expanding Universe, these ubiquitous photons travel nearly unhindered through space and are redshifted, but since their distribution function does not change its form, they still have an almost perfect black body spectrum originating in the primordial plasma. The existence of this relic radiation as cosmic microwave background (CMB) with the today's temperature of $T = 2.725\,\mathrm{K}$, discovered serendipitously by Penzias & Wilson (1965), is one of the strongest proofs in favor of the Hot Big Bang model (Sect. 1.1.6).

Apart from a dipole moment of the order of 10^{-3} which is interpreted as due to the peculiar velocity of our galaxy, the CMB temperature is isotropic, and shows deviations from isotropy only at the level of one part in 10^5, therefore confirming the cosmological principle. At the same time, these tiny temperature anisotropies have turned out to be of great importance for cosmological observations and are interpreted in the following manner. At the time of decoupling, the photons emerged out of the gravitational potential of under- and overdense regions originating from small initial density perturbations and a signature of these perturbations

was imprinted as small temperature differences in the photon background. Moreover, these perturbations gave rise to acoustic waves in the primordial electron-photon plasma which froze in at decoupling. The largest possible wavelength of these oscillations is the sound horizon at decoupling $d_{\rm s}(z_*)$, which forms a characteristic scale in the angular power spectrum of the temperature fluctuations manifesting in a series of harmonic "acoustic peaks". Since the absolute physical scale of $d_{\rm s}(z_*)$ can be determined from relatively simple linear physics, the observation of the first acoustic peak together with an independent measurement of the Hubble constant yields the geometry of the Universe. The detection of the first peak at an angular scale of about 1° in the early 2000s by various experiments, most notably the balloon-based probe BOOMERanG[6] (Netterfield et al. 2002) and its confirmation by the satellite WMAP (Bennett et al. 2003), is a direct evidence for a flat Universe.

The exact position and the amplitude of the peaks depend on the composition and the properties of the Universe at high redshifts, which can be probed in great detail by observing the angular power spectrum of the CMB fluctuations. Further, the so-called secondary anisotropies are introduced in the CMB between redshifts of about 20 and zero, by processes such as scattering off electrons in the reionized Universe or a net gravitational redshift due to the time-varying gravitational potential of the large-scale structure. Thus, the observation of the CMB fluctuations provides a wealth of information about cosmology at early and late epochs.

1.1.9 Dark matter

There is a striking number of independent observations that indicate that the matter in the Universe is not dominated by ordinary baryonic matter but by a non-relativistic, weakly or non-interacting matter component, called *cold dark matter* (CDM). This matter unveils its presence only through gravitational interaction. In virtually all systems of galactic and supergalactic scales where the luminous and the gravitational masses have been measured, there is a discrepancy between the two, the latter always being larger than the former. The velocity dispersion of stars in elliptical and the rotation velocity in spiral galaxies are too high to be explained by the luminous matter of stars and gas; the same holds for the velocity dispersion of galaxies in clusters. Clusters without dark matter are not able to bind the hot X-ray gas observed in the intra-cluster medium. They also fail to produce the observed multiple images of background galaxies by the strong gravitational lensing effect (see Fig. 1.6). All these observations can be reconciled by postulating a dark matter component to constitute the missing mass in these objects.

From theoretical combined with observational studies of the primordial nucleosynthesis, this matter is deemed to be non-baryonic. A further inference derived from structure formation is that this matter is non-relativistic. Independently, from the measurement of the amplitude of the peaks in the CMB anisotropy power spectrum (Sect. 1.1.8), we know that baryonic matter alone is not sufficient to form high-density contrast objects such as galaxies and clusters at present time. Other alternative theories, trying to avoid the ad-hoc postulation of a non-baryonic dark matter component, require a modification of General Relativity and even Newtonian mechanics. All these attempts up to now have failed and their predictions are in contradiction to observations.

The nature of this dark matter component is still unknown, although a number of elemen-

[6]Balloon Observations Of Millimetric Extragalactic Radiation and Geophysics

tary particle candidates have been proposed. Experiments to directly detect these "weakly interacting massive particles" (WIMPs) have started to operate and will soon reach sensitivities to protude into interesting regions in the plane of mass and interaction strength. Until then, all information about dark matter and its properties has to be obtained from astronomical observations. For example, weakly-interacting relativistic particles as dark matter candidates ("Hot Dark Matter", HDM) such as massless neutrinos are excluded by the observational fact that small structures formed first and subsequently, larger objects like clusters evolved ("bottom-up" structure formation). In a HDM-dominated Universe, large objects form first since small-scale perturbations cannot grow due to free-streaming of HDM particles ("top-down" scenario). Next, limits on the self-interaction cross-section and on the mass of dark matter particles can be obtained by (the non-detection of) the annihilation γ-ray signal from high-density regions, e.g. the galactic center (Stoehr et al. 2003), and by measuring the profile and population of dark matter halos (Firmani et al. 2001). The question whether dark matter is in the form of particles at all, or maybe consists of compact massive objects such as black holes, can be answered by observations of micro-lensing events of stars in neighboring galaxies by MAssive Halo Objects (MACHOs).

1.1.10 Dark energy

The cosmological constant Λ was initially introduced ad hoc by Einstein in order to allow a static Universe as a solution for his field equations. Interpreted as part of the energy-momentum tensor, it has the form of a vacuum energy density, with an equation-of-state (1.8) parameter of $w = -1$ and acts as a repulsive force. Only recently, observations have shown Λ to be non-zero with high significance (Perlmutter et al. 1999; Riess et al. 1998; Spergel et al. 2003). Most importantly, the dimming of distant type Ia supernovae indicate an accelerated expansion of the Universe due to a positive cosmological constant. The physical nature of this quantity is still a mystery. Predictions from quantum field theory are in stark discordance with observations, the number of attempts to (until today unsuccessfully) model this "dark energy" are as numerous as names attributed to this phenomenon. More general models allow w to be different from -1 and vary in time. However, no observation up to now has indicated a deviation from a pure cosmological constant with $w = -1$.

1.1.11 Inflation

Apart from the yet unsolved nature of 96% of the Universe content (dark matter and dark energy), the standard model of cosmology presented so far faces still a few more difficulties and shortcomings. Firstly, it cannot explain the isotropy of the CMB for regions which had not been in causal contact to each other before recombination ("horizon problem"). Secondly, grand unified theories predict a ridiculously high density of magnetic monopoles (not a single one of which has been observed), and thirdly, the flatness of the Universe demands an extreme fine-tuning of the initial conditions.

The postulate of a so-called inflationary epoch (e.g. Linde 2005, and references therein) solves all these problems. According to this postulate, about 10^{-35} s after the Big Bang there was a very short period of exponential expansion like in a vacuum energy-dominated era. The scale factor increased by several dozens of e-foldings, and the Universe within the present horizon originated out of a tiny causal pre-inflationary region, which is a sturdy solution to the horizon problem. Furthermore, in the framework of inflationary theory virtually all

monopoles are "inflated away" and any initial curvature is flattened by the huge expansion so that after inflation the departure from flatness is minuscule.

The hypothesis of inflation has the advantage that it also explains the origin and the shape of the initial density perturbations, from which today's large-scale structures emerged (Sect. 1.2). This success of the inflationary model came as a complete surprise and was unintended when the idea of inflation was first introduced. Inflation is driven by a scalar field; quantum fluctuations in this field are inflated to macroscopic scales and form the seeds of the initial density perturbations. These are adiabatic or iso-entropy fluctuations, i.e. the density perturbations in matter and radiation are proportional and have equal phase. They are Gaussian, thus perturbations on different scales are uncorrelated and they are scale-invariant with a spectral index of the fluctuation power spectrum a little less than unity (see also Sect. 1.2.6.1). These properties of density perturbations have been confirmed by CMB measurements. Next-generation polarization CMB observations will be able to test further predictions of inflation such as the presence and form of the gravitational wave background created at the end of inflation.

It should be stressed here that inflation is really a paradigm and not a theory. There exist a large number of models within the inflationary framework, each one making slightly different predictions, and not a single one has yet convincingly emerged as a standard model of inflation.

1.2 Large-scale structure

The Universe is homogeneous and isotropic on very large scales and the Friedmann-Lemaître-Robertson-Walker cosmology (Sect. 1.1) accounts for the overall dynamics of the Universe. On smaller scales, below some hundred Mpc, one observes a vast variety of structures such as "walls" of matter, filaments, galaxy clusters and galaxies. The common picture is that all these objects originated from very small density perturbations, possibly of quantum origin, and later formed and evolved via gravitational instability. In this section, I will give a brief overview of the current theories and models that describe how density perturbations grow and evolve.

1.2.1 The fluid equations

On scales much smaller than the horizon and for non-relativistic matter, Newtonian physics can be used to describe the structure evolution. The standard equations for an ideal fluid of zero pressure in comoving coordinates are (Peebles 1980)

$$\text{Continuity eq.} \qquad \dot{\rho} + 3H\rho + \frac{1}{a}\nabla(\rho\,\boldsymbol{v}) = 0, \tag{1.15}$$

$$\text{Euler eq.} \qquad \dot{\boldsymbol{v}} + H\boldsymbol{v} + \frac{1}{a}(\boldsymbol{v}\nabla)\boldsymbol{v} = -\frac{1}{a}\nabla\phi, \tag{1.16}$$

$$\text{Poisson eq.} \qquad \Delta\phi = 4\pi G\rho a^2 + 3a\ddot{a} - \Lambda a^2, \tag{1.17}$$

where the dot indicates the time-derivative and ∇ is the gradient with respect to comoving spatial coordinates. These non-linear equations govern the evolution of the density ρ, the potential ϕ and the peculiar velocity field $\boldsymbol{v}$ in a dynamical Universe. The peculiar velocity is defined as the proper velocity $\boldsymbol{u}$ minus the Hubble flow, $\boldsymbol{v} = \boldsymbol{u} - H\boldsymbol{x}$.

It is useful to define the *density contrast*

$$\delta = \frac{\rho - \bar{\rho}}{\bar{\rho}}, \tag{1.18}$$

where $\bar{\rho}$ is the background density of a homogeneous and isotropic Universe. With that, the continuity equation (1.15) becomes

$$\dot{\delta} + \frac{1}{a}\nabla[(1+\delta)\boldsymbol{v}] = 0, \tag{1.19}$$

and the Poisson equation (1.17), using the second Friedmann equation (1.4), can be written as

$$\Delta\phi = \frac{3H_0^2\Omega_\mathrm{m}}{2a}\delta. \tag{1.20}$$

Note that any explicit dependence on the cosmological constant which was still present in (1.17) has vanished — a uniform matter/energy background has no influence on the evolution of density perturbations.

1.2.2 Linear perturbation theory

In order to solve the equations of motion (1.16), (1.19) and (1.20), one assumes that the matter is only slightly perturbed away from the homogeneous and isotropic background cosmology, corresponding to $\delta \ll 1$ and $v \ll u$. Considering only terms linear in δ and v, eqs. (1.16) and (1.19) become $\dot{\boldsymbol{v}} + H\boldsymbol{v} = -a^{-1}\nabla\phi$ and $\dot{\delta} + a^{-1}\nabla\boldsymbol{v} = 0$, respectively. Together with the Poisson equation (1.20), they can be combined to the second-order linear differential equation for the linear density contrast

$$\ddot{\delta} + 2H\dot{\delta} - \frac{3H_0^2\Omega_\mathrm{m}}{2a^3}\delta = 0. \tag{1.21}$$

This equation does not contain derivatives with respect to spatial coordinates, thus the general solution has the form

$$\delta(\boldsymbol{x}, a) = D_+(a)\delta_{1+}(\boldsymbol{x}) + D_-(a)\delta_{1-}(\boldsymbol{x}). \tag{1.22}$$

An index '1' is attached to the spatial parts of the solutions to indicate that these correspond to linear order. Higher-order solutions to the fluid equations are found in Sect. 1.2.4. One solution to (1.21) is the Hubble parameter H; since it is a decreasing function of a, $D_-(a) = H(a)$ is called the *decaying mode* and can be neglected for later times. Density perturbations in this mode quickly decay and play no role in structure formation. The second solution is the *growing mode*

$$D_+(a) = \frac{5\Omega_\mathrm{m}}{2}\frac{H(a)}{H_0}\int_0^a \mathrm{d}a' \left[1 + \Omega_\mathrm{m}\left(\frac{1}{a'} - 1\right) + \Omega_\Lambda\left(a'^2 - 1\right)\right]^{-3/2}. \tag{1.23}$$

For a Universe with $\Omega_\mathrm{m} = 1$ and $\Omega_\Lambda = 0$, called *Einstein-de-Sitter* (EdS) Universe, $D_+(a) = a$ for $a > a_\mathrm{eq}$, the latter being the scale factor at the time of matter-radiation equality, $a_\mathrm{eq} \approx 3.2 \times 10^{-5}\,\Omega_\mathrm{m}^{-1}h^{-2}$. In the radiation dominated era ($a < a_\mathrm{eq}$), the growing mode D_+ is constant since for radiation inhomogeneities, pressure prevents further collapse and for collisionless matter, the expansion of the Universe is too fast for the perturbations to grow via self-gravitation.

From the observations of CMB fluctuations (Sect. 1.1.8), one knows that at recombination ($z_* \approx 1090$) the baryon density perturbations were of the order of 10^{-5}. Since then, they only could have grown by a factor of $1/a_* \approx 10^3$ to a today's density contrast of about 10^{-2}, which corresponds to matter inhomogeneities still in the linear regime. Despite of that, non-linear perturbations much larger than this exist. This discrepancy is resolved by non- or weakly interacting dark matter (Sect. 1.1.9), which already at recombination had a much higher density contrast than baryons. This is one of the strongest arguments for the existence of non-baryonic matter in the Universe.

1.2.3 Non-linear equations of motion

In order to describe higher-order correlations and non-linear evolution, one has to depart from linear perturbation theory. Since it is not possible to find exact solutions of the non-linear dynamical equations (1.16, 1.19, 1.20), a perturbative approach is chosen, which will be developed in the next section. With $\theta = \nabla \boldsymbol{v}$ being the divergence of the peculiar velocity field, the full non-linear continuity equation (1.19) is written in Fourier space as

$$a\dot{\hat{\delta}}(\boldsymbol{k}, a) + \hat{\theta}(\boldsymbol{k}, a) = -\int \mathrm{d}^3x\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\boldsymbol{x}} \nabla(\boldsymbol{v}\delta)(\boldsymbol{x}, a), \tag{1.24}$$

where a hat indicates the Fourier transform (see eq. (A.4) for the conventions used in this work). After integration by parts, one writes the fields $\boldsymbol{v}$ and δ as Fourier integrals and gets

$$a\dot{\hat{\delta}}(\boldsymbol{k}, a) = -\int \frac{\mathrm{d}^3k_1}{(2\pi)^3} \int \frac{\mathrm{d}^3k_2}{(2\pi)^3} \mathrm{i}\boldsymbol{k}\hat{\boldsymbol{v}}(\boldsymbol{k}_1, a)\hat{\delta}(\boldsymbol{k}_2, a) \int \mathrm{d}^3x\, \mathrm{e}^{\mathrm{i}\boldsymbol{x}(\boldsymbol{k}-\boldsymbol{k}_1-\boldsymbol{k}_2)}. \tag{1.25}$$

One now assumes the velocity field $\boldsymbol{v}$ to be curl-free. This is justified since for a pressureless ideal fluid, linear vorticity perturbations are decaying modes. For a non-ideal fluid which may have a non-zero stress tensor, vorticity can be non-linearly amplified by shell-crossing effects but this happens only on very small scales, where perturbation theory breaks down anyway (Bernardeau et al. 2002). The velocity then has only a divergence-part $\theta = \nabla \boldsymbol{v}$; translated in Fourier space this means that $\hat{\boldsymbol{v}}(\boldsymbol{k}_1) \propto \boldsymbol{k}_1$ and $\boldsymbol{k} \cdot \hat{\boldsymbol{v}}(\boldsymbol{k}_1, a) = (\boldsymbol{k} \cdot \boldsymbol{k}_1/k_1)(\boldsymbol{k}_1/k_1 \cdot \hat{\boldsymbol{v}}(\boldsymbol{k}_1, a))$. The x-integral in (1.25) is simply $(2\pi)^3$ times Dirac's delta-function δ_D, and one gets

$$a\dot{\hat{\delta}}(\boldsymbol{k}, a) + \hat{\theta}(\boldsymbol{k}, a) = -\int \frac{\mathrm{d}^3k_1}{(2\pi)^3} \int \mathrm{d}^3k_2\, \delta_\mathrm{D}(\boldsymbol{k} - \boldsymbol{k}_1 - \boldsymbol{k}_2)\alpha(\boldsymbol{k}_1, \boldsymbol{k}_2)\hat{\theta}(\boldsymbol{k}_1, a)\hat{\delta}(\boldsymbol{k}_2, a) \tag{1.26}$$

with

$$\alpha(\boldsymbol{k}_1, \boldsymbol{k}_2) = \frac{(\boldsymbol{k}_1 + \boldsymbol{k}_2) \cdot \boldsymbol{k}_1}{k_1^2}. \tag{1.27}$$

In a similar way, (1.16) is written in Fourier space using Poisson's equation (1.20),

$$a\dot{\hat{\theta}}(\boldsymbol{k}, a) + \dot{a}\hat{\theta}(\boldsymbol{k}, a) + \frac{3H_0^2\Omega_\mathrm{m}}{2a}\hat{\delta}(\boldsymbol{k}, a) = -\int \mathrm{d}^3x\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\boldsymbol{x}} \left[\partial_i \left(v_j \partial_j\right) v_i\right](\boldsymbol{x}, a), \tag{1.28}$$

where summation over multiple indices is assumed. Integration by parts and expanding the fields $\boldsymbol{v}$ and θ into Fourier integrals, the right-hand side of the previous equation becomes

$$-\int \frac{\mathrm{d}^3k_1}{(2\pi)^3} \int \frac{\mathrm{d}^3k_2}{(2\pi)^3} \mathrm{i}\boldsymbol{k}\hat{\boldsymbol{v}}(\boldsymbol{k}_1, a) \left[\mathrm{i}\boldsymbol{k}\hat{\boldsymbol{v}}(\boldsymbol{k}_2, a) + \hat{\theta}(\boldsymbol{k}_2, a)\right] \int \mathrm{d}^3x\, \mathrm{e}^{\mathrm{i}\boldsymbol{x}(\boldsymbol{k}-\boldsymbol{k}_1-\boldsymbol{k}_2)}.$$

If the curl-part of the velocity field $\boldsymbol{v}$ is neglected as before, eq. (1.28) is written as

$$\begin{aligned} &a\dot{\hat{\theta}}(\boldsymbol{k},a) + \dot{a}\hat{\theta}(\boldsymbol{k},a) + \frac{3H_0^2\Omega_\mathrm{m}}{2a}\hat{\delta}(\boldsymbol{k}) \\ &\quad = -\int \frac{\mathrm{d}^3k_1}{(2\pi)^3}\int \mathrm{d}^3k_2\,\delta_\mathrm{D}(\boldsymbol{k}-\boldsymbol{k}_1-\boldsymbol{k}_2)\beta(\boldsymbol{k}_1,\boldsymbol{k}_2)\hat{\theta}(\boldsymbol{k}_1,a)\hat{\theta}(\boldsymbol{k}_2,a) \end{aligned} \tag{1.29}$$

with

$$\beta(\boldsymbol{k}_1,\boldsymbol{k}_2) = \frac{|\boldsymbol{k}_1+\boldsymbol{k}_2|^2(\boldsymbol{k}_1\cdot\boldsymbol{k}_2)}{2k_1^2k_2^2}. \tag{1.30}$$

The last expression is obtained with the requirement taken into account that the integrand in (1.29) is symmetric in $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$.

The kernels α (1.27) and β (1.30) describe the coupling between different Fourier modes which arises from the non-linear terms in the dynamical equations (1.16, 1.19). The evolution of both $\hat{\delta}$ and $\hat{\theta}$ at a given Fourier wavevector $\boldsymbol{k}$ is determined by the mode-coupling of the fields at all pairs of wave vectors $(\boldsymbol{k}_1,\boldsymbol{k}_2)$, the sum of which is equal to $\boldsymbol{k}$ as required by spatial homogeneity.

The non-linear equations of motion (1.26) and (1.29) as derived in this section cannot be solved analytically in general. For an EdS Universe, it is possible to obtain solutions with separate temporal and spatial dependence using a perturbative ansatz. This is sketched in the following section.

1.2.4 Second-order or quasi-linear solutions

In order to solve eqs. (1.26) and (1.29), the fields $\hat{\delta}$ and $\hat{\theta}$ are expanded around the linear solution, using the ansatz

$$\hat{\delta}(\boldsymbol{k},a) = \sum_{n=1}^{\infty} D_+^n(a)\hat{\delta}_n(\boldsymbol{k}), \qquad \hat{\theta}(\boldsymbol{k},a) = -\dot{a}\sum_{n=1}^{\infty} D_+^n(a)\hat{\theta}_n(\boldsymbol{k}), \tag{1.31}$$

where $\hat{\delta}_1$ and $\hat{\theta}_1$ are linear in the initial density field, $\hat{\delta}_2$ and $\hat{\theta}_2$ are the quadratic terms, etc. The motivation for this ansatz is the following: For the linear density contrast $\hat{\delta}_1$, its dependence on the scale factor is recovered, which is the linear growth factor $D_+(a)$ (1.22), and for every order n, the perturbative ansatz assures $\hat{\delta}_n = \mathcal{O}([\hat{\delta}_1]^n)$. For early times, the growth factor is much smaller than unity and the series is dominated by the first, linear term. The additional factor $-\dot{a}$ for $\hat{\theta}$ appears because from eq. (1.19), one obtains $\hat{\delta}_1 = \hat{\theta}_1$ to first order.

For an EdS Universe ($\Omega_\mathrm{m} = 1, \Omega_\Lambda = 0$), with $D_+(a) = a, \dot{a}^2 = -2\ddot{a}a$ and $H_0^2 = \dot{a}^2 a$, the dependence on time (or, equivalently, on the scale factor) can be separated out from eqs. (1.26) and (1.29), and one obtains to second order

$$\begin{aligned} 2\hat{\delta}_2(\boldsymbol{k}) &= \hat{\theta}_2(\boldsymbol{k}) + \int \frac{\mathrm{d}^3k_1}{(2\pi)^3}\int \mathrm{d}^3k_2\,\delta_\mathrm{D}(\boldsymbol{k}-\boldsymbol{k}_1-\boldsymbol{k}_2)\alpha(\boldsymbol{k}_1,\boldsymbol{k}_2)\hat{\delta}_1(\boldsymbol{k}_1)\hat{\delta}_1(\boldsymbol{k}_2); \\ \frac{5}{2}\hat{\theta}_2(\boldsymbol{k}) &= \int \frac{\mathrm{d}^3k_1}{(2\pi)^3}\int \mathrm{d}^3k_2\,\delta_\mathrm{D}(\boldsymbol{k}-\boldsymbol{k}_1-\boldsymbol{k}_2)\beta(\boldsymbol{k}_1,\boldsymbol{k}_2)\hat{\delta}_1(\boldsymbol{k}_1)\hat{\delta}_1(\boldsymbol{k}_2) + \frac{3}{2}\hat{\delta}_2(\boldsymbol{k}). \end{aligned}$$

Combining the two equations, one can write the second-order density contrast $\hat{\delta}_2$ as a function of the linear one $\hat{\delta}_1$,

$$\hat{\delta}_2(\boldsymbol{k}) = \frac{1}{2}\int \frac{\mathrm{d}^3k_1}{(2\pi)^3}\int \mathrm{d}^3k_2\delta_\mathrm{D}(\boldsymbol{k}-\boldsymbol{k}_1-\boldsymbol{k}_2)F_2(\boldsymbol{k}_1,\boldsymbol{k}_2)\hat{\delta}_1(\boldsymbol{k}_1)\hat{\delta}_1(\boldsymbol{k}_2), \tag{1.32}$$

and finds[7]

$$\frac{1}{2}F_2(\boldsymbol{k}_1,\boldsymbol{k}_2) = \frac{5}{7}\alpha(\boldsymbol{k}_1,\boldsymbol{k}_2) + \frac{2}{7}\beta(\boldsymbol{k}_1,\boldsymbol{k}_2) = \frac{5}{7} + \frac{1}{2}\frac{\boldsymbol{k}_1\cdot\boldsymbol{k}_2}{k_1k_2}\left(\frac{k_1}{k_2}+\frac{k_2}{k_1}\right) + \frac{2}{7}(\boldsymbol{k}_1\cdot\boldsymbol{k}_2)^2, \quad (1.33)$$

where the last expression is obtained by demanding F_2 to be symmetric in its arguments. I parametrize F_2 in the following way,

$$\begin{aligned} F_2(\boldsymbol{k}_1,\boldsymbol{k}_2) = F_2(k_1,k_2,\cos\varphi) &= \sum_{m=0}^{2} F_2^{(m)}(k_1,k_2)\cos^m\varphi; \\ F_2^{(0)} = \frac{10}{7}, \quad F_2^{(1)} &= \frac{k_1}{k_2}+\frac{k_2}{k_1}, \quad F_2^{(2)} = \frac{4}{7}, \end{aligned} \quad (1.34)$$

where φ is the angle between $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$. F_2 only depends on the scalar product of the wave vectors defining a triangle in Fourier space, therefore it is invariant under translations and rotations of the triangle.

For a cosmology with arbitrary Ω_m and Ω_Λ, the above ansatz is modified with an additional multiplicative function $f(\Omega_\mathrm{m},\Omega_\Lambda)$ for $\hat{\theta}$. Separable solutions are obtained for $f(\Omega_\mathrm{m},\Omega_\Lambda) \approx \Omega_\mathrm{m}^{1/2}$ which is a very good approximation in practice and leads to changes in the kernel F_2 of only a couple of percent (Bernardeau et al. 2002, Sect. 2.4.4). In Sect. 1.2.7.2, the non-linear extension of the results presented above will be applied to more general cosmologies.

The kernel F_2 describes the coupling between different Fourier modes which is characteristic of non-linear theories. The individual contributions to this kernel can be interpreted as follows. The constant term represents the isotropic contribution, which is independent of the configuration of the triangle in Fourier space, spanned by the wavevectors $\boldsymbol{k}, \boldsymbol{k}_1$ and $\boldsymbol{k}_2$. The linear cos-term arises from the gradient of the density field in direction of the flow (described by the $(\nabla\delta\cdot\boldsymbol{v})$-term in the continuity equation 1.19). The third, quadratic term represents the velocity gradient in the flow direction (the $(\boldsymbol{v}\cdot\nabla)\boldsymbol{v}$-term in the Euler equation 1.16). The last two terms give more weight to collinear wavevectors ($\varphi = 0, \pi$). This reflects the anisotropy of structures and flows, which can be seen in the filamentary structure of cold dark matter (see Fig. 1.3). Gradients in density and in velocity divergence are mainly generated parallel to the flow by gravitational instability. In Fig. 1.2, the reduced bispectrum, which will be defined in Sect. 1.2.6.2 is plotted, which gives an account of the shape of F_2. In Sect. 1.2.7.2, it is shown how deviations from this behavior in the non-linear regime are modeled.

1.2.5 Suppression of growth

For a density perturbation in radiation or baryonic matter, there exists a critical scale called the *Jeans length*, which is the minimum length at which gravity exceeds the opposing pressure gradient. Before matter-radiation equality, this length is about the size of the horizon d_h. In the matter dominated era, it drops basically to zero. Perturbations on scales smaller than the Jeans length are strongly suppressed because of radiation pressure which prevents them from growing. On larger scales, relativistic theory predicts perturbations to grow as $\delta(a) \propto a^2$ before equality and as $\delta(a) \propto a$ for $a > a_\mathrm{eq}$.

For a perturbation with a given comoving wavelength $\lambda > d_\mathrm{h}$, the horizon will eventually get larger than the perturbation because of the expanding Universe — the perturbation will "enter the horizon" for some scale factor a_enter. If this happens before a_eq, this perturbation

[7] this definition of F_2 differs from some previous work (e.g. Bernardeau et al. 2002) by a factor of two.

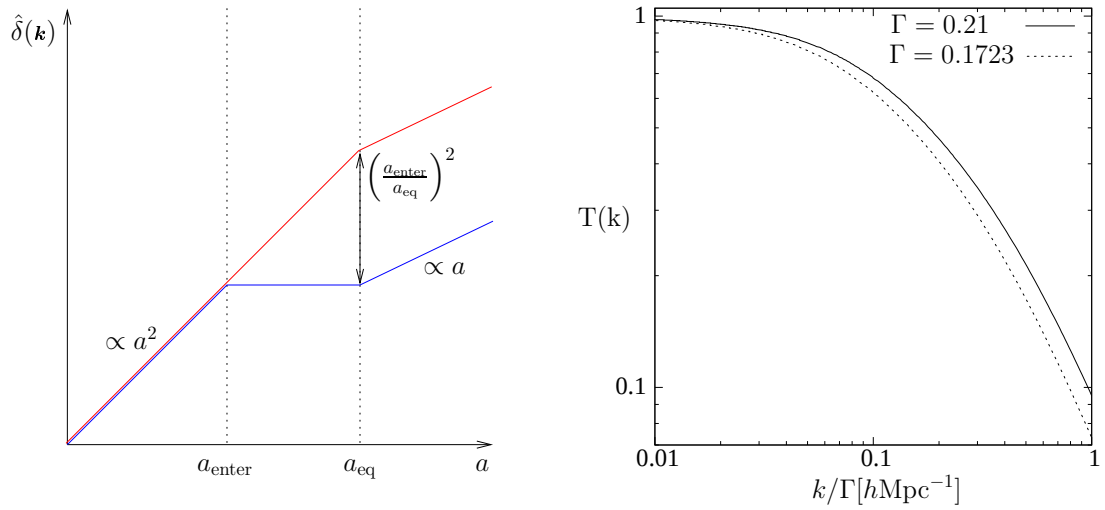

Figure 1.1: *Suppression of growth.* Left-hand side: *The small-scale perturbation which enters the horizon before $a_{\rm eq}$ is suppressed in growth relative to the large-scale fluctuation which does not enter before the matter dominated era starts. The suppression is $(a_{\rm enter}/a_{\rm eq})^2$.* Right-hand side: *The transfer function from eq. (1.37) for the two different values of Γ used in this work. For $\Omega_{\rm m} = 0.3$ and $h = 0.7$, the solid line corresponds to a baryon-free model, the dotted line approximates the transfer function with $\Omega_{\rm b} = 0.04$.*

will be suppressed with respect to a perturbation which enters after equality. The suppression factor as given by the growth factor in the radiation dominated era is $(a_{\rm enter}/a_{\rm eq})^2$, see Fig. 1.1. This scale-dependent evolution of perturbations defines a characteristic length scale in the large-scale structure which is the horizon at equality, $d_{\rm h}(a_{\rm eq}) = 12\,(\Omega_{\rm m}h^2)^{-1}$ Mpc.

In order to combine perturbations of all scales, one defines the *transfer function*

$$T(k) = \left(\frac{\hat{\delta}(k, a=1)}{\hat{\delta}(k, a_{\rm i})}\right)\left(\frac{\hat{\delta}(k=0, a=1)}{\hat{\delta}(k=0, a_{\rm i})}\right)^{-1}, \tag{1.35}$$

where $k = 0$ indicates an arbitrarily large scale. The scale factor $a_{\rm i}$ of the "initial" density fluctuation $\hat{\delta}(k, a_{\rm i})$ has to be chosen such that at this initial time no scale of interest has entered the horizon. Then, $T(k)$ is independent of $a_{\rm i}$ for all relevant scales.

For adiabatic cold dark matter perturbations, the transfer function asymptotically approaches unity for small k, since large scales are not suppressed. The behavior of $T(k)$ for large k can be inferred from the entering condition of a perturbation of scale λ:

$$\lambda = \frac{1}{k} \stackrel{!}{=} d_{\rm h}(a_{\rm enter}) \approx \frac{c}{a_{\rm enter}H(a_{\rm enter})}. \tag{1.36}$$

For $a \ll a_{\rm eq}$, the Hubble parameter scales like a^{-2}, thus $k \propto a_{\rm enter}^{-1}$. Therefore, $T(k) \propto k^{-2}$ for large k.

In this work, I use the fitting formula given in Bardeen et al. (1986),

$$T(k) = \frac{\ln(1+2.34\,q)}{2.34\,q}\left[1 + 3.89\,q + (16.1\,q)^2 + (5.46\,q)^3 + (6.71\,q)^4\right]^{-1/4}, \tag{1.37}$$

with $q = k/(\Gamma h\,\mathrm{Mpc}^{-1})$. The *shape parameter* Γ is initially set to

$$\Gamma = \Omega_\mathrm{m} h. \tag{1.38}$$

Peacock & Dodds (1994) have shown that the fitting formula (1.37) is also valid for a small baryon contribution to Ω_m if one sets $\Gamma = \Omega_\mathrm{m} h \exp(-2\Omega_\mathrm{b})$. Accordingly, an increasing baryonic density does not change the overall shape of the transfer function but shifts it towards smaller scales; therefore, on a given scale its amplitude is reduced. A generalization to models with $\Omega_\mathrm{m} \neq 1$ is provided by Sugiyama (1995),

$$\Gamma = \Omega_\mathrm{m} h \exp[-\Omega_\mathrm{b}(1 + \sqrt{2h}/\Omega_\mathrm{m})]. \tag{1.39}$$

In this work, both expressions (1.38, 1.39) for Γ will be used.

This approach to describe baryonic contributions to the matter component is only an approximation, since it does not take into account the acoustic oscillations in the baryon-photon plasma before decoupling. The resulting oscillations in the transfer function are only of the order of a couple of percent. More important is the decrease in amplitude, which for this work is taken into account by (1.39) with sufficient accuracy, see also Eisenstein & Hu (1998).

1.2.6 Moments of the density fields

Since the matter density contrast δ in the Universe has to be described as a random field (see App. A), the studies of δ are limited to its statistical properties. A random field is given by its moments; for its complete determination an infinite number of moments is needed in general.

For the special case of a Gaussian random field, the first- and second-order moments are sufficient to completely describe the field — all higher connected moments vanish. In the early Universe, the initial inhomogeneities (quantum fluctuations amplified during inflation) are Gaussian. As long as these perturbations grow linearly, they remain Gaussian. This has been confirmed observationally: no deviations from Gaussianity in the CMB fluctuations have been detected. Only in the quasi-linear and, more prominently, in the highly non-linear regime via gravitational collapse, non-Gaussian features in the matter field arise.

The first moment, or mean, of the matter density contrast δ (1.18) vanishes, since $\langle\rho\rangle = \bar{\rho}$. In this section, the second and third moments (two- and three-point correlation functions) of δ are discussed. In Fourier space, these are the power spectrum and the bispectrum, respectively. Up to now, in most of the cosmological studies and experiments, the power spectrum of δ has been examined. However, if one wants to detect deviations from non-Gaussianity, the inclusion of higher-order statistics is inevitable.

1.2.6.1 The power spectrum

The power spectrum of the density fluctuations P_δ is defined by

$$\left\langle \hat{\delta}(\boldsymbol{k}, a)\hat{\delta}^*(\boldsymbol{k}', a) \right\rangle = (2\pi)^3 \delta_\mathrm{D}(\boldsymbol{k} - \boldsymbol{k}') P_\delta(k, a), \tag{1.40}$$

where angular brackets denote the ensemble average (see App. A). For a given scale factor a, the power spectrum only depends on the modulus of the wavevector $\boldsymbol{k}$. If one chooses the

perturbative ansatz (1.31) for δ, only the first, linear order is needed in order to get the linear power spectrum.

The power spectrum $P_\delta(k,a)$ for some later time can be calculated from the initial one, $P_\delta(k,a) = D_+^2(a)\,T^2(k)\,P_\delta(k,a_\mathrm{i})/D_+^2(a_\mathrm{i})$. For the initial power spectrum $P_\delta(k,a_\mathrm{i})$ one usually assumes the simple and scale-free power law $P_\delta(k,a_\mathrm{i}) \propto k^{n_\mathrm{s}}$, with the scalar spectral index $n_\mathrm{s} \lesssim 1$ motivated by inflationary scenarios (Sect. 1.1.11) and confirmed observationally (e.g. Seljak et al. 2005). For $n_\mathrm{s} = 1$, the final linear power spectrum has the asymptotic behavior

$$P_\delta(k) \propto \begin{cases} k & \text{for small } k \\ k^{-3} & \text{for large } k \end{cases} . \tag{1.41}$$

The steep decrease of the power spectrum towards high k is due to the suppression of small-scale perturbations which entered the horizon d_h before matter-radiation equality a_eq (Sect. 1.2.5). The turnover scale depends on $d_\mathrm{h}(a_\mathrm{eq})$ and is the most characteristic scale in the dark matter power spectrum.

The power spectrum *normalization* is fixed by the parameter σ_8, which is defined as the variance of density fluctuations in spheres of radius $8\,h^{-1}$ Mpc. This particular scale was motivated observationally by the variance of galaxy counts, which is about unity at this scale, $\sigma_{8,\mathrm{gal}} \sim 1$. In order to relate this to the (dark) matter fluctuation amplitude, one has to model the bias between galaxies and dark matter, and assume that galaxies are tracers of the underlying dark matter. A linear stochastical bias simply translates into $\sigma_{8,\mathrm{gal}}^2 = b^2\sigma_8^2$, where b is the linear bias factor. The determination of the bias is observationally very challenging. One method is to combine cosmic shear measurements with galaxy number counts, see e.g. Simon (2004). Recent observations, including cosmic shear (e.g. van Waerbeke et al. 2005) indicate a value for σ_8 slightly smaller than unity.

1.2.6.2 The bispectrum

The bispectrum B_δ of the density field is the Fourier transform of the 3PCF of δ. It is defined by the third-order analog of (1.40), see also eq. (A.7),

$$\begin{aligned}\left\langle \hat\delta(\boldsymbol{k}_1,a)\hat\delta(\boldsymbol{k}_2,a)\hat\delta(\boldsymbol{k}_3,a)\right\rangle \\ = (2\pi)^3\delta_\mathrm{D}(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)[B_\delta(\boldsymbol{k}_1,\boldsymbol{k}_2,a)+B_\delta(\boldsymbol{k}_2,\boldsymbol{k}_3,a)+B_\delta(\boldsymbol{k}_3,\boldsymbol{k}_1,a)].\end{aligned} \tag{1.42}$$

Inserting the series expansion (1.31) for the density contrast, the triple-correlator $\langle\hat\delta\hat\delta\hat\delta\rangle$ on the left-hand side of this equation is, to lowest non-trivial order[8], $D_+^4(a)\langle\hat\delta_1(\boldsymbol{k}_1)\hat\delta_1(\boldsymbol{k}_2)\hat\delta_2(\boldsymbol{k}_3)\rangle +$ 2 permutations. But the second-order density contrast can be expressed in terms of the linear one using (1.32) and be rewritten in terms of fourth power in $\hat\delta_1$. Using (1.40) up to linear order, and noting that the connected part of the fourth order correlator vanishes for a Gaussian field, one can evaluate the integral in (1.32) and get

$$D_+(a)^4\langle\hat\delta_1(\boldsymbol{k}_1)\hat\delta_1(\boldsymbol{k}_2)\hat\delta_2(\boldsymbol{k}_3)\rangle = (2\pi)^3\,\delta_\mathrm{D}(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)F_2(\boldsymbol{k}_1,\boldsymbol{k}_2)P_\delta(k_1,a)P_\delta(k_2,a). \tag{1.43}$$

Comparing this expression with (1.42), one finds

$$B_\delta(\boldsymbol{k}_1,\boldsymbol{k}_2,a) = b_\delta(k_1,k_2,\cos\varphi,a) = F_2(k_1,k_2,\cos\varphi)P_\delta(k_1,a)P_\delta(k_2,a). \tag{1.44}$$

[8] the triple-correlator of the linear field $\hat\delta_1$ vanishes because of its Gaussian nature

Unlike the power spectrum which spatially depends only on one scalar quantity (the modulus of the wave vector), the bispectrum is a function of three variables, which uniquely describe a triangle. It is invariant under translations and rotations of this Fourier space triangle, which originates from the statistical homogeneity and isotropy of the density contrast δ. It is an even function of φ, corresponding to the parity-invariance of δ.

It is convenient to define the reduced bispectrum Q as

$$Q(k_1, k_2, \cos\varphi, a) = \frac{B_\delta(\boldsymbol{k}_1, \boldsymbol{k}_2, a) + B_\delta(\boldsymbol{k}_2, \boldsymbol{k}_3, a) + B_\delta(\boldsymbol{k}_3, \boldsymbol{k}_1, a)}{P_\delta(k_1, a)P_\delta(k_2, a) + P_\delta(k_2, a)P_\delta(k_3, a) + P_\delta(k_3, a)P_\delta(k_1, a)}. \tag{1.45}$$

Since the bispectrum is proportional to P_δ^2, the reduced bispectrum Q is mostly independent of time, scale and cosmology and displays the dependence on the triangle configuration of the bispectrum, given by the three Fourier vectors $\boldsymbol{k}_1, \boldsymbol{k}_2$ and $\boldsymbol{k}_3$, see Fig. 1.2.

1.2.7 Non-linear evolution

Bound objects in the low-redshift Universe have a density contrast much higher than unity, of up to $\delta \sim 1000$ for galaxy clusters. These objects cannot longer be described with linear perturbation theory. Several approaches exist to model non-linear evolution and clustering of these objects. In the highly non-linear regime, the collapse of objects can be calculated analytically for the simple case of an EdS cosmology. The transition zone between linear and highly non-linear objects is fitted using high-resolution N-body simulations.

1.2.7.1 The HKLM-method

For an EdS-Universe, Hamilton, Kumar, Lu, & Matthews (1991, HKLM) found a relation between the linear and the non-linear correlation functions of the density contrast. Peacock & Dodds (1994) generalized this relation to $\Omega_\mathrm{m} \neq 1$ and applied it to the power spectrum. I will shortly discuss the idea behind this approach.

A spherical region of volume r_L^3 with $\delta_\mathrm{L} \ll 1$ is assumed to collapse via non-linear evolution to a smaller, denser region of volume r_NL^3 and density contrast δ_NL. In reality, dissipation of the infalling matter into random motions will prevent collapse to a singularity and stabilize the halo, a process which is called *virialization*. Matter conservation yields $r_\mathrm{L}^3 = (1+\delta_\mathrm{NL})r_\mathrm{NL}^3$ which quantifies the change in scale of the mass clump. In order to express this equation in a statistical sense, one makes the transition from the density δ to its two-point correlation function (2PCF) ξ. From the 2PCF one can obtain a typical halo profile. However, for a collapsing object, it is the integrated overdensity which has to reach a critical value, thus instead of ξ one should consider the volume-averaged correlation function $\bar{\xi}(r) = r^{-3}\int_0^r \mathrm{d}x^3\xi(x)$ and write

$$r_\mathrm{L} = [1 + \bar{\xi}_\mathrm{NL}(r_\mathrm{NL})]^{1/3} r_\mathrm{NL}. \tag{1.46}$$

The next step in the HKLM method is the conjecture of a universal function f_NL which relates the linear and the non-linear volume-averaged correlation functions,

$$\bar{\xi}_\mathrm{NL}(r_\mathrm{NL}) = f_\mathrm{NL}[\bar{\xi}_\mathrm{L}(r_\mathrm{L})]. \tag{1.47}$$

The hypothesis of such a universal function was verified with EdS N-body simulations and later for general ΛCDM cosmologies. The asymptotic behavior of f_NL can be deduced readily. For $x \ll 1$, where non-linearities are negligible, the non-linear correlation function should just

be the linear one, $f_{\rm NL}(x) \propto x$. For large x, the corresponding non-linear process is *stable clustering*, which denotes the hypothesis that the internal structure of a collapsed object will be constant in proper time. Since the (non-linear) correlation function on small scales is the normalized auto-correlation of the halo profile, its only time-dependence comes from the varying number density of halos which scales as a^3, therefore $\bar{\xi}_{\rm NL} \propto a^3$. The linear density contrast (for EdS) is proportional to a (Sect. 1.2.2) which implies $\bar{\xi}_{\rm L} \propto a^2$, thus $f_{\rm NL}(x) \propto x^{3/2}$. The intermediate region is obtained by numerical fits.

In order to make the transition to Fourier space, one interprets $\bar{\xi}$ as a measure of the power at some effective wavenumber $k_{\rm eff}$. This is a valid approximation, because $\bar{\xi}$ is the integral over the usually rapidly rising dimensionless power spectrum $\Delta^2 = k^3/(2\pi^2)P_\delta(k)$ with a sharp cut-off window function. One can directly transfer eqs. (1.46) and (1.47) into Fourier space,

$$\begin{aligned} \Delta^2_{\rm NL}(k_{\rm NL}) &= f_{\rm NL}[\Delta^2_{\rm L}(k_{\rm L})]; \\ k_{\rm L} &= [1 + \Delta^2_{\rm NL}(k_{\rm NL})]^{-1/3} k_{\rm NL}. \end{aligned} \tag{1.48}$$

A fit for the universal function $f_{\rm NL}$, generalized to ΛCDM models, is obtained by using N-body simulations (Peacock & Dodds 1996).

1.2.7.2 Hyper-extended perturbation theory (HEPT)

Similar to the HKLM method, hyper-extended perturbation theory (HEPT) is an attempt to construct a non-linear (numerical) model, this time for the bispectrum of density fluctuations, by modifying perturbation theory predictions and using N-body simulations to fit the regime between linear and stable-clustering scales. HEPT was first introduced in Scoccimarro & Frieman (1999) for an EdS Universe and later generalized to ΛCDM models by Scoccimarro & Couchman (2001).

In HEPT, the kernel (1.34) is replaced by an effective kernel with modified coefficients,

$$F_2^{(0)} = \frac{10}{7} a(k_1) a(k_2), \quad F_2^{(1)} = \left(\frac{k_1}{k_2} + \frac{k_2}{k_1} \right) b(k_1) b(k_2), \quad F_2^{(2)} = \frac{4}{7} c(k_1) c(k_2). \tag{1.49}$$

The fitting functions a, b and c are given eqs. (10-12) of Scoccimarro & Couchman (2001) and interpolate between the linear and highly non-linear regime; their numerical coefficients are determined by fitting N-body simulations. The average deviation between the fit and the simulations are 15%, which is a sufficient accuracy for the purpose of this thesis. The fitting functions a, b and c not only depend on the Fourier mode k, but also on some (redshift-dependent) non-linear scale $k_{\rm NL}$, the local spectral index of the linear power spectrum $n(k)$ and weakly on σ_8 and the linear growth factor $D_+(a)$.

The effect of this modified effective kernel is shown in Fig. 1.2, where the reduced bispectrum Q (1.45) as a function of the angle between the two wave vectors $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ is plotted. Q reproduces the behavior of the kernel F_2, since after dividing the bispectrum by the square of the power spectrum, it is mainly depending on the triangle configuration and only marginally on cosmology.

The anisotropy which enhances collinear configurations ($\varphi = 0, \pi$), present in quasi-linear perturbation theory (Sect. 1.2.4), is recovered on large scales where non-linear contributions are unimportant. On small, non-linear scales however, the anisotropy is washed out and Q is more and more independent of the triangle configuration as the scale decreases. The physical

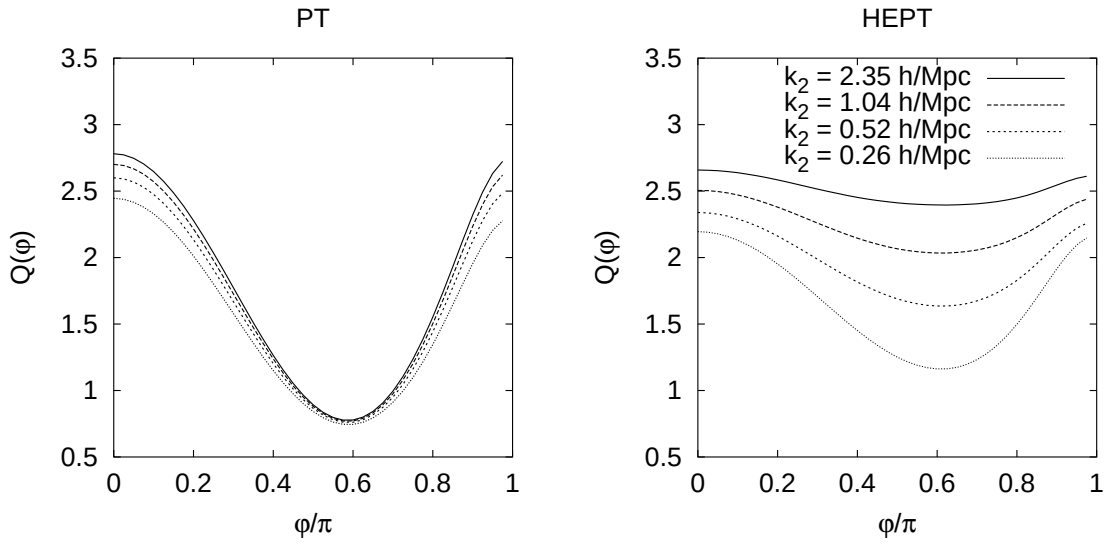

Figure 1.2: *The reduced bispectrum (1.45) as a function of the angle φ between $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$. In all cases, $k_2 = 2k_1$, the different curves are for different k_2 as shown in the panel. The cosmology is a ΛCDM Universe (model 2, see Table B.1) at a redshift of zero. The left panel shows the prediction of perturbation theory, the right panel takes the fitting formulae of HEPT into account.*

process responsible for this small-scale trend is virialization of dark matter halos: random motions cause structures and flows to lose coherence, and the anisotropy of Q for different triangle configurations diminishes. In the strongly non-linear regime, hierarchical clustering predicts Q to be constant. This behavior reflects the anisotropic, filamentary structure on very large scales as well as the isotropic, virialized clusters dominating the matter distribution on small, non-linear scales, see Fig. 1.3.

1.3 Gravitational lensing

In this section I will give an overview of the theory of gravitational lensing. After a short introduction to the basic ideas and notations used in gravitational lensing, the focus is set on the weak lensing effect caused by the large-scale structure of the Universe. A thorough treatment of the subjects addressed in this section can be found in Schneider et al. (1992, focusing on strong lensing) and Bartelmann & Schneider (2001, weak lensing). A recent seminal text covering most of the aspects of gravitational lensing and their applications exhaustively is Schneider (2005).

The modern theory of gravitational lensing is based on the General Theory of Relativity. In 1915, Einstein predicted a shift in the position of stars near the Sun due to the deflection of light caused by the Sun's gravitational field, and in 1920, measurements during a Solar eclipse confirmed the prediction. Soon afterwards, it was realized that for certain lens configurations the deflection can be large enough to produce multiple images of background sources. Subsequently in 1979, the first double image, the lensed quasar 0957+561, was discovered (Walsh et al. 1979). Since then, gravitational lensing has become a major research area with huge impact on a variety of astronomical fields. Lens systems on all scales, from stars over galaxies

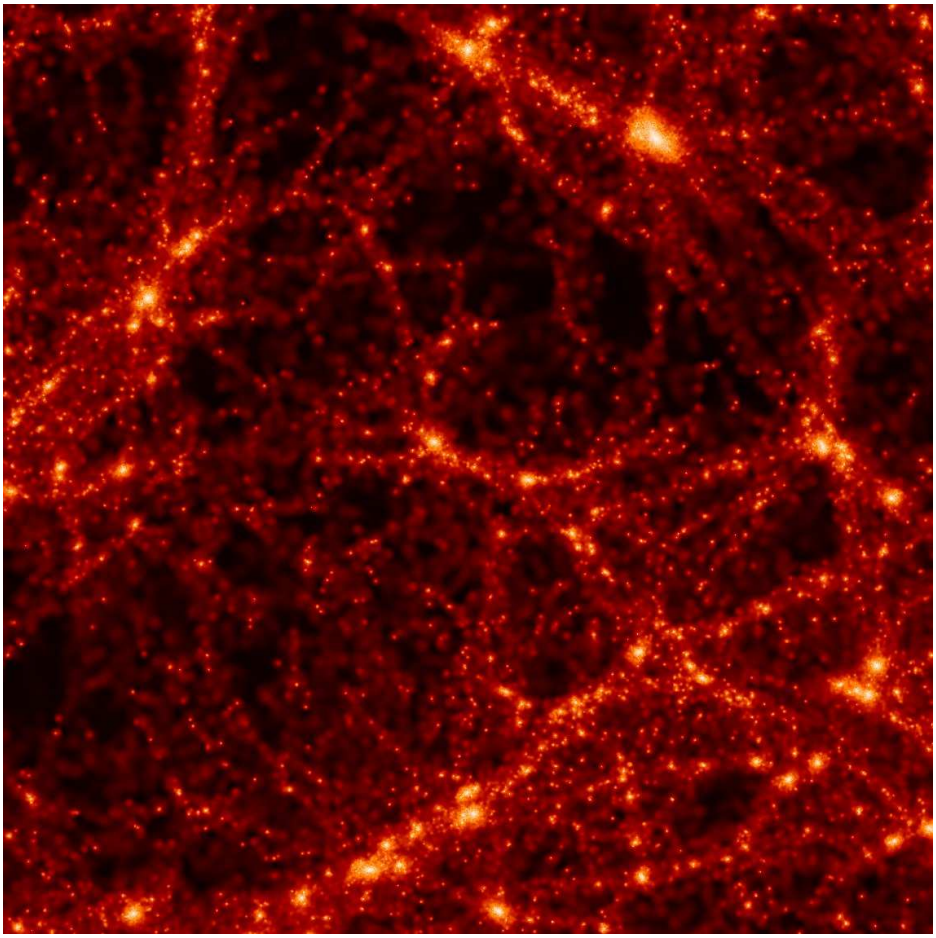

Figure 1.3: *The filamentary structure of the matter distribution on cosmic scales. The figure shows the density distribution of dark matter from a ΛCDM N-body simulation On large scales, matter forms anisotropic, collinear structures in the form of filaments. At the intersections of these filaments, the dark matter density is high enough to virialize into clusters. Source: The VIRGO consortium (Jenkins et al. 1998).*

and clusters up to the large-scale structure of the Universe are nowadays studied in great detail. A wealth of information about the composition of dark matter, the mass and profile of dark matter halos of galaxies and clusters, galaxy evolution and dynamics and, last but not least, cosmological parameters is obtained by these studies.

1.3.1 The deflection angle

The most basic parameter in gravitational lensing is the deflection angle, which quantifies the light deflection from the tidal field ϕ of a mass distibution. In the limit of weak gravitational fields, $\phi \ll c^2$, one can choose a metric which is only slightly perturbed away from that of Minkowski spacetime, and write the line element as follows,

$$\mathrm{d}s^2 = \left(1 + 2\phi/c^2\right) c^2 \mathrm{d}t^2 - \left(1 - 2\phi/c^2\right) \mathrm{d}x_i \mathrm{d}x^i. \tag{1.50}$$

For light-rays, $\mathrm{d}s^2 = 0$, and one gets $c\,\mathrm{d}t \approx (1 - 2\phi/c^2)\mathrm{d}\ell$ with $\mathrm{d}\ell^2 = \mathrm{d}x_i \mathrm{d}x^i$. The travel time for a light ray along a path γ is

$$t = \frac{1}{c}\int_\gamma \left(1 - 2\phi/c^2\right)\mathrm{d}\ell = \frac{1}{c}\int_\gamma n(\boldsymbol{x})\mathrm{d}\ell, \tag{1.51}$$

where in analogy to geometrical optics the index of refraction n is introduced in the second step. One now applies Fermat's principle, which states that the actual light path is stationary in the travel time, or $\delta t = 0$. The Euler-Lagrange equations of the variational principle $\delta \int n\mathrm{d}\ell = 0$ are

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(n\,\frac{\mathrm{d}x^i}{\mathrm{d}\ell}\right) = \frac{\partial n}{\partial x^i}\frac{\mathrm{d}\ell}{\mathrm{d}s}, \tag{1.52}$$

where the curve γ is parametrized with the arc length s. Next, these equations are integrated along the line of sight x^3 between the observer (O) and the source (S). Defining the vector $\boldsymbol{\xi} = (x^1, x^2)$ as being perpendicular to the x^3-direction between O and S, one gets

$$n\frac{\mathrm{d}\boldsymbol{\xi}}{\mathrm{d}\ell}\bigg|_{\mathrm{S}}^{\mathrm{O}} = \int_{\mathrm{S}}^{\mathrm{O}} \nabla_\perp n\,\frac{\mathrm{d}\ell}{\mathrm{d}s}\,\mathrm{d}s. \tag{1.53}$$

If the lens is far from the source and observer, the refractive index is unity at the end points O and S. The *deflection angle* $\hat{\boldsymbol{\alpha}}$ is the difference between the directions of the emitted and received light ray, and using the definition of n (1.51), one gets

$$\hat{\boldsymbol{\alpha}} = \frac{\mathrm{d}\boldsymbol{\xi}}{\mathrm{d}s}\bigg|_{\mathrm{O}} - \frac{\mathrm{d}\boldsymbol{\xi}}{\mathrm{d}s}\bigg|_{\mathrm{S}} = -\frac{2}{c^2}\int_{\mathrm{S}}^{\mathrm{O}} \nabla_\perp \phi\,\mathrm{d}\ell. \tag{1.54}$$

1.3.2 Point lens and thin lens approximation

The simplest case of a lens system consists of a point mass M and a point source. Assuming the impact parameter of the light ray ξ to be much larger than the Schwarzschild radius $R_{\mathrm{S}} = 2GM/c^2$ of the lens and much smaller than the distances between source, lens and observer, the path of the light ray can be approximated by piecewise straight lines. The deflection angle in this case is very accurately obtained by inserting the potential of a point mass $\phi = -GM/\xi$ into (1.54) and evaluating the integral along the straight lines. The result is

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \frac{4GM}{c^2\xi}\frac{\boldsymbol{\xi}}{\xi}. \tag{1.55}$$

Because $\xi \gg R_{\mathrm{S}}$, the deflection angle is much smaller than unity and indeed, for all known isolated lens systems, the deflection is smaller than an arc minute.

In the next step, a distribution of point masses m_i is considered. If again the impact parameter with respect to all point masses is much larger than the corresponding Schwarzschild radius, the resulting deflection angle is the (vectorial) sum of the deflections caused by the individual lenses,

$$\begin{aligned}\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) &= \frac{4G}{c^2}\sum_i m_i(\boldsymbol{\xi}_i, x_i^3)\frac{\boldsymbol{\xi} - \boldsymbol{\xi}_i}{|\boldsymbol{\xi} - \boldsymbol{\xi}_i|^2} \\ &= \frac{4G}{c^2}\int \mathrm{d}^2\xi' \int \mathrm{d}x'^3\,\rho(\boldsymbol{\xi}', x'^3)\frac{\boldsymbol{\xi} - \boldsymbol{\xi}'}{|\boldsymbol{\xi} - \boldsymbol{\xi}'|^2}.\end{aligned} \tag{1.56}$$

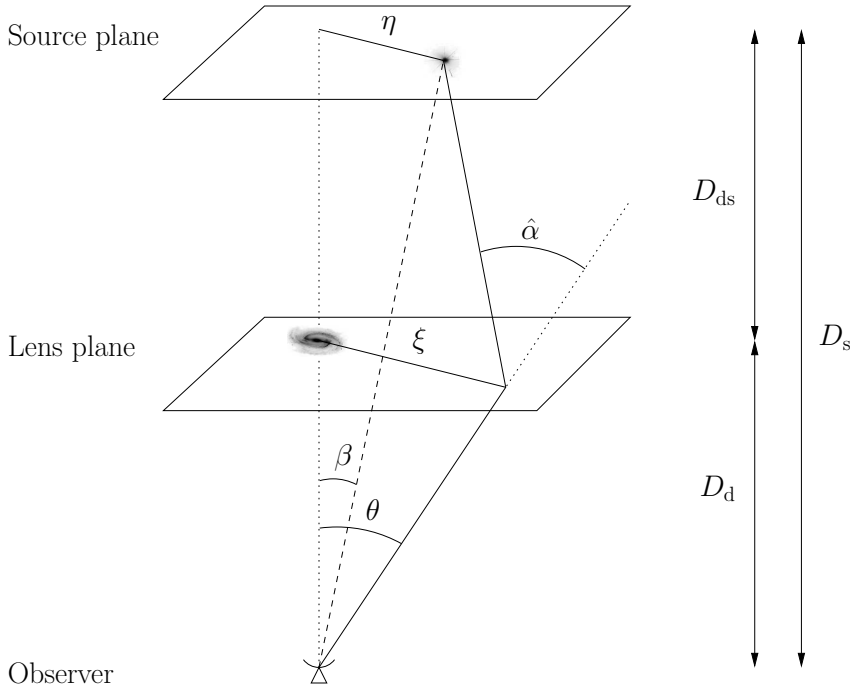

Figure 1.4: *Geometry of a lens system.*

In the second step, the transition from discrete point masses to a continuous mass distribution with density ρ was made. Defining the *surface mass density* $\Sigma(\boldsymbol{\xi}) = \int \mathrm{d}x^3\, \rho(\boldsymbol{\xi}, x^3)$, (1.56) can be rewritten as

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \frac{4G}{c^2} \int \mathrm{d}^2\xi'\, \Sigma(\boldsymbol{\xi}') \frac{\boldsymbol{\xi} - \boldsymbol{\xi}'}{|\boldsymbol{\xi} - \boldsymbol{\xi}'|^2}. \tag{1.57}$$

This expression is valid as long as the extent of the mass distribution along the line of sight is much smaller than the distances between source, lens and observer. This *thin lens approximation* is well satisfied when considering lensing by single galaxies or clusters, but fails in the case of cosmic shear where the mass distribution causes multiple deflections all along the line of sight.

1.3.3 The lens equation

An equation can be set up which relates the positions of the images and the source by simply looking at the geometrical configuration of the lens system as sketched in Fig. 1.4. Let $D_{\rm d}$ be the distance from the observer to the lens, or deflector, $D_{\rm s}$ the distance between the observer and the source, and $D_{\rm ds}$ the distance between lens and source. The lens and the source plane are defined as being perpendicular to the optical axis which is the line connecting observer and lens.

The source is located at a distance $\boldsymbol{\eta}$ from the optical axis, the impact vector is $\boldsymbol{\xi}$ as before. By using the intercept theorem, the following equation is found,

$$\boldsymbol{\eta} = \frac{D_{\rm s}}{D_{\rm d}} \boldsymbol{\xi} - D_{\rm ds} \hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}). \tag{1.58}$$

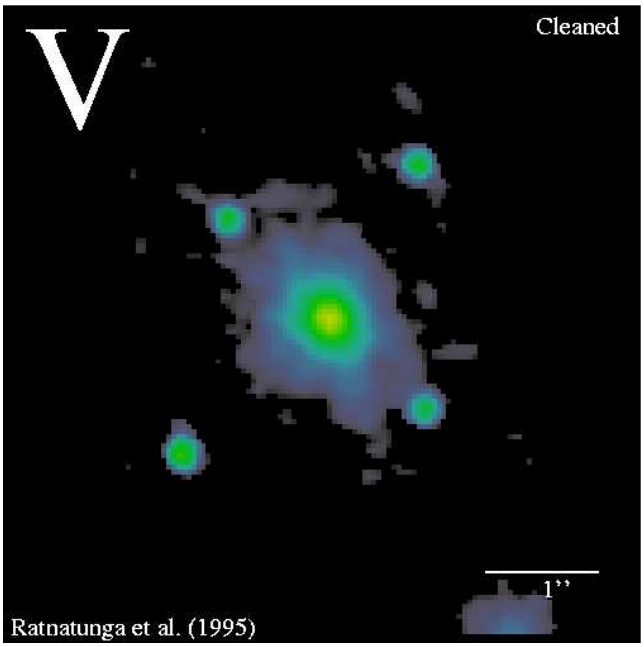

Figure 1.5: *The quadruple lens system HST14176+5226 with a red elliptical galaxy as lensing galaxy, from CASTLES[9]. The observation was made by Ratnatunga et al. (1995).*

This simple geometrical relation even holds for systems on cosmological scales, if angular diameter distances are used, see eq. (1.14). Relating distances to angles via $\boldsymbol{\eta} = D_\mathrm{s}\,\boldsymbol{\beta}$, $\boldsymbol{\xi} = D_\mathrm{d}\,\boldsymbol{\theta}$ and defining the scaled deflection angle $\boldsymbol{\alpha}(\boldsymbol{\theta}) = D_\mathrm{ds}/D_\mathrm{s} \cdot \hat{\boldsymbol{\alpha}}(D_\mathrm{d}\,\boldsymbol{\theta})$ results in the very concise form of the lens equation

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta}). \tag{1.59}$$

The solution of the lens equation for a given mass distribution determining the deflection angle and a given source position $\boldsymbol{\beta}$ is the observed position $\boldsymbol{\theta}$ of the image. The lens equation can have more than one solution which corresponds to multiple images of the source.

1.3.4 Convergence and shear

Introducing the *critical surface density*

$$\Sigma_\mathrm{cr} = \frac{c^2}{4\pi G}\frac{D_\mathrm{s}}{D_\mathrm{d} D_\mathrm{ds}}, \tag{1.60}$$

one can define a dimensionless surface mass density $\kappa(\boldsymbol{\theta}) = \Sigma(D_\mathrm{d}\,\boldsymbol{\theta})/\Sigma_\mathrm{cr}$, also called *convergence*. A mass distribution with $\kappa \geq 1$ at some position $\boldsymbol{\theta}$ produces multiple images provided there is a source at an appropriate position $\boldsymbol{\beta}$. Today, around 80 multiple image systems are known. In most of these cases, a high redshift quasar, typically at $z \gtrsim 2$ is lensed by a foreground galaxy of redshift $z \lesssim 1$; an example of a quadruple lens is shown in Fig. 1.5.

Using (1.60), equation (1.57) can be rewritten as

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{1}{\pi}\int \mathrm{d}^2\theta' \kappa(\boldsymbol{\theta}')\frac{\boldsymbol{\theta} - \boldsymbol{\theta}'}{\left|\boldsymbol{\theta} - \boldsymbol{\theta}'\right|^2}. \tag{1.61}$$

[9] CfA-Arizona Space Telescope LEns Survey, `http://cfa-www.harvard.edu/castles/`

The deflection angle is the gradient of the so-called *deflection potential*, $\boldsymbol{\alpha} = \nabla\psi$, which is

$$\psi(\boldsymbol{\theta}) = \frac{1}{\pi}\int \mathrm{d}^2\theta'\kappa(\boldsymbol{\theta}')\ln\left|\boldsymbol{\theta}-\boldsymbol{\theta}'\right|, \tag{1.62}$$

and satisfies the 2-D Poisson equation

$$\nabla^2\psi = (\partial_1\partial_1 + \partial_2\partial_2)\psi = 2\kappa. \tag{1.63}$$

The Jacobian of the mapping (1.59) is

$$\frac{\partial\beta_i}{\partial\theta_j} = \mathcal{A}_{ij} = \delta_{ij} - \frac{\partial^2\psi}{\partial\theta_i\partial\theta_j}. \tag{1.64}$$

By introducing the *shear*

$$\gamma_1 = \frac{1}{2}\left(\partial_1\partial_1\psi - \partial_2\partial_2\psi\right), \quad \gamma_2 = \partial_1\partial_2\psi, \tag{1.65}$$

the symmetric Jacobian matrix can be parametrized with the convergence and the shear as

$$\mathcal{A} = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix}. \tag{1.66}$$

The two-component shear $\gamma = (\gamma_1, \gamma_2)$ is a polar or spin-2 field, because it transforms identically under rotations by an angle π. It is very convenient to write the shear as a complex number, $\gamma = \gamma_1 + \mathrm{i}\,\gamma_2 = |\gamma|\mathrm{e}^{2\,\mathrm{i}\,\varphi}$ where the polar transformation property of the shear has been accounted for by the factor 2 in front of the polar angle φ.

Liouville's theorem guarantees surface brightness conservation by gravitational lensing, thus the observed intensity I at a position $\boldsymbol{\theta}$ is related to the intensity in the source plane I^{s} via $I(\boldsymbol{\theta}) = I^{\mathrm{s}}(\boldsymbol{\beta}(\boldsymbol{\theta}))$. If the angular extent of the source is smaller than the scale over which the lens properties change, this relation can be linearized in a neighborhood of the image position $\boldsymbol{\theta}_0$ to

$$I(\boldsymbol{\theta}) = I^{\mathrm{s}}\Big(\boldsymbol{\beta}(\boldsymbol{\theta}_0) + \mathcal{A}(\boldsymbol{\theta}_0)\,(\boldsymbol{\theta}-\boldsymbol{\theta}_0)\Big). \tag{1.67}$$

In general, this equation maps a circular source to an elliptical image. The convergence κ, being the diagonal part of the Jacobian, isotropically magnifies the image whereas the shear γ, being the trace-free part of $\mathcal{A}$, distorts the image. The axis ratio of the resulting ellipse is equal to the ratio of the eigenvalues of $\mathcal{A}$, which are $\lambda_{1,2} = 1-\kappa\pm\sqrt{\gamma_1^2+\gamma_2^2}$. An image is magnified by a factor equal to the ratio of the integrated intensity distributions $I(\boldsymbol{\theta})$ and $I^{\mathrm{s}}(\boldsymbol{\theta})$, which is the inverse of the determinant of $\mathcal{A}$,

$$\mu = \frac{1}{\det\mathcal{A}} = \frac{1}{(1-\kappa)^2-(\gamma_1^2+\gamma_2^2)}. \tag{1.68}$$

A lens system where $\kappa \geq 1$ at some position θ is called a *strong* lens. For typical strong lenses, the shear γ is also large, thus highly distorted images are produced. Very prominent examples of strong lenses are the centers of rich clusters where background galaxies show up as large luminous arcs, see Fig. 1.6.

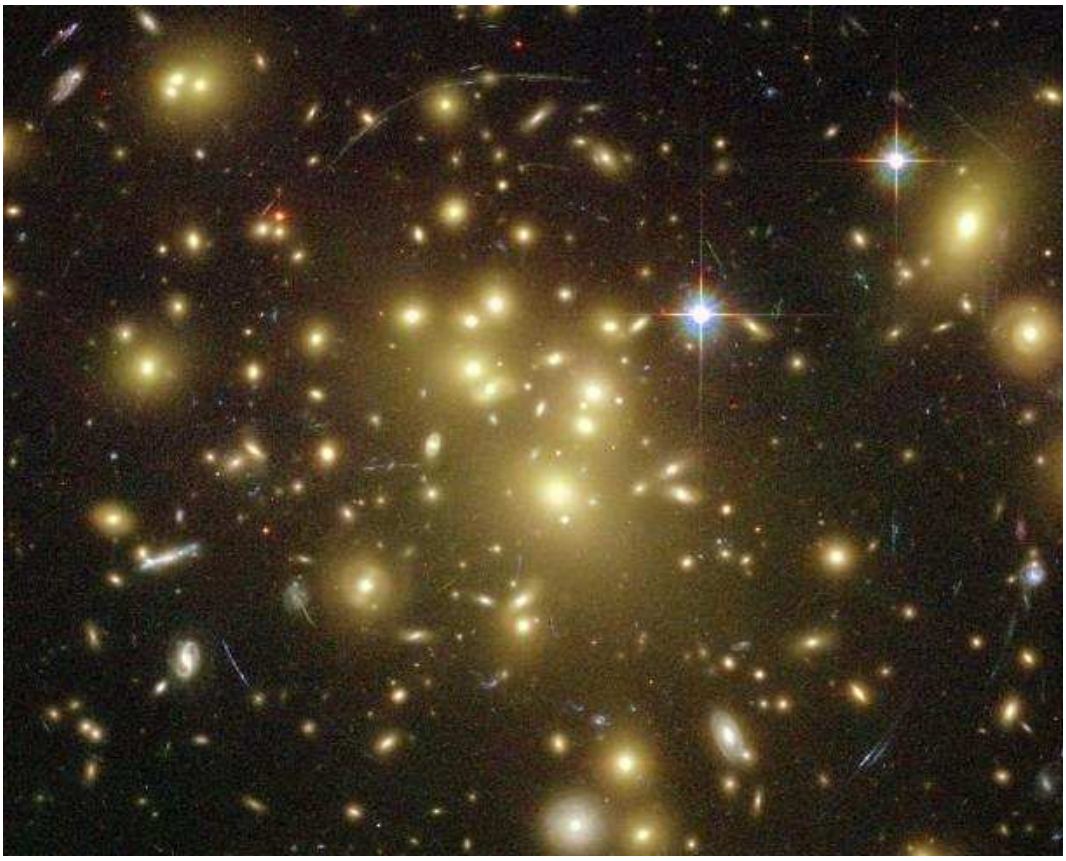

Figure 1.6: *The center of the galaxy cluster Abell 1689 as seen with the ACS camera on board of the HST. Source: NASA press release, N. Benitez et al. (2003).*

In order to get a relation between the shear and the convergence, one rewrites the expressions (1.63) and (1.65) in Fourier space,

$$\hat{\kappa}(\boldsymbol{\ell}) = -\frac{1}{2}\left(\ell_1^2 + \ell_2^2\right)\hat{\psi}(\boldsymbol{\ell}), \quad \hat{\gamma}_1(\boldsymbol{\ell}) = -\frac{1}{2}\left(\ell_1^2 - \ell_2^2\right)\hat{\psi}(\boldsymbol{\ell}), \quad \hat{\gamma}_2(\boldsymbol{\ell}) = -\ell_1\ell_2\hat{\psi}(\boldsymbol{\ell}) \tag{1.69}$$

and obtains

$$\hat{\gamma}(\boldsymbol{\ell}) = \hat{\gamma}_1(\boldsymbol{\ell}) + \mathrm{i}\hat{\gamma}_2(\boldsymbol{\ell}) = \frac{(\ell_1 + \mathrm{i}\ell_2)^2}{\ell^2}\hat{\kappa}(\boldsymbol{\ell}) = \mathrm{e}^{2\mathrm{i}\beta}\hat{\kappa}(\boldsymbol{\ell}), \tag{1.70}$$

where β is the polar angle of the Fourier vector $\boldsymbol{\ell}$. These relations are not defined for $\ell = 0$, corresponding to an arbitrarily large scale. The physical interpretation of this is that a homogeneous constant surface mass density causes no shear and κ can be determined from the shear only up to an additive constant. In analogy to (1.40), the power spectra of the convergence and the shear are defined and found to be equal, $P_\kappa = P_\gamma$ with

$$\left\langle \hat{\kappa}(\boldsymbol{\ell})\hat{\kappa}^*(\boldsymbol{\ell}')\right\rangle = \left\langle \hat{\gamma}(\boldsymbol{\ell})\hat{\gamma}^*(\boldsymbol{\ell}')\right\rangle = (2\pi)^2\delta_{\mathrm{D}}(\boldsymbol{\ell} - \boldsymbol{\ell}')P_\kappa(\ell). \tag{1.71}$$

The relation between the third-order moments of κ and γ is

$$\langle \hat{\kappa}(\boldsymbol{\ell}_1)\hat{\kappa}(\boldsymbol{\ell}_2)\hat{\kappa}(\boldsymbol{\ell}_3)\rangle = \langle \hat{\gamma}(\boldsymbol{\ell}_1)\hat{\gamma}(\boldsymbol{\ell}_2)\hat{\gamma}(\boldsymbol{\ell}_3)\rangle\, \mathrm{e}^{-2\mathrm{i}(\beta_1+\beta_2+\beta_3)}, \tag{1.72}$$

where β_i is the polar angle of the wavevector $\boldsymbol{\ell}_i$.

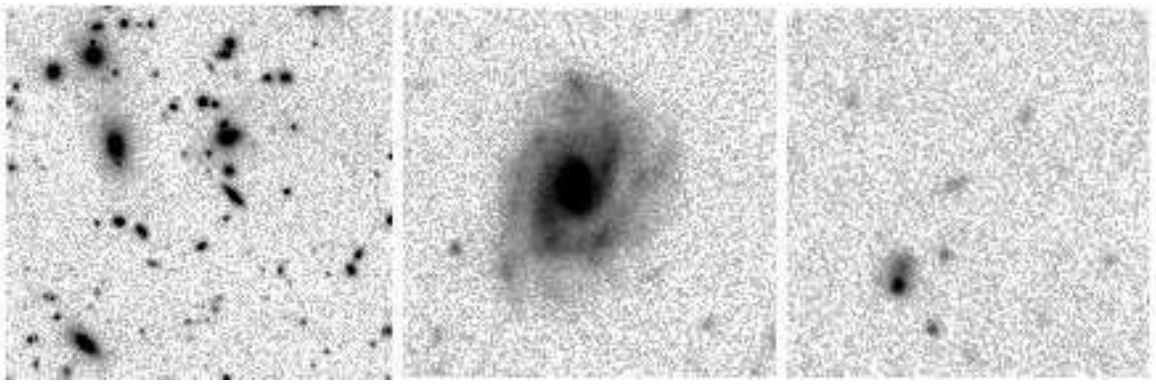

Figure 1.7: *Example of images of faint, high-redshift galaxies on a CCD chip. Source Y. Mellier.*

1.3.5 Weak lensing measurements

The regime where $\kappa \ll 1$ and $|\gamma| \ll 1$ is called the *weak lensing* regime. The distortions of distant galaxies are typically smaller than their intrinsic ellipticities which makes it impossible to tell to what extend an individual source is sheared. However, by observing a large number of galaxies and measuring their shapes, one can statistically detect this weak shear. For that, one has to quantify the shape of faint, often quite irregular galaxies which in addition are given as brightness distributions over only a handful of CCD pixels, see Fig. 1.7 for an example.

First, the center of an isolated brightness distribution I is determined as

$$\bar{\boldsymbol{\theta}} = \frac{\int \mathrm{d}^2\theta \, q_I[I(\boldsymbol{\theta})] \, \boldsymbol{\theta}}{\int \mathrm{d}^2\theta \, q_I[I(\boldsymbol{\theta})]}, \tag{1.73}$$

where q_I is a suitably chosen weight function in order to reduce the noise due to the sky background. Next, the second-moment tensor of the brightness distribution is defined as

$$Q_{ij} = \frac{\int \mathrm{d}^2\theta \, q_I[I(\boldsymbol{\theta})] \, (\theta_i - \bar{\theta}_i)(\theta_j - \bar{\theta}_j)}{\int \mathrm{d}^2\,\theta \, q_I[I(\boldsymbol{\theta})]}, \quad i,j = 1,2, \tag{1.74}$$

and the complex *ellipticity* of an object describing its shape is deduced as

$$\varepsilon = \frac{Q_{11} - Q_{22} + 2\mathrm{i}Q_{12}}{Q_{11} + Q_{22} + 2(Q_{11}Q_{22} - Q_{12}^2)^{1/2}}. \tag{1.75}$$

For an object with elliptical isophotes, the axis ratio r of the isophotes is related to the above defined ellipticity as $|\varepsilon| = (1-r)/(1+r)$. The following relation between the intrinsic or source ellipticity ε^{s} of an object and its observed ellipticity ε holds:

$$\varepsilon^{\mathrm{s}} = \begin{cases} \dfrac{\varepsilon - g}{1 - g^*\varepsilon} & \text{for} \quad |g| \le 1 \\ \dfrac{1 - g\varepsilon^*}{\varepsilon^* - g^*} & \text{for} \quad |g| > 1 \end{cases}, \tag{1.76}$$

where g is the *reduced shear*,

$$g = \frac{\gamma}{1-\kappa}. \tag{1.77}$$

This transformation becomes particularly simple for $\kappa \ll 1$ and $|\gamma| \ll 1$. Then, $g \approx \gamma$ and, if additionally $|\varepsilon^{\rm s}| \lesssim 1/2$, the observed ellipticity is just the sum of the intrinsic ellipticity and the shear,

$$\varepsilon \approx \varepsilon^{\rm s} + \gamma. \tag{1.78}$$

This is the basic equation of weak lensing. Under the simple assumption that the intrinsic orientation of galaxies is random, which is justifiable on grounds of statistical isotropy of the large-scale structure, the expectation value of the intrinsic ellipticity vanishes, $\langle \varepsilon^{\rm s} \rangle = 0$. Therefore, the observed ellipticity is an albeit very noisy but unbiased estimator of the shear,

$$\langle \varepsilon \rangle = \gamma. \tag{1.79}$$

1.3.6 The mass-sheet degeneracy

For a general lens system the source position β is not observable and the source size and luminosity is not known unless in the rare case of a standard candle such as a supernova of type Ia. This leads to a non-uniqueness of the lens surface mass density κ, which reproduces the observational constraints of this system, be it in the strong or weak lensing regime. A family of lens models

$$\kappa_\lambda = (1 - \lambda) + \lambda\kappa, \tag{1.80}$$

rescales the convergence by the multiplicative factor λ and adds a constant term or mass sheet $(1 - \lambda)$ to it. This ambiguity in the surface mass density is therefore called the *mass-sheet degeneracy*. The family κ_λ satisfies the Poisson equation $\Delta\psi_\lambda = 2\kappa_\lambda$ (cf. eq. 1.63) with $\psi_\lambda(\boldsymbol{\theta}) = (1 - \lambda)\theta^2/2 + \lambda\psi(\boldsymbol{\theta})$. Defining a family of deflection angles $\boldsymbol{\alpha}_\lambda(\boldsymbol{\theta}) = \nabla\psi_\lambda(\boldsymbol{\theta}) = (1 - \lambda)\boldsymbol{\theta} + \lambda\boldsymbol{\alpha}(\boldsymbol{\theta})$, one inserts (1.80) into the lens equation (1.59) and gets

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \boldsymbol{\alpha}_\lambda(\boldsymbol{\theta}) = \lambda\left[\boldsymbol{\theta} - \kappa(\boldsymbol{\theta})\right], \tag{1.81}$$

which has the same form than the 'original' lens equation (1.59) except for an unobservable rescaling of the source position. The Jacobi matrix (1.66) changes to $\mathcal{A}_\lambda = \lambda\mathcal{A}$ and therefore, the (unobservable) shear γ (1.65) is rescaled in the same way, but the (observable) reduced shear (1.77) is invariant. In weak lensing systems, the additional use of redshifts of the background galaxies can break the mass-sheet degeneracy (Bradač et al. 2004). Since the magnification changes to $\mu_\lambda = \mu/\lambda^2$, the number density of background images is modified as a consequence, and number counts could in principle lift the degeneracy (e.g. Dye et al. 2002); however, since firm number count observations are particularly difficult they are rarely made. In Sect. 2.1 the aperture mass $M_{\rm ap}$ will be introduced as a measure of the local surface mass density which is independent of the mass-sheet degeneracy in the weak lensing limit. Second and third moments of $M_{\rm ap}$ in the cosmic shear context (see next section) will be used extensively in this work.

1.4 Cosmic shear

The term *cosmic shear* describes the coherent distortion of light coming from high-redshift galaxies caused by matter inhomogeneities on very large cosmological scales. The distortions are very minute, of the order of a few percent and therefore much smaller than the typical intrinsic ellipticity of a galaxy. For that reason, cosmic shear has to be detected in a statistical way, using high-quality images of a large number of faint galaxies.

Cosmic shear has been detected for the first time nearly simultaneously by four independent groups in 2000 (Bacon et al. 2000; Kaiser et al. 2000; van Waerbeke et al. 2000; Wittman et al. 2000). Since these pilot studies with the main goal to demonstrate that measurements of cosmic shear are possible and feasible, weak gravitational lensing by the large-scale matter distribution in the Universe has become an important tool for cosmology. On the one hand, the observed sky area and thus the number of faint background galaxies increased dramatically with the advent of wide-field imaging cameras mounted onto large telescopes, both ground-based and space-based. On the other hand, measurement errors have decreased with further understanding of systematics together with new image analysis methods. These two advances were crucial in the evolution of cosmic shear towards a high-precision cosmology probe.

Deep cosmic shear surveys out to redshifts beyond unity covering sky areas from a few dozen square arc minutes to more than a hundred square degrees have yielded measurements of cosmological parameters without the need for modeling the relation between luminous and dark matter (bias). Most notably, interesting constraints on Ω_m and σ_8 have been derived (Hoekstra et al. 2002; Jarvis et al. 2003; van Waerbeke et al. 2005). Cosmic shear is probably the most important tool in order to determine the dark matter power spectrum normalization σ_8. Its superiority over other techniques becomes evident by noting that the method works without having to model the bias between galaxies and dark matter as is required for galaxy surveys. It also probes the normalization on scales which are relevant for non-linear structure formation in the recent past in contrast with CMB anisotropy experiments which are sensitive on much larger, linear scales.

Depending on the depth of the survey, cosmic shear is sensitive to inhomogeneities in the projected matter distribution out to redshifts of order unity. It probes scales where fluctuations started to grow non-linearly due to gravitational instabilities. These non-linearities along with projection effects erase most of the primordial features such as baryon wiggles in the power spectrum due to the acoustic waves in the electron-photon plasma prior to decoupling (Sect. 1.1.8). Thus, cosmological parameters cannot be determined uniquely from cosmic shear; there exist substantial near-degeneracies, e.g. between Ω_m and σ_8, or Γ and n_s.

Some of these near-degeneracies can be broken on combining weak lensing with other cosmological measurements like CMB anisotropy experiments, the statistics of the Lyman-α forest or galaxy redshift surveys. The parameter dependencies are very different for the individual methods, for example the Ω_m-σ_8-degeneracy is nearly orthogonal between cosmic shear and CMB (van Waerbeke et al. 2002). Consequently, even the most precise measurements of cosmological parameters can be improved substantially when weak lensing data is added (Contaldi et al. 2003; Hu & Tegmark 1999).

These degeneracies manifest themselves in a different way also for shear statistics of different order, and they can be lifted by combining e.g. second- and third-order statistics. An example is the reduced skewness of the convergence κ, which is a (non-linear) combination of the skewness and the variance of κ. This quantity has been shown to not, or only weakly, depend on σ_8 and thus to be able to break the near-degeneracy with Ω_m (Bernardeau et al. 1997; van Waerbeke et al. 1999). An important result of the present work is that degeneracies between cosmological parameters, most notably between Ω_m and σ_8, can be lifted substantially by linearly combining second- and third-order statistics of cosmic shear (Chapter 5).

For cosmic shear, the thin lens approximation fails, because the light of background galaxies is distorted continuously on its path by intervening matter. Therefore, a rigorous description of the propagation of light is needed. In the following, I use the notation from Bartelmann & Schneider (2001).

1.4.1 Light propagation

The propagation of light bundles through a general space-time manifold is described by the *geodesic deviation equation*, which relates the deviation of two neigboring geodesics to the space-time geometry via the Riemann tensor (for a derivation, see e.g. Peebles 1993). Using this equation, one can derive a propagation equation for the comoving transverse separation $\boldsymbol{x}$ of neighboring light rays in an infinitesimally thin light bundle. In a Friedmann-Lemaître Universe, this is simply

$$\frac{\mathrm{d}^2\boldsymbol{x}}{\mathrm{d}w^2} + K\boldsymbol{x} = 0. \tag{1.82}$$

Here, w is the comoving distance along the light ray and K is the spatial curvature of the Universe. All rays of the light bundle have a common intersection point, which is the position of the observer at $w = 0$. The solution of (1.82) is $\boldsymbol{x}(\boldsymbol{\theta}, w) = f_K(w)\boldsymbol{\theta}$, where $\boldsymbol{\theta}$ is the angle of two light rays separated by $\boldsymbol{x}$, as seen by the observer.

Next, one adds density perturbations to the cosmological background. It is assumed that their Newtonian potentials Φ are weak, $|\Phi| \ll c^2$, which is fulfilled for large-scale density fluctuations. Furthermore, the largest perturbations have to be much smaller than the horizon which is also satisfied: the density fluctuation power spectrum has a very small amplitude on large scales. Applying Fermat's principle to the light rays, one finds that an additional term $-2/c^2\,\nabla_\perp\Phi$ has to be added to the right-hand side of equation (1.82), where $\nabla_\perp$ denotes differentiation perpendicular to the light path. This modified propagation equation describes the deviation of a light ray in a perturbed Universe with respect to that in an unperturbed one. Since in reality *all* light rays are perturbed, one has to replace the gradient of the potential by the *difference* of potential gradients $\Delta(\nabla_\perp\Phi)$ between points of two neighboring light rays. Equation (1.82) then becomes

$$\frac{\mathrm{d}^2\boldsymbol{x}}{\mathrm{d}w^2} + K\boldsymbol{x} = -\frac{2}{c^2}\Delta\Big(\nabla_\perp\Phi(\boldsymbol{x}(\boldsymbol{\theta}, w), w)\Big). \tag{1.83}$$

The comoving separation of the light rays $\boldsymbol{x}$ depends on $\boldsymbol{\theta}$ and the comoving distance w, the potential depends on $\boldsymbol{x}$ and w.

Using Green's function, (1.83) can be written as an integral equation,

$$\boldsymbol{x}(\boldsymbol{\theta}, w) = f_K(w)\boldsymbol{\theta} - \frac{2}{c^2}\int_0^w dw' f_K(w - w')\Delta\Big(\nabla_\perp\Phi(\boldsymbol{x}(\boldsymbol{\theta}, w'), w')\Big). \tag{1.84}$$

In order to simplify this equation, the zeroth-order solution, $\boldsymbol{x}(\boldsymbol{\theta}, w) = f_K(w)\,\boldsymbol{\theta}$, is inserted into the right-hand side, which is equivalent to integrating over the potential along the unperturbed ray. This is the analogy of the *Born approximation* used in scattering theory and permits to interchange the potential gradient differences by the gradient of the differences of the potential. But taking the potential difference boils down to adding a term independent of $\boldsymbol{\theta}$, the perpendicular gradient of which vanishes. Therefore, one replaces $\Delta\Phi$ by Φ. One can now introduce a quantity analogous to the deflection angle (see Sect. 1.3.1), which is the unperturbed separation minus the real separation, divided by the angular diameter distance, hence

$$\boldsymbol{\alpha}(\boldsymbol{\theta}, w) = \frac{f_K(w)\boldsymbol{\theta} - \boldsymbol{x}(\boldsymbol{\theta}, w)}{f_K(w)} = \frac{2}{c^2}\int_0^w \mathrm{d}w'\,\frac{f_K(w - w')}{f_K(w)}\nabla_\perp\Phi(f_K(w')\boldsymbol{\theta}, w'). \tag{1.85}$$

An effective convergence can be defined which depends on the comoving distance, via $2\kappa(\boldsymbol{\theta}, w) = \nabla_\theta\,\boldsymbol{\alpha}(\boldsymbol{\theta}, w)$. Applied to eq. (1.85), the differentiation only acts on the potential,

resulting in the two-dimensional Laplacian of ψ. One can add at each point the second derivative of the potential with respect to the line-of-sight direction to get the three-dimensional Laplacian. Integrated along the light path, this extra term averages to zero (see White & Hu (2000) for a validation of this assumption). This Laplacian can then be written in terms of the linear density perturbations using the Poisson equation (1.20), and one gets

$$\kappa(\boldsymbol{\theta}, w) = \frac{3H_0^2\Omega_\mathrm{m}}{2c^2} \int_0^w \mathrm{d}w' \frac{f_K(w-w')f_K(w')}{f_K(w)a(w')} \delta\left(f_K(w')\boldsymbol{\theta}, w\right). \tag{1.86}$$

Although the thin-lens approximation is no longer valid for cosmic shear as mentioned before, one can think of the three-dimensional perturbations as a stack of multiple lens planes, each plane being a thin lens. Then, using the Born approximation and retaining only the first-order terms in Φ is equivalent to adding up the contributions to the convergence of all lens planes linearly. The higher-order terms which describe the coupling of lens planes are neglected. One can introduce a deflection potential

$$\psi(\boldsymbol{\theta}, w) = \frac{2}{c^2} \int_0^w \mathrm{d}w' \frac{f_K(w-w')}{f_K(w')f_K(w)} \Phi(f_K(w')\boldsymbol{\theta}, w'), \tag{1.87}$$

with fulfills $\kappa = \nabla^2\psi/2$ (cf. 1.63), and define the shear $\gamma_1 = \left(\partial_1\partial_1\psi - \partial_2\partial_2\psi\right)/2$ and $\gamma_2 = \partial_1\partial_2\psi$ according to (1.65). Given the weak potentials on large scales and small deflection angles, the Born approximation is justifiable. The validity of this treatment has also been verified by ray-tracing light rays through large-scale N-body simulations, see e.g. Jain et al. (2000).

In order to get the total convergence, one integrates the effective convergence (1.86), weighted by the distribution function of observed source galaxies in comoving distance $p(w)\mathrm{d}w$, up to some sufficiently large distance w_lim at which p basically drops to zero, and obtains

$$\kappa(\boldsymbol{\theta}) = \int_0^{w_\mathrm{lim}} \mathrm{d}w\, p(w)\, \kappa(\boldsymbol{\theta}, w) = \int_0^{w_\mathrm{lim}} \mathrm{d}w\, G(w)\, f_K(w)\, \delta\left(f_K(w)\boldsymbol{\theta}, w\right), \tag{1.88}$$

with the *lens efficiency* function

$$G(w) = \frac{3}{2}\left(\frac{H_0}{c}\right)^2 \frac{\Omega_\mathrm{m}}{a(w)} \int_w^{w_\mathrm{lim}} \mathrm{d}w'\, p(w') \frac{f_K(w-w')}{f_K(w')}. \tag{1.89}$$

In this work, three fiducial cosmological models (App. B) with two different source redshift distributions are employed. For model 1 (Table B.1), the following realistic redshift distribution, taken from Brainerd et al. (1996), is used,

$$p(w)\mathrm{d}w = p(z)\mathrm{d}z = \frac{\beta}{z_0\Gamma\left(3/\beta\right)} \left(\frac{z}{z_0}\right)^2 \mathrm{e}^{-(z/z_0)^\beta} \mathrm{d}z, \tag{1.90}$$

where Γ denotes the Eulerian Gamma function. The parameter β specifies how fast the distribution falls off towards the high-redshift end. z_0 is related to the mean redshift $\bar{z}$ of the distribution; for $\beta = 1.5$, it is $\bar{z} \approx 1.5\, z_0$. Both model 2 and model 3 assume a single redshift plane for the source galaxies at $z \approx 1$ corresponding to the ray-tracing simulations.

The product $G \cdot f_K$ determines the redshift sampling of the density field δ in the projection for the convergence κ (1.88). This integration kernel follows the source galaxy distribution but

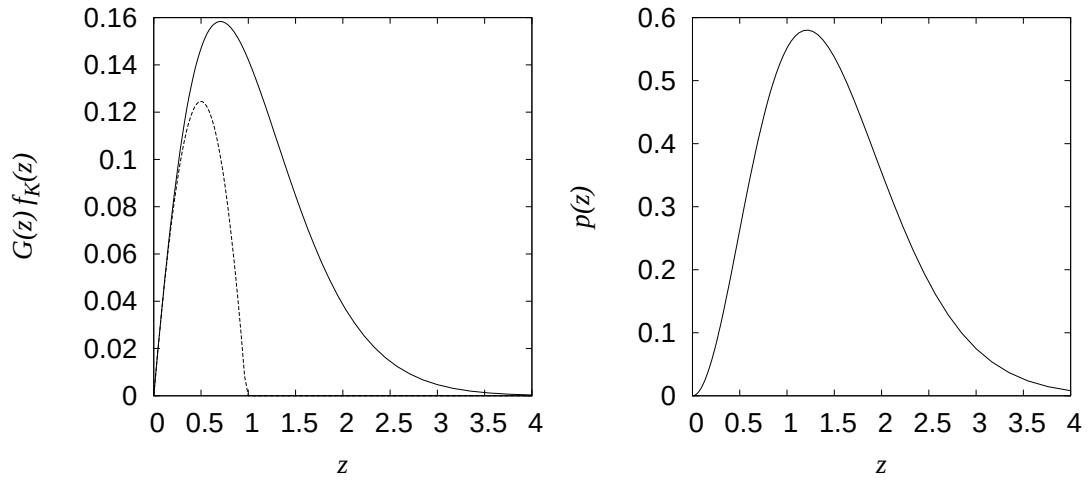

Figure 1.8: *The integration kernel $G \cdot f_K$ in (1.88) as a function of the redshift z is shown in the left panel for the two ΛCDM cosmologies model 1 (solid line) and model 2 (dashed line), see Table B.1. Model 1 assumes a continuous redshift distribution (1.90) with $\beta = 1.5$ and $z_0 = 1$, which is shown in the right panel. For model 2, all source are situated at redshift of about unity.*

also depends on the geometrical factors f_K, which are the direct correspondence of the distances between source, lens and observer in the single thin lens case (Sect. 1.3.2). Accordingly, if all source galaxies are at the same redshift, as it is the case for models 2 and 3, intermediate redshifts between observer and sources contribute most to the effective convergence. The lens efficiency and the source distribution are plotted in Fig. 1.8.

1.4.2 Moments of the convergence

Since the convergence κ is a linear projection of the density contrast δ, it also has to be described as a random field. Consequently, the power spectrum and the bispectrum of κ can be calculated from the corresponding Fourier moments of δ.

1.4.2.1 The convergence power spectrum

The convergence power spectrum is defined in (1.71) and can be expressed in terms of the power spectrum of the density perturbations using Limber's equation (Limber 1953) in Fourier space (Kaiser 1992),

$$P_\kappa(\ell) = \int \mathrm{d}w\, G^2(w) P_\delta\left(\frac{\ell}{f_K(w)}\right). \tag{1.91}$$

The dimensionless power spectrum $\ell^2 P_\kappa(\ell)$ is plotted in Fig. 1.9. The cosmological and redshift distribution parameters can be divided into groups, depending on how they influence the power spectrum. One group, consisting of $\Omega_\mathrm{m}, \sigma_8, z_0$ and β roughly cause a shift only of the amplitude. A change in Γ, n_s or Ω_Λ results in a tilt of the power spectrum. Since the functional dependence of the power spectrum is very similar for each of the parameters in one group, these parameters are highly degenerate. These relationships and degeneracies between parameters will be analyzed in more detail in Chapters 3 and 5.

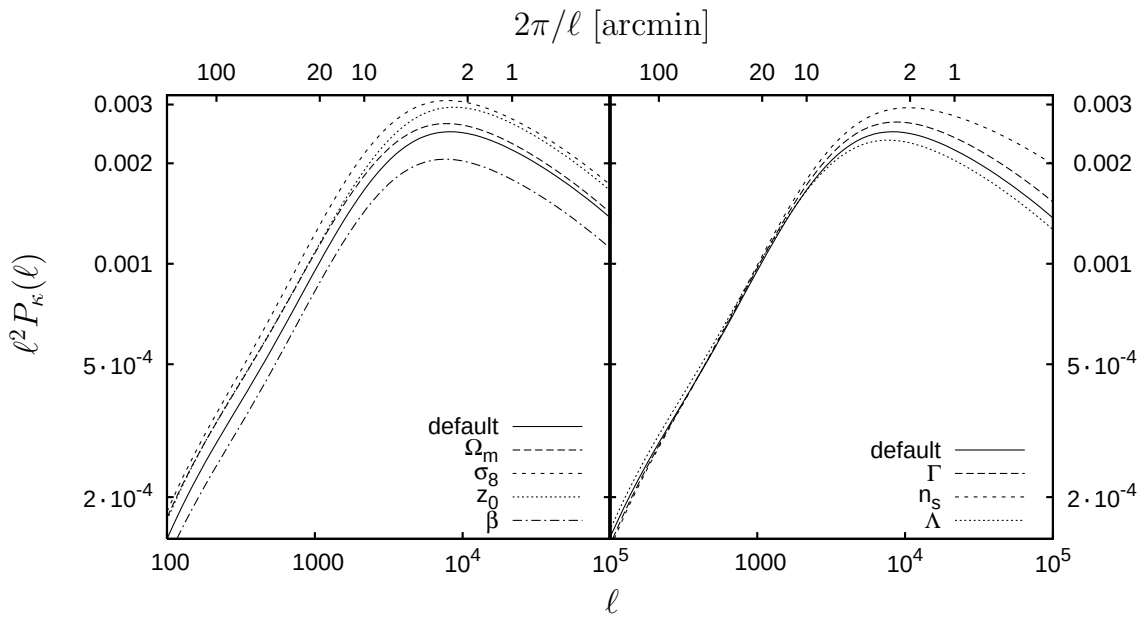

Figure 1.9: *The dependence of the dimensionless convergence power spectrum $\ell^2 P_\kappa(\ell)$ on cosmological parameters. For each curve, the corresponding parameter as indicated in the panel is increased by 10% with respect to the default ΛCDM model (model 1, see Table B.1). The conjugate angle $2\pi/\ell$ of the Fourier scale ℓ is given on the top axes.*

1.4.2.2 The convergence bispectrum

The bispectrum B_κ of the convergence is defined by the following equation, in analogy to (1.42):

$$\langle\hat{\kappa}(\boldsymbol{\ell}_1)\hat{\kappa}(\boldsymbol{\ell}_2)\hat{\kappa}(\boldsymbol{\ell}_3)\rangle = (2\pi)^2\delta_{\rm D}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3)\left[B_\kappa(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2)+B_\kappa(\boldsymbol{\ell}_2,\boldsymbol{\ell}_3)+B_\kappa(\boldsymbol{\ell}_3,\boldsymbol{\ell}_1)\right]. \quad (1.92)$$

The convergence inherits the statistical isotropy and parity-invariance from the density contrast, thus the bispectrum only depends on the scalar product of its two arguments, $B_\kappa(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2) = B_\kappa(\boldsymbol{\ell}_1\cdot\boldsymbol{\ell}_2) = b_\kappa(\ell_1,\ell_2,\varphi)$. I write

$$b_\kappa(\ell_1,\ell_2,\varphi) = \sum_{m=0}^{2} F_2^{(m)}(\ell_1,\ell_2)\cos^m(\varphi)\,\bar{b}_\kappa^{(m)}(\ell_1,\ell_2), \quad (1.93)$$

where the coefficients $F_2^{(m)}$ are taken either from PT (1.34) or from HEPT (1.49). The functions $\bar{b}_\kappa^{(m)}$ are projections of the 3-D bispectrum of density fluctuations δ. The projection is obtained using Limber's equation, applied to the bispectrum (App. C). Inserting the result of this projection (C.8) for the convergence bispectrum (1.92) and using the second-order relation between the bispectrum and the power spectrum (1.44) yields

$$\bar{b}_\kappa^{(m)}(\ell_1,\ell_2) = \int_0^{w_{\rm hor}} \frac{{\rm d}w}{f_K(w)} G^3(w) f^{(m)}(w,\ell_1) f^{(m)}(w,\ell_2) P_\delta\left(\frac{\ell_1}{f_K(w)}\right) P_\delta\left(\frac{\ell_2}{f_K(w)}\right). \quad (1.94)$$

If PT is used as a model for the δ-bispectrum, the coefficients $f^{(m)}$ for $m=0,1,2$, are equal to unity, and P_δ corresponding to the linear power spectrum (Sect. 1.2.6.1) has to be inserted.

For the non-linear model HEPT (Scoccimarro & Couchman 2001) with the modified kernel (1.49), one has to replace $f^{(m)}$, $m = 0, 1, 2$, with the fitting functions a, b and c respectively, see Sect. 1.2.7.2. Note that the coefficients $F_2^{(m)}$ only depend on the ratio of its arguments, therefore they can be drawn out of the integration over the comoving distance w justifying the form of the bispectrum given in (1.93).

1.4.3 The E- and B-mode of the shear field

From the relations between the shear, the convergence and the gravitational potential, one can derive the following relation (Kaiser 1995; Schneider et al. 2002),

$$\nabla\kappa = \begin{pmatrix} \partial_1\gamma_1 + \partial_2\gamma_2 \\ \partial_2\gamma_1 - \partial_1\gamma_2 \end{pmatrix} = \boldsymbol{u}, \tag{1.95}$$

which defines the vector $\boldsymbol{u}$ as the gradient of the "potential" κ. However, when measuring $\boldsymbol{u}$ from data, it will have a non-gradient component due to noise and systematic measurement errors. Moreover, other sources than measurement effects can cause a non-zero curl component of $\boldsymbol{u}$ which will be discussed below.

In order to separate the gradient from the curl part of $\boldsymbol{u}$, one introduces the new quantities κ^{E} and κ^{B}, which represent the so-called E- and B-mode of the convergence (in analogy to the electromagnetic field) by

$$\begin{aligned} \nabla^2\kappa^{\mathrm{E}} &= \nabla\cdot\boldsymbol{u}; \\ \nabla^2\kappa^{\mathrm{B}} &= \nabla\times\boldsymbol{u} = \partial_1 u_2 - \partial_2 u_1. \end{aligned} \tag{1.96}$$

Further, one defines the E- and B-mode potentials ψ^{E} and ψ^{B} via the Poisson equation

$$\nabla^2\psi^{\mathrm{E,B}} = 2\,\kappa^{\mathrm{E,B}}. \tag{1.97}$$

For notation and calculation simplification, the E- and B-modes are combined into complex quantities,

$$\psi = \psi^{\mathrm{E}} + \mathrm{i}\,\psi^{\mathrm{B}}, \quad \kappa = \kappa^{\mathrm{E}} + \mathrm{i}\,\kappa^{\mathrm{B}}. \tag{1.98}$$

The Poisson equation (1.97) then takes its simple form $\nabla^2\psi = 2\,\kappa$, and the complex shear $\gamma = \frac{1}{2}\left(\partial_1\partial_1 - \partial_2\partial_2\right)\psi + \mathrm{i}\,\partial_1\partial_2\psi$ can be written as

$$\gamma_1 + \mathrm{i}\,\gamma_2 = \frac{1}{2}\left(\partial_1\partial_1\psi^{\mathrm{E}} - \partial_2\partial_2\psi^{\mathrm{E}}\right) - \partial_1\partial_2\psi^{\mathrm{B}} + \mathrm{i}\left[\partial_1\partial_2\psi^{\mathrm{E}} + \frac{1}{2}\left(\partial_1\partial_1\psi^{\mathrm{B}} - \partial_2\partial_2\psi^{\mathrm{B}}\right)\right]. \tag{1.99}$$

These definitions are consistent with eqs. (1.63) and (1.65) in the absence of B-modes.

The shear and convergence power spectrum as defined in (1.71) can be generalized to a field which has both an E- and a B-mode. One defines the E-mode, the B-mode and the mixed power spectra of the convergence as follows,

$$\begin{aligned} \mathrm{EE} &= \langle\hat{\kappa}^{\mathrm{E}}(\boldsymbol{\ell})\hat{\kappa}^{\mathrm{E}}(\boldsymbol{\ell}')\rangle = (2\pi)^2\delta_{\mathrm{D}}(\boldsymbol{\ell}-\boldsymbol{\ell}')P_\kappa^{\mathrm{E}}(\ell), \\ \mathrm{BB} &= \langle\hat{\kappa}^{\mathrm{B}}(\boldsymbol{\ell})\hat{\kappa}^{\mathrm{B}}(\boldsymbol{\ell}')\rangle = (2\pi)^2\delta_{\mathrm{D}}(\boldsymbol{\ell}-\boldsymbol{\ell}')P_\kappa^{\mathrm{B}}(\ell), \\ \mathrm{EB} &= \langle\hat{\kappa}^{\mathrm{E}}(\boldsymbol{\ell})\hat{\kappa}^{\mathrm{B}}(\boldsymbol{\ell}')\rangle = (2\pi)^2\delta_{\mathrm{D}}(\boldsymbol{\ell}-\boldsymbol{\ell}')P_\kappa^{\mathrm{EB}}(\ell), \end{aligned} \tag{1.100}$$

and can derive for the correlators of γ in Fourier space

$$\begin{aligned}\langle\hat{\gamma}(\boldsymbol{\ell})\hat{\gamma}^*(\boldsymbol{\ell})\rangle &= (2\pi)^2\delta_{\mathrm{D}}(\boldsymbol{\ell}-\boldsymbol{\ell}')[P_\kappa^{\mathrm{E}}(\ell)+P_\kappa^{\mathrm{B}}(\ell)],\\ \langle\hat{\gamma}(\boldsymbol{\ell})\hat{\gamma}(\boldsymbol{\ell})\rangle &= (2\pi)^2\delta_{\mathrm{D}}(\boldsymbol{\ell}-\boldsymbol{\ell}')[P_\kappa^{\mathrm{E}}(\ell)-P_\kappa^{\mathrm{B}}(\ell)+2\mathrm{i}P_\kappa^{\mathrm{EB}}(\ell)].\end{aligned} \tag{1.101}$$

A similar generalization can be done for the convergence bispectrum (1.92) and the terms (in the above introduced short notation) EEE, EEB, EBB and BBB can be defined. Schneider (2003) has shown that under parity transformation (see also Sect. 2.3.1.2) ψ^{E} is invariant and ψ^{B} changes sign. The same holds for κ and therefore, since the large-scale structure is parity-symmetric, the above defined correlators EB, EEB and BBB which have an odd number of 'B's vanish.

In principle, there are three different sources of a B-mode in the shear field. Firstly, it can arise from measurement errors such as imperfect PSF anisotropy correction or inaccurate determination of the galaxy shape parameters. Secondly, lensing itself is a cause: source redshift clustering (Schneider et al. 2002) and lens-lens coupling which is not accounted for in the Born approximation of the deflection angle (1.85), see Schneider et al. (1998), can introduce a B-mode. However, these higher-order effect are assumed to be very small; their contribution to the signal can most probably be neglected even for the nowadays achieved measurement accuracies. Thirdly, the intrinsic correlation of galaxy orientations, also called *intrinsic alignment*, is supposed to show up in both the E- and the B-mode shear signal. Galaxies can be intrinsically aligned, even over large physical distances, by spin-correlations with their (common) dark-matter halo or by tidal torques (see e.g. Catelan et al. 2001; Crittenden et al. 2001; Croft & Metzler 2000; Heavens et al. 2000; Jing 2002). The mechanisms are still not very well understood and consequently, theoretical predictions of this effect differ by orders of magnitudes (e.g. Brown et al. 2002, and references therein). Observations have not yielded stringent constraints on this effect — moreover, basically no intrinsic alignment was found in the SDSS[10] galaxy sample (Hirata et al. 2004). For a shear survey where no redshift information of the source galaxies is available, galaxy pairs which are close on the sky have a smaller probability to be also physically close the deeper the survey is. Consequently, the intrinsic alignment contamination to shear diminishes with the survey depth. Since the dependence on angular separation is different for shear and intrinsic correlation, it is possible to separate the two effects (King & Schneider 2002, 2003).

It is believed that for a deep cosmic shear survey the main if not only significant source of a B-mode in the shear field are measurement errors. The amplitude of the residual B-mode can therefore be interpreted as a quality attribute of the data reduction. Recently, a re-analysis of the VIRMOS-DESCART data using a refined PSF correction method succeeded in eliminating the B-mode whose presence in previous analyses of the data was most probable due to an incorrect PSF model (van Waerbeke et al. 2005).

[10]Sloan Digital Sky Survey, `http://www.sdss.org`

Chapter 2

Cosmic shear statistics

In this chapter I introduce and define the basic quantities which are the prerequisites for the treatment of cosmic shear statistics (Sect. 2.1). Several second- and third-order shear statistics and their interrelations are defined in Sects. 2.2 and 2.3. The dependence of these statistics on the underlying Fourier-space moments of the convergence, which are the power spectrum and the bispectrum, respectively, are given. These relations will be used in Chapter 4 in order to compare third-order shear statistics from theoretical models of the large-scale structure with numerical ray-tracing simulations. Further, in Chapters 3 and 5, these relations are essential since they determine the dependence of second- and third-order cosmic shear statistics on cosmological parameters.

In Sects. 2.4 and 2.5, estimators and their covariance, respectively, of the shear statistics are discussed. These expressions will be used in Chapters 3 and 5 to quantitatively address the problem of constraining cosmological parameters using cosmic shear measurements.

2.1 Definitions of the basic quantities of cosmic shear statistics

Projection of shear. Shear is a two-component quantity, which transforms like a spin-2 or polar field under rotations. It can be written as a complex number $\gamma = \gamma_1 + \mathrm{i}\gamma_2 = |\gamma| \exp(2\mathrm{i}\phi)$, where γ_1 and γ_2 are the two Cartesian components, $|\gamma|$ the modulus and ϕ the polar angle of the shear (1.65). In Fig. 2.1 the orientation of the shear illustrated as an ellipse is shown as a function of its Cartesian coordinates.

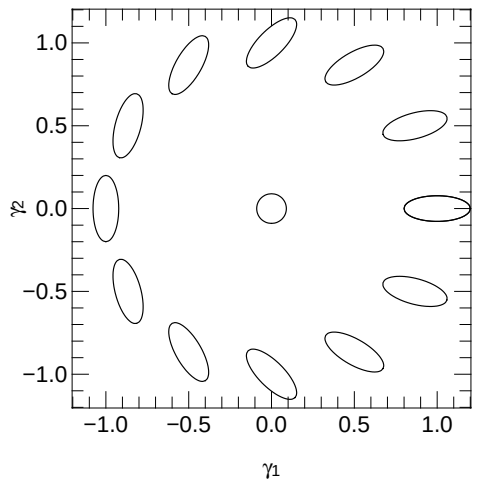

Figure 2.1: *The orientation of the ellipses given by the Cartesian coordinates γ_1 and γ_2 of the shear. While the polar angle φ passes through the range $[0;\, 2\pi]$, the shear ellipse rotates around π.*

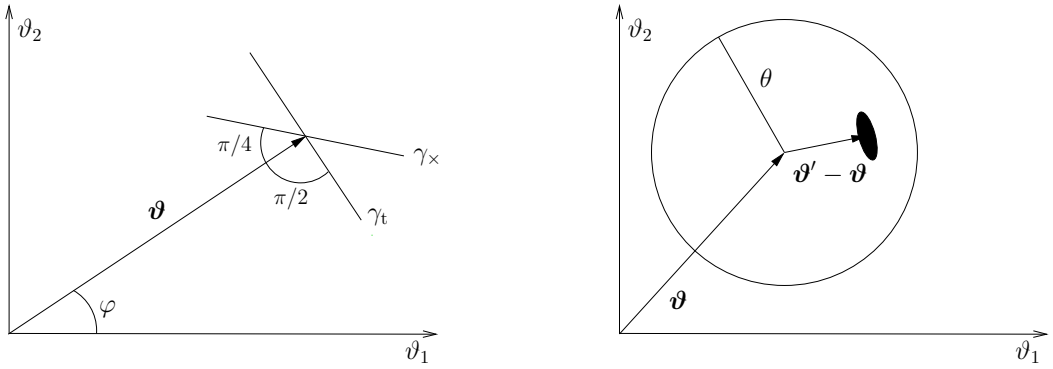

Figure 2.2: Left panel: *The tangential and cross-component of the shear, as defined in eq. (2.1).* Right panel: *Sketch for the definition of the aperture mass statistics (2.2).*

A very useful representation of shear is the decomposition of γ into its *tangential* and *cross*-component relative to some direction $\boldsymbol{\vartheta}$,

$$\gamma_\mathrm{t} = -\,\Re\left(\gamma \mathrm{e}^{-2\mathrm{i}\varphi}\right) \qquad \text{and} \qquad \gamma_\times = -\,\Im\left(\gamma \mathrm{e}^{-2\mathrm{i}\varphi}\right), \tag{2.1}$$

where φ is the polar angle of the vector $\boldsymbol{\vartheta}$, see Fig. 2.2. Around a spherical mass overdensity, the shear field shows a tangential pattern, thus background galaxies are tangentially aligned with respect to the mass center. In the above defined notation, this corresponds to $\gamma_\mathrm{t} > 0$ and $\gamma_\times = 0$.

The aperture mass statistics. Linear combinations of shear which are scalar quantities can be found and will turn out to be very useful in various aspects. One important measure of shear and the convergence is the aperture mass M_ap which was first introduced by Kaiser et al. (1994) and Schneider (1996) as a quantity which is not affected by the mass-sheet degeneracy (Sect. 1.3.6). In the following years its relevance and usefulness for cosmic shear was discovered. M_ap at a point $\boldsymbol{\vartheta}$ is defined as the integral over the filtered surface mass density κ in an aperture of radius θ, centered at $\boldsymbol{\vartheta}$. Alternatively it can be expressed in terms of the tangential shear γ_t (2.1), where 'tangential' is now understood with respect to the aperture center $\boldsymbol{\vartheta}$ (Kaiser et al. 1994; Schneider 1996),

$$M_\mathrm{ap}(\theta, \boldsymbol{\vartheta}) = \int \mathrm{d}^2\vartheta'\, U_\theta(|\boldsymbol{\vartheta} - \boldsymbol{\vartheta}'|)\kappa(\boldsymbol{\vartheta}') = \int \mathrm{d}^2\vartheta'\, Q_\theta(|\boldsymbol{\vartheta} - \boldsymbol{\vartheta}'|)\gamma_\mathrm{t}(\boldsymbol{\vartheta}'), \tag{2.2}$$

the second equality holds if U_θ is a compensated filter function, i.e.

$$\int \mathrm{d}\vartheta\, \vartheta\, U_\theta(\vartheta) = 0. \tag{2.3}$$

Obviously, this property of U_θ leaves M_ap invariant to a constant additive surface mass density and is therefore insensitive to the mass-sheet degeneracy in the weak lensing regime ($\kappa \ll 1$). Further, Q is obtained from U by

$$Q_\theta(\vartheta) = \frac{2}{\vartheta^2}\int_0^\vartheta \mathrm{d}\vartheta'\, \vartheta'\, U_\theta(\vartheta') - U_\theta(\vartheta). \tag{2.4}$$

Here and in the following I assume, if not stated otherwise, that the convergence is free of a B-mode, $\kappa = \kappa^{\mathrm{E}}$, and omit the subscript 'E'.

It is the second equality in (2.2) which makes the aperture mass so useful from an observational point of view. Accordingly, the total (weighted) projected mass density κ in a region on the sky can be estimated by averaging over the (weighted) tangential ellipticities of background galaxies in this region, since the observed ellipticity is an estimator of the shear in the weak-lensing limit (1.79). An aperture mass-map can be compiled for a sky region, where each point $\boldsymbol{\vartheta}$ is assigned the value of $M_{\mathrm{ap}}(\theta, \boldsymbol{\vartheta})$ for some filter scale θ, see Fig. 5.1. Mass overdensities such as galaxy clusters manifest themselves as peaks in the M_{ap}-map. With this method, not only clusters which have been found via galaxy overdensities or X-ray emission can be confirmed, but blind searches can be done to identify galaxy clusters (e.g. Schirmer et al. 2004).

Analogous to the second part of (2.2), the weighted cross-component of shear in an aperture can be defined as

$$M_{\perp}(\theta, \boldsymbol{\vartheta}) = \int \mathrm{d}^2\vartheta\, Q_\theta(|\boldsymbol{\vartheta} - \boldsymbol{\vartheta}'|)\gamma_\times(\boldsymbol{\vartheta}), \tag{2.5}$$

the corresponding "convergence" is κ^{B} (1.98) which, however, does not represent a physical surface mass density. $M_\perp$ is a measure of the B-mode, and can be used to quantify systematic measurement errors (see Sect. 1.4.3), since the equality $M_\perp(\theta, \boldsymbol{\vartheta}) = 0$ holds for every $\boldsymbol{\vartheta}$ in the absence of a B-mode. Note that this equality is not satisfied when the aperture mass is estimated using observed background galaxy ellipticities as tracers of the shear field because of noise. However, the expectation value of $M_\perp$ is still zero for every $\boldsymbol{\vartheta}$ when no B-mode is present.

The integrals (2.2) and (2.5) can be written as convolutions,

$$\begin{aligned} M_{\mathrm{ap}}(\theta, \boldsymbol{\vartheta}) &= (U_\theta * \kappa)(\boldsymbol{\vartheta}) = \Re\left(Q'_\theta * \gamma\right)(\boldsymbol{\vartheta}), \\ M_\perp(\theta, \boldsymbol{\vartheta}) &= \Im\left(Q'_\theta * \gamma\right)(\boldsymbol{\vartheta}), \end{aligned} \tag{2.6}$$

where the modified filter function Q'_θ is defined as follows,

$$Q'_\theta(\vartheta) = -Q_\theta(\vartheta)\mathrm{e}^{-2\mathrm{i}\arctan(\vartheta_2/\vartheta_1)}. \tag{2.7}$$

One can combine (2.2) and (2.5) and define the complex aperture mass

$$M(\theta, \boldsymbol{\vartheta}) = M_{\mathrm{ap}}(\theta, \boldsymbol{\vartheta}) + \mathrm{i}M_\perp(\theta, \boldsymbol{\vartheta}) = \left(Q'_\theta * \gamma\right)(\boldsymbol{\vartheta}). \tag{2.8}$$

Two sets of filter functions for U and the derived function Q (2.4) have been used frequently in previous studies; both will be employed for this work. With the definition $U_\theta(\vartheta) = u(\vartheta/\theta)/\theta^2$ and $Q_\theta(\vartheta) = q(\vartheta/\theta)/\theta^2$, the first set of functions considered here are the polynomial filter functions defined in Schneider et al. (1998),

$$u(x) = \frac{9}{\pi}(1-x^2)\left(\frac{1}{3} - x^2\right)\mathrm{H}(1-x), \quad q(x) = \frac{6}{\pi}x^2(1-x^2)\mathrm{H}(1-x), \tag{2.9}$$

where H is the Heaviside step function, $\mathrm{H}(x) = 1$ for $x > 0$ and 0 elsewhere. These functions have the nice property of compact support, thus the integrals (2.2) are performed over a finite region. Another set of functions was defined by Crittenden et al. (2002),

$$u(x) = \frac{1}{2\pi}\left(1 - \frac{x^2}{2}\right)\mathrm{e}^{-\frac{x^2}{2}}, \quad q(x) = \frac{x^2}{4\pi}\mathrm{e}^{-\frac{x^2}{2}}. \tag{2.10}$$

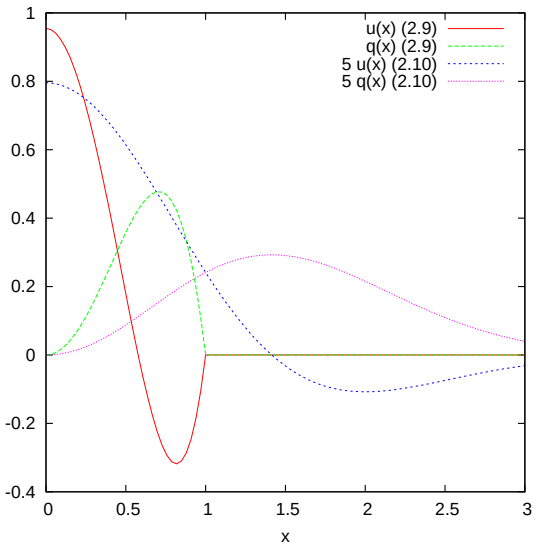

Figure 2.3: *The polynomial (2.9) and Gaussian (2.10) filter functions for $M_{\rm ap}$.*

Both of these Gaussian filter functions decrease exponentially and are significantly non-zero up to about three times the aperture radius θ. Therefore, the formally infinite integrals (2.2) can be evaluated over a finite range with sufficient accuracy, however, the integration range is much more extended than for the polynomial filter. All four filter functions are plotted in Fig. 2.3.

2.2 Second-order statistics

The ensemble average of the convergence κ and the shear γ caused by matter inhomogeneities on large scales vanishes because $\langle\delta\rangle = 0$ (Sect. 1.2.6) and because the large-scale structure is statistically isotropic. The same holds for all linear combinations of κ and γ such as the aperture mass statistics (2.2). For all cosmological experiments, the ensemble average of a quantity is replaced by a spatial average over a sufficiently large volume or sky area. This corresponds to the ergodicity hypothesis, see App. A.

If averaged over only a small area, $\langle\kappa\rangle$ and $\langle\gamma\rangle$ will have values different from zero, representing local over- or underdensities. For cosmic shear, these fluctuations typically are much smaller than the noise due to the intrinsic ellipticities of the observed galaxies. Most of the cosmic shear studies do not aim to map the distribution of the projected mass in some sky region, but rather try to determine its statistical properties. The first non-trivial moment of κ and γ is the second-order moment, and most observations up to now have focused on second-order shear statistics. The detection of third and higher-order moments is very challenging. Their amplitude is very small, since the deviation of κ and γ from a Gaussian field is small on the scales where cosmic shear is observed. Only recently with the up-coming of large and deep surveys have higher-order statistics and the non-Gaussianity of the shear field started to play a significant role for cosmic shear.

2.2.1 The two-point correlation function

Given a pair of angular positions on the sky $\boldsymbol{\vartheta}$ and $\boldsymbol{\vartheta}+\boldsymbol{\theta}$, the tangential and cross-component of the shear (2.1) with respect to the connecting vector $\boldsymbol{\theta}$ is the natural representation of

shear. The three components of the two-point correlation function (2PCF) are defined as the following linear combinations of products of γ_t and $\gamma_\times$,

$$\begin{aligned}\xi_+(\theta) &= \langle\gamma_\mathrm{t}(\boldsymbol{\vartheta})\gamma_\mathrm{t}(\boldsymbol{\vartheta}+\boldsymbol{\theta})\rangle + \langle\gamma_\times(\boldsymbol{\vartheta})\gamma_\times(\boldsymbol{\vartheta}+\boldsymbol{\theta})\rangle\,,\\ \xi_-(\theta) &= \langle\gamma_\mathrm{t}(\boldsymbol{\vartheta})\gamma_\mathrm{t}(\boldsymbol{\vartheta}+\boldsymbol{\theta})\rangle - \langle\gamma_\times(\boldsymbol{\vartheta})\gamma_\times(\boldsymbol{\vartheta}+\boldsymbol{\theta})\rangle\,,\\ \xi_\times(\theta) &= \langle\gamma_\mathrm{t}(\boldsymbol{\vartheta})\gamma_\times(\boldsymbol{\vartheta}+\boldsymbol{\theta})\rangle = \langle\gamma_\times(\boldsymbol{\vartheta})\gamma_\mathrm{t}(\boldsymbol{\vartheta}+\boldsymbol{\theta})\rangle.\end{aligned} \tag{2.11}$$

Because the large-scale structure is statistically homogeneous and isotropic, the 2PCF only depends on the modulus θ of the connecting vector of the two points. Due to parity symmetry, the mixed 2PCF $\xi_\times$ vanishes, since under parity transformation (Sect. 2.3.1.2), γ_t is invariant and $\gamma_\times$ changes its sign.

The Fourier transform of the 2PCF of shear is the shear power spectrum. Since $P_\gamma = P_\kappa$ (eq. 1.71), the 2PCF can be written in terms of the convergence power spectrum. The more general case of a convergence which contains not only the gradient-part κ^E but also the curl-component κ^B shall be included here, see Sect. 1.4.3. Using (1.101), the 2PCF (2.11) can be written as

$$\begin{aligned}\xi_+(\theta) &= \frac{1}{2\pi}\int \mathrm{d}\ell\,\ell \mathrm{J}_0(\ell\theta)[P_\kappa^\mathrm{E}(\ell) + P_\kappa^\mathrm{B}(\ell)],\\ \xi_-(\theta) &= \frac{1}{2\pi}\int \mathrm{d}\ell\,\ell \mathrm{J}_4(\ell\theta)[P_\kappa^\mathrm{E}(\ell) - P_\kappa^\mathrm{B}(\ell)],\\ \xi_\times(\theta) &= \frac{1}{2\pi}\int \mathrm{d}\ell\,\ell \mathrm{J}_4(\ell\theta)P_\kappa^\mathrm{EB}(\ell),\end{aligned} \tag{2.12}$$

where J_ν is the first-kind Bessel function of order ν. In the absence of a B-mode, $\xi_\pm(\theta) = 1/2\int \mathrm{d}\ell\,\ell P_\kappa^\mathrm{E}(\ell)\mathrm{J}_{0,4}(\ell\theta)$ and $\xi_\times = 0$. In that case I will drop the superscript denoting the E-mode and write $P_\kappa = P_\kappa^\mathrm{E}$.

The 2PCF and the corresponding filter functions J_0 and J_4 are shown in Figs. 2.4 and 2.5, respectively. At the end of the next section, I will discuss various second-order shear statistics and how they sample the convergence power spectrum which is characterized by the corresponding filter function.

It is desirable to possess a clean measurement which separates the E- from the B-mode power spectrum, since (to leading order, see Sect. 1.4.3) only the first contains cosmological information — the latter is usually solely due to unwanted systematic effects. It is possible to define correlation functions $\xi_\pm^\mathrm{E}$ and $\xi_\pm^\mathrm{B}$ which only depend on the E- and B-mode power spectrum, respectively. These quantities are not observable directly but have to be calculated from the above defined 2PCF $\xi_\pm$. However, the knowledge of either ξ_- for arbitrary large angular separations or ξ_+ on arbitrarily small scales is necessary in order to obtain the E- and B-mode correlation functions (Schneider et al. 2002). This is not feasible since data fields are finite and shapes of very close galaxy pairs cannot be determined reliably. In the next section, it is demonstrated that the dispersion of the aperture mass statistics (2.2) provides a useful way to isolate the E- and the B-mode power spectrum (Crittenden et al. 2002).

2.2.2 Second-order aperture mass statistics

The second moment or dispersion of (2.2) was introduced by Schneider et al. (1998) and has been used with great success in numerous cosmic shear surveys (Hamana et al. 2003; Hoekstra et al. 2002; Jarvis et al. 2003; van Waerbeke et al. 2005). Because it separates the E- from the

B-mode, it is an extremely valuable tool to assess systematic measurement errors. Moreover, $\langle M_{\rm ap}^2\rangle$ is a local measure of the power spectrum and therefore very sensitive to cosmological parameters.

Using the definition of the complex aperture mass (2.8), one can define the two second-order quantities $\langle MM\rangle(\theta)$ and $\langle MM^*\rangle(\theta)$, where the ensemble average is replaced by a spatial average over aperture centers $\boldsymbol{\vartheta}$ when these quantities are calculated from data. Using the relations (2.2) leads to

$$\begin{aligned}\langle MM\rangle(\theta) &= \langle M_{\rm ap}^2\rangle(\theta) + \langle M_\perp^2\rangle(\theta) = \frac{1}{2\pi}\int \mathrm{d}\ell\,\ell\left[P_\kappa^{\rm E}(\ell) + P_\kappa^{\rm B}(\ell)\right]\hat{U}^2(\theta\ell),\\ \langle MM^*\rangle(\theta) &= \langle M_{\rm ap}^2\rangle(\theta) - \langle M_\perp^2\rangle(\theta) + 2\mathrm{i}\langle M_{\rm ap}M_\perp\rangle(\theta)\\ &= \frac{1}{2\pi}\int \mathrm{d}\ell\,\ell\left[P_\kappa^{\rm E}(\ell) - P_\kappa^{\rm B}(\ell) + 2\mathrm{i}P_\kappa^{\rm EB}(\ell)\right]\hat{U}^2(\theta\ell).\end{aligned} \tag{2.13}$$

Combining these expressions, one gets

$$\begin{aligned}\langle M_{\rm ap}^2\rangle(\theta) &= \frac{1}{2\pi}\int \mathrm{d}\ell\,\ell\,P_\kappa^{\rm E}(\ell)\hat{U}^2(\theta\ell),\\ \langle M_\perp^2\rangle(\theta) &= \frac{1}{2\pi}\int \mathrm{d}\ell\,\ell\,P_\kappa^{\rm B}(\ell)\hat{U}^2(\theta\ell).\end{aligned} \tag{2.14}$$

These important equations state that $\langle M_{\rm ap}^2\rangle$ depends on the E-mode and $\langle M_\perp^2\rangle$ on the B-mode power spectrum only, respectively. Here, $\hat{U}(\theta\ell) = \mathcal{F}[U_\theta](\ell)$ is the Fourier transform of the filter function U_θ. For the polynomial filter (2.9), it is

$$\hat{U}(\eta) = \frac{24\mathrm{J}_4(\eta)}{\eta^2}, \tag{2.15}$$

and

$$\hat{U}(\eta) = \frac{\eta^2}{2}\mathrm{e}^{-\eta^2/2} \tag{2.16}$$

in the case of the Gaussian filter (2.10). The aperture mass dispersion for the two filter functions is plotted in Fig. 2.4.

All of the above defined second-order statistics of shear are linear functions of the convergence power spectrum P_κ. The corresponding filter functions determine how these statistics sample the convergence power on different scales. In Fig. 2.5, the behavior of these filter functions is shown. Both filters for the correlations function are broad and decrease very slowly towards large arguments, the amplitude of the oscillations of both $\mathrm{J}_{0,4}(\eta)$ decreases with $\eta^{-1/2}$. Because $\mathrm{J}_0(\eta)$ is constant for small η, ξ_+ picks up a lot of power at low wavenumbers or large scales, whereas $\mathrm{J}_4(\eta) \propto \eta^4$ for small η, thus ξ_- samples the power spectrum more localized than ξ_+. Both filters of the aperture mass statistics decrease quickly for large η, as η^{-5} for the polynomial filter (2.15) and exponentially for the Gaussian filter (2.16). Thus, the aperture mass statistics is the most localized measure of the power spectrum and is very sensitive to its shape and therefore on cosmological parameters. On the other hand, a statistics with a non-local, broad filter function collects power over a large range yielding a higher signal-to-noise than a very localized statistics.

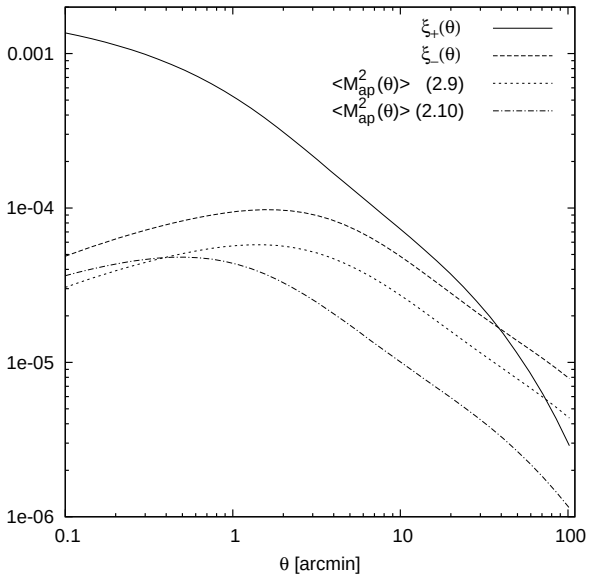

Figure 2.4: *The second-order statistics of cosmic shear used in this work, for a ΛCDM model (Table B.1, model 1). The solid and long-dashed lines show the two components of the 2PCF, ξ_+ and ξ_-, respectively. The two aperture mass dispersions correspond to the polynomial (2.9, short-dashed) and Gaussian filter (2.10, dash-dotted line), respectively.*

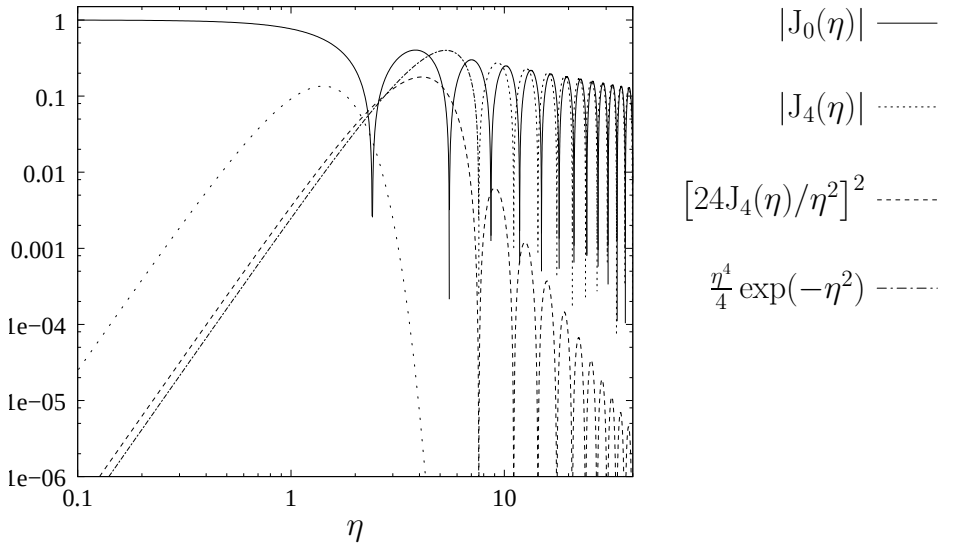

Figure 2.5: *Filter functions for the different second-order shear statistics.*

2.2.3 Interrelations

The dispersion of the aperture mass can in principle be measured directly from data by placing apertures onto the observed area. However, this method is very ineffective – regions with bright stars, foreground galaxies or telescope reflections have to be omitted in order not to bias the result. Moreover, for the Gaussian filter (2.10), apertures of radius θ cannot be put closer than about 3θ from the image border because of the significant exponential tail of the filter.

A more effective way to get $\langle M_{\rm ap}^2 \rangle$ from data is by integration over the 2PCF. Since $\langle M_{\rm ap}^2 \rangle$ is given in terms of the power spectrum (2.14), and the equation which relates the 2PCF to the power spectrum (2.12) can be inverted, one can express $\langle M_{\rm ap}^2 \rangle$ (and any other second-order statistics) in terms of the 2PCF. The following relation can be derived (Schneider et al. 2002;

Crittenden et al. 2002)

$$\left\langle M^2_{\mathrm{ap},\perp}(\theta)\right\rangle = \frac{1}{2\theta^2}\int_0^\infty \mathrm{d}\vartheta\,\vartheta\,\left[\xi_+(\vartheta)T_+\left(\frac{\vartheta}{\theta}\right) \pm \xi_-(\vartheta)T_-\left(\frac{\vartheta}{\theta}\right)\right], \tag{2.17}$$

where the functions T_+ and T_- are given by

$$T_\pm(x) = \int_0^\infty \mathrm{d}t\,t\,\mathrm{J}_{0,4}(xt)\hat{U}^2(t). \tag{2.18}$$

Explicit expressions for $T_\pm$ can be found in Schneider et al. (2002) for the polynomial and Crittenden et al. (2002) in the case of the Gaussian filter. The functions $T_\pm(x)$ vanish for $x \to 0$ in both cases and have compact support of $x \in [0;2]$ for the polynomial filter and decay exponentially for the Gaussian filter. Thus, the correlation functions have to be known only for a finite range of angular scales in order to infer the aperture mass statistics and separate the E- from the B-mode power spectrum. This fact is the major advantage of the aperture mass dispersion over the 2PCF. For the latter statistics alone and for a realistic shear data set, it is not possible to separate the E- from the B-mode without extrapolating the correlation function to scales where it has not been measured, see Sect. 2.2.1.

Note that in the absence of a B-mode, both terms in (2.17) have equal amplitude.

2.3 Third-order statistics

2.3.1 The three-point correlation function

The shear two-point correlation function (2PCF, Sect. 2.2.1) is defined for a pair of points; analogously, the three-point correlation function (3PCF) is given for a triplet of points $\boldsymbol{X}_1$, $\boldsymbol{X}_2$, and $\boldsymbol{X}_3$ forming a triangle. Let $\boldsymbol{x}_1 = \boldsymbol{X}_3 - \boldsymbol{X}_2$, $\boldsymbol{x}_2 = \boldsymbol{X}_1 - \boldsymbol{X}_3$ and $\boldsymbol{x}_3 = \boldsymbol{X}_2 - \boldsymbol{X}_1$ be the vectors forming the three sides of the triangle. Moreover, let φ_i be the polar angle of the side vector $\boldsymbol{x}_i$ and ϕ_i the interior angle at the triangle corner $\boldsymbol{X}_i$, see Fig. 2.6. One then gets the relations

$$\varphi_3 - \varphi_2 = \pi - \phi_1, \quad \varphi_1 - \varphi_3 = \pi - \phi_2, \quad \varphi_2 - \varphi_1 = \pi - \phi_3. \tag{2.19}$$

The Cartesian components of the shear 3PCF (Schneider & Lombardi 2003) are defined as

$$\gamma_{\mu\nu\lambda}(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3) = \langle\gamma_\mu(\boldsymbol{X}_1)\gamma_\nu(\boldsymbol{X}_2)\gamma_\lambda(\boldsymbol{X}_3)\rangle, \tag{2.20}$$

where $\gamma_\mu(\boldsymbol{X}_i), \mu = 1, 2$ is the Cartesian component of the shear at position $\boldsymbol{X}_i$.

One would like to write the 3PCF in terms of the tangential and cross-component of the shear, in analogy to the 2PCF. However, this task is not straightforward at all in the third-order case. The first difficulty is to define a direction with respect to which γ_t and $\gamma_\times$ are understood. For a triangle no obvious projection direction exists in contrast to the two-point case where the connecting vector of the two points is a natural choice. The second obstacle is that it is not possible to define a scalar from the product of three two-component quantities, which again is in contrast to the second-order case. In 2003, three papers (Schneider & Lombardi 2003; Takada & Jain 2003; Zaldarriaga & Scoccimarro 2003) introduced polar components of the 3PCF projected to different centers of the corresponding triangle. Furthermore, Schneider & Lombardi (2003) defined the so-called *natural components* (see below) of the 3PCF, whose transformation properties are particularly simple and which are relatively easily expressed in terms of the convergence bispectrum (Schneider, Kilbinger, & Lombardi 2005), see Sect. 2.3.1.5.

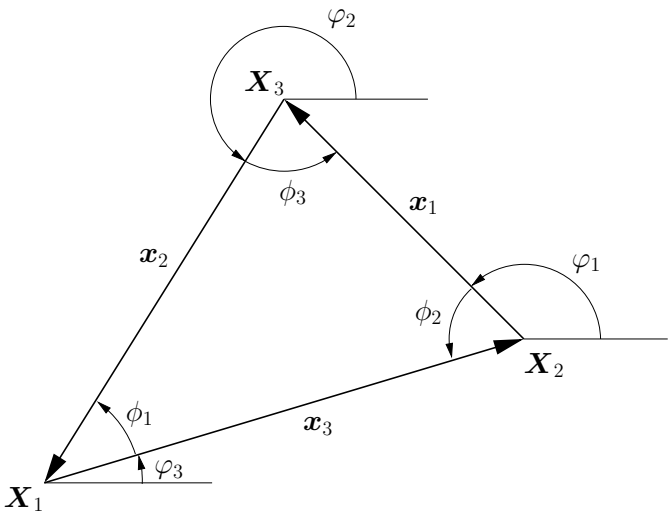

Figure 2.6: *The labeling of the corner points, sides and angles of the triangles used for the definition of the 3PCF.*

2.3.1.1 Projection onto triangle centers

Given a triangle, there are a number of special points or triangle "centers", such as the orthocenter (intersection point of the altitudes), the centroid (intersection point of the side-bisectors), the barycenter etc. All these points (in fact, any point which is fixed with respect to the triangle) can be used to define a direction to project the shear. For example, for every triangle corner $\boldsymbol{X}_i$, one can project the corresponding shear $\gamma_\mu(\boldsymbol{X}_i)$ to the altitude of the opposite side $\boldsymbol{x}_i$; the common projection point of all three corners is then the orthocenter. Eq. (2.1), $\gamma_\mathrm{t} + \mathrm{i}\gamma_\times = -(\gamma_1 + \mathrm{i}\gamma_2)\exp(-2\mathrm{i}\alpha)$, is more conveniently written in components, $\gamma_\tau = -R_{\tau\mu}(2\alpha)\gamma_\mu$, where $\tau = 1, 2$ corresponds to 't' and '×', respectively, and the general rotation matrix R is

$$\mathsf{R}(\phi) = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix}. \tag{2.21}$$

The polar angle of the altitude of $\boldsymbol{x}_i$ is $\alpha_i = \varphi_i + \pi/2$; inserting the three angles into (2.21), one gets the 3PCF in the orthocenter projection,

$$\gamma^{(\mathrm{o})}_{\tau\sigma\rho}(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3) = R_{\tau\mu}(2\varphi_1)R_{\sigma\nu}(2\varphi_2)R_{\rho\lambda}(2\varphi_3)\gamma^{\mathrm{cart}}_{\mu\nu\lambda}(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3). \tag{2.22}$$

Here, and in the following, I will denote the projected 3PCF with a superscript indicating the projection center (e.g. '(o)' for orthocenter, '(b)' for barycenter) in case of ambiguity. If the projection center is not specified, the superscript will be 'P'. The 3PCF in Cartesian coordinates will be denoted by 'cart'.

The 3PCF in a particular projection (to any triangle center) is invariant under a rotation of the whole triangle, since the projection center is fixed with respect to the triangle and performs the same rotation alongside. The 3PCF thus only depends on three independent quantities which uniquely describe a triangle. These three quantities are e.g. the moduli of the three sides x_1, x_2, x_3, or the moduli of two sides x_1, x_2 and one angle ϕ_3. In addition, the parity of the triangle has to be fixed, see next section. If not stated otherwise, positive parity is assumed in all cases, i.e. the vectors $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3$ describe an anti-clockwise or cyclic closed path. Negative parity corresponds to a clockwise or anti-cyclic closed path of these vectors.

2.3.1.2 Parity-modes of the 3PCF

The shear has two components at each of the three triangle corners, therefore the 3PCF has $2^3 = 8$ components. The question arises whether all of these components are important and carry cosmological information, or whether maybe some of them vanish because of parity symmetry, as for example the function $\xi_\times$ (2.11) in the case of the two-point correlation function.

A parity transform P — e.g. the flipping of the triangle along one of its sides — keeps the tangential shear component at each triangle corner invariant and causes the cross-component to change sign, which can be expressed as $\mathrm{P}\gamma = \gamma^*$ in complex notation. Thus, any component of the 3PCF which consists of a product of three γ is either invariant or changes sign under a parity operation. According to this transformation behavior, the eight components of the 3PCF are divided into the two groups,

$$\begin{aligned} &\gamma_{\mathrm{ttt}}, \gamma_{\mathrm{t}\times\times}, \gamma_{\times\mathrm{t}\times}, \gamma_{\times\times\mathrm{t}} \quad \text{(parity-even)}, \\ &\gamma_{\mathrm{tt}\times}, \gamma_{\mathrm{t}\times\mathrm{t}}, \gamma_{\times\mathrm{tt}}, \gamma_{\times\times\times} \quad \text{(parity-odd)}. \end{aligned} \tag{2.23}$$

where the parity-even (-odd) modes consist of an even (odd) number of $\gamma_\times$'s and accordingly their eigenvalue for the parity operator P is $+1$ (-1), respectively. The parity-modes are plotted for special triangles in Figs. 4.4 and 4.6.

2.3.1.3 Natural components of the 3PCF

A rotation of the 3PCF will transform $\gamma_{\upsilon\varsigma\varrho}$ into $\gamma'_{\tau\sigma\rho} = R_{\tau\upsilon}(2\alpha_1)R_{\sigma\varsigma}(2\alpha_2)R_{\rho\varrho}(2\alpha_3)\gamma_{\upsilon\varsigma\varrho}$. One cannot expect any of the parity-modes as defined in the last section to be invariant under general rotation. However, special rotations can be found which leave linear combinations of these modes invariant. For example, the quantity

$$\langle\gamma(\boldsymbol{X}_1)\gamma(\boldsymbol{X}_2)\gamma(\boldsymbol{X}_3)\rangle^{\mathrm{P}} = \langle[(\gamma_\mathrm{t} + \mathrm{i}\gamma_\times)(\boldsymbol{X}_1)]\,[(\gamma_\mathrm{t} + \mathrm{i}\gamma_\times)(\boldsymbol{X}_2)]\,[(\gamma_\mathrm{t} + \mathrm{i}\gamma_\times)(\boldsymbol{X}_3)]\rangle$$

is multiplied by the phase factor $\exp[-2\mathrm{i}(\alpha_1 + \alpha_2 + \alpha_3)]$ according to (2.1) when rotated as described above and it is seen immediately that it is invariant for $\alpha_1 + \alpha_2 + \alpha_3 = 0$. Analogously, $\langle\gamma^*(\boldsymbol{X}_1)\gamma(\boldsymbol{X}_2)\gamma(\boldsymbol{X}_3)\rangle$ is invariant for the special case $-\alpha_1 + \alpha_2 + \alpha_3 = 0$. Two more "invariants" can be defined; because of their special properties they are called the *natural components* of the 3PCF (Schneider & Lombardi 2003),

$$\begin{aligned} \Gamma^{(0)}(x_1, x_2, x_3) &= \langle\gamma(\boldsymbol{X}_1)\gamma(\boldsymbol{X}_2)\gamma(\boldsymbol{X}_3)\rangle^{\mathrm{p}} \\ &= -\left\langle(\gamma(\boldsymbol{X}_1)\gamma(\boldsymbol{X}_2)\gamma(\boldsymbol{X}_3))^{\mathrm{cart}}\,\mathrm{e}^{-2\mathrm{i}(\alpha_1+\alpha_2+\alpha_3)}\right\rangle \\ &= \gamma_{\mathrm{ttt}} - \gamma_{\mathrm{t}\times\times} - \gamma_{\times\mathrm{t}\times} - \gamma_{\times\times\mathrm{t}} + \mathrm{i}(\gamma_{\mathrm{tt}\times} + \gamma_{\mathrm{t}\times\mathrm{t}} + \gamma_{\times\mathrm{tt}} - \gamma_{\times\times\times}), \end{aligned}$$

$$\begin{aligned} \Gamma^{(1)}(x_1, x_2, x_3) &= \langle\gamma^*(\boldsymbol{X}_1)\gamma(\boldsymbol{X}_2)\gamma(\boldsymbol{X}_3)\rangle^{\mathrm{p}} \\ &= -\left\langle(\gamma(\boldsymbol{X}_1)\gamma(\boldsymbol{X}_2)\gamma(\boldsymbol{X}_3))^{\mathrm{cart}}\,\mathrm{e}^{-2\mathrm{i}(-\alpha_1+\alpha_2+\alpha_3)}\right\rangle \\ &= \gamma_{\mathrm{ttt}} - \gamma_{\mathrm{t}\times\times} + \gamma_{\times\mathrm{t}\times} + \gamma_{\times\times\mathrm{t}} + \mathrm{i}(\gamma_{\mathrm{tt}\times} + \gamma_{\mathrm{t}\times\mathrm{t}} - \gamma_{\times\mathrm{tt}} + \gamma_{\times\times\times}), \end{aligned}$$

$$\begin{aligned} \Gamma^{(2)}(x_1, x_2, x_3) &= \langle\gamma(\boldsymbol{X}_1)\gamma^*(\boldsymbol{X}_2)\gamma(\boldsymbol{X}_3)\rangle^{\mathrm{p}} \\ &= -\left\langle(\gamma(\boldsymbol{X}_1)\gamma(\boldsymbol{X}_2)\gamma(\boldsymbol{X}_3))^{\mathrm{cart}}\,\mathrm{e}^{-2\mathrm{i}(\alpha_1-\alpha_2+\alpha_3)}\right\rangle \end{aligned}$$

$$= \gamma_{\mathrm{ttt}} + \gamma_{\mathrm{t}\times\times} - \gamma_{\times\mathrm{t}\times} + \gamma_{\times\times\mathrm{t}} + \mathrm{i}(\gamma_{\mathrm{tt}\times} - \gamma_{\mathrm{t}\times\mathrm{t}} + \gamma_{\times\mathrm{tt}} + \gamma_{\times\times\times}),$$

$$\begin{aligned}\Gamma^{(3)}(x_1,x_2,x_3) &= \langle\gamma(\boldsymbol{X}_1)\gamma(\boldsymbol{X}_2)\gamma^*(\boldsymbol{X}_3)\rangle^{\mathrm{p}} \\ &= -\left\langle(\gamma(\boldsymbol{X}_1)\gamma(\boldsymbol{X}_2)\gamma(\boldsymbol{X}_3))^{\mathrm{cart}}\,\mathrm{e}^{-2\mathrm{i}(\alpha_1+\alpha_2-\alpha_3)}\right\rangle \\ &= \gamma_{\mathrm{ttt}} + \gamma_{\mathrm{t}\times\times} + \gamma_{\times\mathrm{t}\times} - \gamma_{\times\times\mathrm{t}} + \mathrm{i}(-\gamma_{\mathrm{tt}\times} + \gamma_{\mathrm{t}\times\mathrm{t}} + \gamma_{\times\mathrm{tt}} + \gamma_{\times\times\times}). \quad (2.24)\end{aligned}$$

Inverting these equations, the parity-modes of the 3PCF can be obtained from the complex natural components as follows,

$$\begin{aligned}\gamma_{\mathrm{ttt}} &= \Re\left(\Gamma^{(0)}+\Gamma^{(1)}+\Gamma^{(2)}+\Gamma^{(3)}\right)/4, & \gamma_{\mathrm{tt}\times} &= \Im\left(\Gamma^{(0)}+\Gamma^{(1)}+\Gamma^{(2)}-\Gamma^{(3)}\right)/4, \\ \gamma_{\mathrm{t}\times\times} &= \Re\left(-\Gamma^{(0)}-\Gamma^{(1)}+\Gamma^{(2)}+\Gamma^{(3)}\right)/4, & \gamma_{\mathrm{t}\times\mathrm{t}} &= \Im\left(\Gamma^{(0)}+\Gamma^{(1)}-\Gamma^{(2)}+\Gamma^{(3)}\right)/4, \\ \gamma_{\times\mathrm{t}\times} &= \Re\left(-\Gamma^{(0)}+\Gamma^{(1)}-\Gamma^{(2)}+\Gamma^{(3)}\right)/4, & \gamma_{\times\mathrm{tt}} &= \Im\left(\Gamma^{(0)}-\Gamma^{(1)}+\Gamma^{(2)}+\Gamma^{(3)}\right)/4, \\ \gamma_{\times\times\mathrm{t}} &= \Re\left(-\Gamma^{(0)}+\Gamma^{(1)}+\Gamma^{(2)}-\Gamma^{(3)}\right)/4, & \gamma_{\times\times\times} &= \Im\left(-\Gamma^{(0)}+\Gamma^{(1)}+\Gamma^{(2)}+\Gamma^{(3)}\right)/4.\end{aligned} \quad (2.25)$$

2.3.1.4 Transformation properties and symmetries of the 3PCF

Defining the complex phase factors

$$\begin{aligned}\lambda_0 &= \exp[-2\mathrm{i}(\alpha_1+\alpha_2+\alpha_3)], & \lambda_1 &= \exp[-2\mathrm{i}(-\alpha_1+\alpha_2+\alpha_3)], \\ \lambda_2 &= \exp[-2\mathrm{i}(\alpha_1-\alpha_2+\alpha_3)], & \lambda_3 &= \exp[-2\mathrm{i}(\alpha_1+\alpha_2-\alpha_3)],\end{aligned} \quad (2.26)$$

the q^{th} natural component is invariant under a special rotation with a choice of the rotation angles $\alpha_1, \alpha_2, \alpha_3$ such that $\lambda_q = 1$ is assured. A general rotation simply changes the phase of the natural components, they do not mix, $(\Gamma^{(q)})' = \lambda_q\Gamma^{(q)}$. Since the 3PCF has eight real components, a general rotation can be written as an 8×8-matrix. This matrix has eight eigenvalues, which are the four λ_q defined in (2.26) and their complex conjugates. Thus, the natural components $\Gamma^{(q)}$ together with their complex conjugates are a complete set of eigenvectors of the triangle rotation, with corresponding complex eigenvalues λ_q and λ_q^*.

For special triangles obeying certain symmetries some of the 3PCF components or their combinations vanish. For isosceles triangles with, say $x_1 = x_2$, one finds that two of the parity-even modes are equal, $\gamma_{\mathrm{t}\times\times} = \gamma_{\times\mathrm{t}\times}$, two of the parity-odd modes vanish, $\gamma_{\mathrm{tt}\times} = \gamma_{\times\times\times} = 0$ and the two remaining odd components have opposite sign, $\gamma_{\mathrm{t}\times\mathrm{t}} = -\gamma_{\times\mathrm{tt}}$. This implies that $\Gamma^{(0)}$ and $\Gamma^{(3)}$ are real and equal and $\Im\Gamma^{(1)} = -\Im\Gamma^{(2)}$. For equilateral triangles, all parity-odd modes vanish and as a consequence, the imaginary part of the natural components vanishes, too.

In Sect. 4.1, I will verify these symmetry properties of the 3PCF using ray-tracing simulations and predictions from non-linear models of the large-scale structure. I will also show that in general, none of the eight components is vanishing, following the work of Takada & Jain (2003). Thus, all eight parity-even and -odd modes (2.23) carry important information about the projected mass density and should be taken into account for shear measurements. Equivalently, the four complex natural components can be used, and in fact, the latter ones are the natural choice when expressing the 3PCF in terms of the convergence bispectrum and thus directly relating the third-order shear statistics to cosmology, see following section.

2.3.1.5 The 3PCF in terms of the convergence bispectrum

The natural components in the orthocenter projection are written in terms of the convergence bispectrum B_κ (Schneider, Kilbinger, & Lombardi 2005). First, the 3PCF is expressed in Cartesian coordinates using (2.22) and transformed into Fourier space. The triple-correlator of $\hat{\gamma}$ is written in terms of $\hat{\kappa}$ using (1.72) and the first natural component is

$$\Gamma^{(0)}(x_1,x_2,x_3) = \langle\gamma(\boldsymbol{X}_1)\gamma(\boldsymbol{X}_2)\gamma(\boldsymbol{X}_3)\rangle^{(\mathrm{o})} = \mathrm{e}^{-2\mathrm{i}(\varphi_1+\varphi_2+\varphi_3)}\int\frac{\mathrm{d}^2\ell_1}{(2\pi)^2}\mathrm{e}^{-\mathrm{i}\boldsymbol{X}_1\boldsymbol{\ell}_1}$$
$$\times\int\frac{\mathrm{d}^2\ell_2}{(2\pi)^2}\mathrm{e}^{-\mathrm{i}\boldsymbol{X}_2\boldsymbol{\ell}_2}\int\frac{\mathrm{d}^2\ell_3}{(2\pi)^2}\mathrm{e}^{-\mathrm{i}\boldsymbol{X}_3\boldsymbol{\ell}_3}\langle\hat{\kappa}(\boldsymbol{\ell}_1)\hat{\kappa}(\boldsymbol{\ell}_2)\hat{\kappa}(\boldsymbol{\ell}_3)\rangle\mathrm{e}^{2\mathrm{i}(\beta_1+\beta_2+\beta_3)}. \tag{2.27}$$

Next, the bispectrum B_κ (1.92) for the triple correlator of the convergence is inserted and one ℓ-integration for each of the three terms is carried out, making use of the delta-function. After renaming the integration variables, one gets

$$\Gamma^{(0)}(x_1,x_2,x_3) = \mathrm{e}^{-2\mathrm{i}(\varphi_1+\varphi_2+\varphi_3)}\int_0^\infty\frac{\mathrm{d}\ell_1\,\ell_1}{(2\pi)^2}\int_0^\infty\frac{\mathrm{d}\ell_2\,\ell_2}{(2\pi)^2}\int_0^{2\pi}\mathrm{d}\beta_1\int_0^{2\pi}\mathrm{d}\beta_2\,b_\kappa(\ell_1,\ell_2,\varphi)$$
$$\times\left[\mathrm{e}^{\mathrm{i}(\boldsymbol{x}_1\boldsymbol{\ell}_2-\boldsymbol{x}_2\boldsymbol{\ell}_1)}+\mathrm{e}^{\mathrm{i}(\boldsymbol{x}_2\boldsymbol{\ell}_2-\boldsymbol{x}_1\boldsymbol{\ell}_1)}+\mathrm{e}^{\mathrm{i}(\boldsymbol{x}_3\boldsymbol{\ell}_2-\boldsymbol{x}_1\boldsymbol{\ell}_1)}\right]\mathrm{e}^{2\mathrm{i}(\beta_1+\beta_2+\beta_3)}. \tag{2.28}$$

For the moment, only the first term of this equation is considered. Since the bispectrum b_κ only depends on $\varphi=\beta_1-\beta_2$ which is the difference of the polar angles of $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$, one more angular integral can be carried out by substituting β_1 and β_2 with φ and θ_3, where

$$\varphi=\beta_1-\beta_2,\quad \theta_3=(\beta_1+\beta_2)/2-(\varphi_1+\varphi_2)/2. \tag{2.29}$$

Focusing on the scalar product $\boldsymbol{x}_1\boldsymbol{\ell}_2$ occurring in the exponential in (2.28), the angle between the two vectors is $\varphi_1-\beta_2=(\varphi_1+\varphi_2)/2+(\varphi_1-\varphi_2)/2+(\beta_1-\beta_2)/2-(\beta_1+\beta_2)/2=-(\pi-\phi_3)/2+\varphi/2-\theta_3$, where use of (2.19) was made. One then can express the scalar product as $\boldsymbol{x}_1\boldsymbol{\ell}_2=-x_1\ell_2\sin[\theta_3-(\varphi+\phi_3)/2]$ and, similarly, $\boldsymbol{x}_2\boldsymbol{\ell}_1=x_2\ell_1\sin[\theta_3+(\varphi+\phi_3)/2]$. The sum of the two scalar products can be written as $\boldsymbol{x}_1\boldsymbol{\ell}_2-\boldsymbol{x}_2\boldsymbol{\ell}_1=-A_3\sin(\theta_3+\alpha_3)$ with

$$\cos\alpha_3=\frac{\ell_1x_2+\ell_2x_1}{A_3}\cos\left(\frac{\varphi+\phi_3}{2}\right),\quad \sin\alpha_3=\frac{\ell_1x_2-\ell_2x_1}{A_3}\cos\left(\frac{\varphi+\phi_3}{2}\right),$$
$$A_3^2=(\ell_1x_2)^2+(\ell_2x_1)^2+2\ell_1x_2\ell_2x_1\cos(\varphi+\phi_3). \tag{2.30}$$

Next, one rewrites the sum of angles in the exponent in (2.28), $\beta_1+\beta_2+\beta_3-\varphi_1-\varphi_2-\varphi_3=2\theta_3+\beta_3+(\varphi_1+\varphi_2)/2-\varphi_3=3\theta_3+\bar{\beta}+(\phi_1-\phi_2)/2$, with

$$\bar{\beta}=\beta_3-\theta_3-(\varphi_1+\varphi_2)/2=\beta_3-\frac{\beta_1+\beta_2}{2}. \tag{2.31}$$

Now, one angular integration of the first term in (2.28) can be performed as follows:

$$\int_0^{2\pi}\mathrm{d}\beta_1\;\mathrm{e}^{2\mathrm{i}(\beta_1+\beta_2+\beta_3-\varphi_1-\varphi_2-\varphi_3)}\mathrm{e}^{\mathrm{i}(\boldsymbol{x}_1\boldsymbol{\ell}_2-\boldsymbol{x}_2\boldsymbol{\ell}_1)}$$
$$=\mathrm{e}^{2\mathrm{i}\bar{\beta}}\,\mathrm{e}^{\mathrm{i}(\phi_1-\phi_2)}\int_0^{2\pi}\mathrm{d}\theta_3\;\mathrm{e}^{6\mathrm{i}\theta_3}\mathrm{e}^{-\mathrm{i}A_3\sin(\theta_3+\alpha_3)}$$
$$=\mathrm{e}^{2\mathrm{i}\bar{\beta}}\,\mathrm{e}^{\mathrm{i}(\phi_1-\phi_2)}\int_0^{2\pi}\mathrm{d}\vartheta\;\mathrm{e}^{6\mathrm{i}(\vartheta-\alpha_3)}\,\mathrm{e}^{-\mathrm{i}A_3\sin\vartheta}=2\pi\,\mathrm{e}^{2\mathrm{i}\bar{\beta}}\,\mathrm{e}^{\mathrm{i}(\phi_1-\phi_2)}\,\mathrm{e}^{-6\mathrm{i}\alpha_3}\;\mathrm{J}_6(A_3). \tag{2.32}$$

Similar calculations for the other two terms lead to the final result

$$\Gamma^{(0)}(x_1,x_2,x_3) = (2\pi)^{-3}\int_0^\infty \mathrm{d}\ell_1\,\ell_1\int_0^\infty \mathrm{d}\ell_2\,\ell_2\int \mathrm{d}\varphi\, b_\kappa(\ell_1,\ell_2,\varphi)\mathrm{e}^{2\mathrm{i}\bar{\beta}} \times\left[\mathrm{e}^{\mathrm{i}(\phi_1-\phi_2-6\alpha_3)}\mathrm{J}_6(A_3)+\mathrm{e}^{\mathrm{i}(\phi_2-\phi_3-6\alpha_1)}\mathrm{J}_6(A_1)+\mathrm{e}^{\mathrm{i}(\phi_3-\phi_1-6\alpha_2)}\mathrm{J}_6(A_2)\right], \quad (2.33)$$

where the A_i and α_i are obtained from (2.30) by cyclic permutations of x_1, x_2 and x_3.

Similarly, the second natural component $\Gamma^{(1)}$ is expressed in terms of the projected bispectrum b_κ (1.92, 1.93). The phase factor responsible for the transformation into Cartesian coordinates in this case is $\exp[-2\mathrm{i}(-\varphi_1+\varphi_2+\varphi_3)]$ (see eq. 2.24), and from the fact that κ is real, one infers $\mathcal{F}[\langle\gamma^*\gamma\gamma\rangle](\boldsymbol{\ell}_1,\boldsymbol{\ell}_2,\boldsymbol{\ell}_3) = \langle\hat{\kappa}(-\boldsymbol{\ell}_1)\hat{\kappa}(\boldsymbol{\ell}_2)\hat{\kappa}(\boldsymbol{\ell}_3)\rangle\exp(2\mathrm{i}(-\beta_1+\beta_2+\beta_3)$. After similar calculations than above, the final result is

$$\Gamma^{(1)}(x_1,x_2,x_3) = (2\pi)^{-3}\int_0^\infty \mathrm{d}\ell_1\,\ell_1\int_0^\infty \mathrm{d}\ell_2\,\ell_2\int \mathrm{d}\varphi\, b_\kappa(\ell_1,\ell_2,\varphi)\mathrm{e}^{2\mathrm{i}\bar{\beta}}\left[\mathrm{e}^{\mathrm{i}(\phi_1-\phi_2+2\alpha_3)}\mathrm{e}^{2\mathrm{i}(\bar{\beta}-\varphi+\phi_3)} \times \mathrm{J}_2(A_3) +\mathrm{e}^{\mathrm{i}(\phi_2-\phi_3-2\alpha_1)}\mathrm{e}^{2\mathrm{i}(-\bar{\beta}+\phi_3-\phi_2))}\mathrm{J}_2(A_1)+\mathrm{e}^{\mathrm{i}(\phi_3-\phi_1-2\alpha_2)}\mathrm{e}^{2\mathrm{i}(\bar{\beta}+\varphi-\phi_2)}\mathrm{J}_2(A_2)\right]. \quad (2.34)$$

The remaining two natural components $\Gamma^{(2)}$ and $\Gamma^{(3)}$ are obtained from (2.34) by cyclic permutations of the indices, according to the transformation laws.

The expressions of the 3PCF as functions of the bispectrum are considerably more complex than those of the 2PCF in terms of the power spectrum. Three-dimensional integrals are involved (integration over the three arguments of the bispectrum), the integrand extends to infinity and is highly oscillating due to the Bessel functions of order two and six. Note that the arguments of the Bessel functions are not analytic functions of the integration variables, and are modulated with the cosine of the integration angle φ. Depending on the triangle configuration and ℓ_1 and ℓ_2, this modulation can cause several hundreds to thousands of zeros of the Bessel functions between 0 and 2π deeming a numerical integration to an ordeal. This makes the predictions of the 3PCF from theoretical models of the bispectrum challenging.

Expressions (2.33) and (2.34) can be inverted to obtain the bispectrum in terms of the shear 3PCF, in analogy to the second-order case, where the power spectrum is written as a function of the shear 2PCF. Consequently, since any third-order statistics (linearly) depends on the bispectrum, it can be expressed as a function of the 3PCF. An application of this fact will be discussed in Sect. 2.3.3, where the third-order aperture mass statistics is written in terms of the 3PCF natural components (although the derivation of these expressions does not use the inversion of the bispectrum as a function of the 3PCF).

2.3.2 Third-order aperture mass statistics

The third moment or skewness of the aperture mass (2.2) was first considered in Schneider et al. (1998) in the framework of quasi-linear perturbation theory (PT). It has been measured for the first time with the VIRMOS-DESCART survey by Pen et al. (2003) who used the reduced skewness $\langle M_\mathrm{ap}^3\rangle/\langle M_\mathrm{ap}^2\rangle^2$ to derive an upper limit for Ω_Λ. However, their data set was significantly contaminated by a B-mode and even by a parity-violating mode (Schneider 2003). A 2σ detection of the M_ap-skewness on scales with no substantial B-mode present was obtained from the CTIO data by Jarvis et al. (2004).

Schneider, Kilbinger, & Lombardi (2005) generalized the expression for the skewness, and defined the three-point correlator of the aperture mass for three different aperture radii,

$\langle M_{\rm ap}^3(\theta_1,\theta_2,\theta_3)\rangle = \langle M_{\rm ap}(\theta_1)\, M_{\rm ap}(\theta_2)\, M_{\rm ap}(\theta_3)\rangle$. From now on, I will denote the skewness of the aperture with $\langle M_{\rm ap,d}^3\rangle = \langle M_{\rm ap}^3(\theta,\theta,\theta)\rangle$ ('d' for diagonal). The hypothesis that the generalized third-order aperture mass statistics contains much more information about cosmology than the skewness is subject of a detailed analysis in Chapter 5. This analysis and the resulting arguments which confirm this hypothesis constitute one of the main parts of this thesis.

Using (2.2), the generalized third-order aperture mass statistics can be written in terms of the convergence bispectrum B_κ as

$$\begin{aligned}\left\langle M_{\rm ap}^3(\theta_1,\theta_2,\theta_3)\right\rangle &= \langle M_{\rm ap}(\theta_1) M_{\rm ap}(\theta_2) M_{\rm ap}(\theta_3)\rangle \\ &= \int\frac{{\rm d}^2\ell_1}{(2\pi)^2}\int\frac{{\rm d}^2\ell_2}{(2\pi)^2}\,B_\kappa(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2)\sum_{(i,j,k)\in S_3}\hat{U}(\theta_i|\boldsymbol{\ell}_1|)\,\hat{U}(\theta_j|\boldsymbol{\ell}_2|)\,\hat{U}(\theta_k|\boldsymbol{\ell}_1+\boldsymbol{\ell}_2|),\end{aligned}\tag{2.35}$$

where S_3 is the symmetric permutation group of (123), thus the summation is performed over even permutations of i,j,k. This expression can be further simplified using polar coordinates for the integration variables $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$. One angular integral is trivial and yields 2π. The second one can be performed analytically because of the separable φ-dependence of the bispectrum (1.93), where φ is the angle between $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$. Integrals of the form

$$K^{(m)}(\theta\ell_1,\theta\ell_2) = \int_0^{2\pi}{\rm d}\varphi\cos^m(\varphi)\,\hat{U}\left(\theta\sqrt{\ell_1^2+\ell_2^2+2\ell_1\ell_2\cos\varphi}\right)\tag{2.36}$$

for $m=0,1,2$ occur; for U_θ being the Gaussian filter (2.10) these are

$$\begin{aligned}K^{(0)}(t_1,t_2) &= \pi{\rm e}^{-\frac{1}{2}(t_1^2+t_2^2)}\left[(t_1^2+t_2^2){\rm I}_0(t_1t_2) - 2t_1t_2{\rm I}_1(t_1t_2)\right],\\ K^{(1)}(t_1,t_2) &= \pi{\rm e}^{-\frac{1}{2}(t_1^2+t_2^2)}\left[2t_1t_2{\rm I}_0(t_1t_2) - (2+t_1^2+t_2^2){\rm I}_1(t_1t_2)\right],\\ K^{(2)}(t_1,t_2) &= \pi{\rm e}^{-\frac{1}{2}(t_1^2+t_2^2)}\left[\left(2+t_1^2+t_2^2\right){\rm I}_0(t_1t_2) - \left(\frac{t_1^2+t_2^2+4}{t_1t_2}+2t_1t_2\right){\rm I}_1(t_1t_2)\right],\end{aligned}\tag{2.37}$$

where ${\rm I}_\nu$ is the modified Bessel function of order ν.

One remark concerning the asymptotic behavior of the integrand shall be made here. For large arguments the modified Bessel functions diverge exponentially, ${\rm I}_\nu(x)\approx{\rm e}^x/\sqrt{2\pi x}$. However, this divergence is absorbed by the exponential prefactor of the $K^{(m)}$ and the functions decrease exponentially as $\exp[(t_1-t_2)^2]$ — this term dominates over the polynomials for large arguments. If the integral is to be calculated numerically, this asymptotic form of the $K^{(m)}$ should be used in order to reduce round-off errors (multiplication of a very large by a very small number).

Inserting (2.37) into (2.35), one gets

$$\left\langle M_{\rm ap}^3(\theta_1,\theta_2,\theta_3)\right\rangle = (2\pi)^{-3}\int{\rm d}\ell_1\,\ell_1\int{\rm d}\ell_2\,\ell_2\sum_{m=0}^{2}I^{(m)}(\theta_1,\theta_2,\theta_3;\ell_1,\ell_2)\,\bar{b}_\kappa^{(m)}(\ell_1,\ell_2)\tag{2.38}$$

with

$$I^{(m)}(\theta_1,\theta_2,\theta_3;\ell_1,\ell_2) = F_2^{(m)}(\ell_1,\ell_2)\sum_{(i,j,k)\in S_3}\hat{U}(\theta_i\ell_1)\,\hat{U}(\theta_j\ell_2)K^{(m)}(\theta_k\ell_1,\theta_k\ell_2).\tag{2.39}$$

The functions $I^{(m)}$ determine how the generalized third-order aperture mass samples the bispectrum and therefore cosmological information — these functions are plotted in Fig. 2.7.

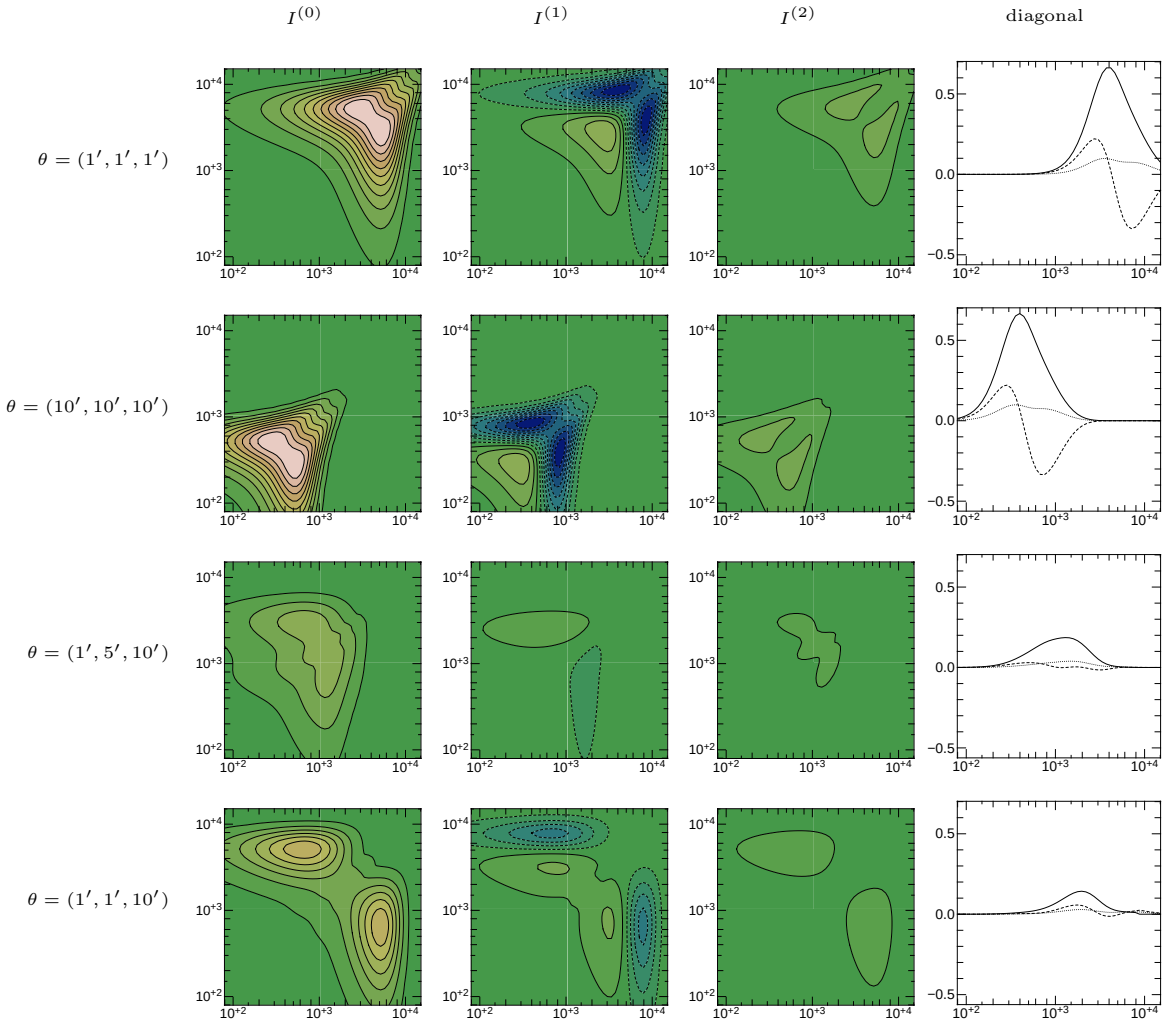

Figure 2.7: *Contours of the filter functions $I^{(m)}$ for $m = 0, 1, 2$ from left to right are plotted for different values of θ_i as a function of ℓ_1 and ℓ_2 (in units of inverse radians). Dashed contours indicate negative values. The right-most panels show the profile along the diagonal of $I^{(m)}$; the solid, dashed and dotted lines correspond to $m = 0, 1$ and 2, respectively.*

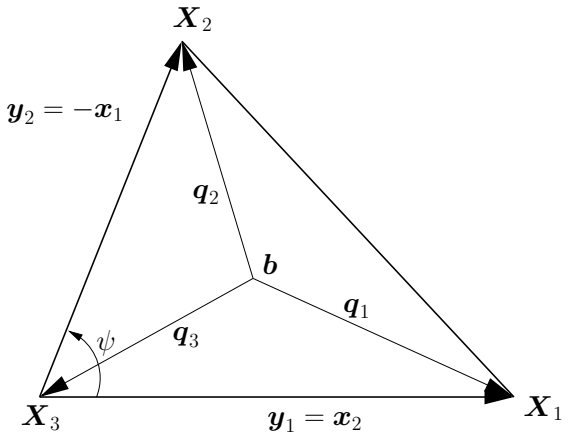

Figure 2.8: *Parametrization of the triangle used for the integration of the 3PCF.* $\boldsymbol{b}$ *is the centroid or barycenter of the triangle. The vectors* $\boldsymbol{q}_i$ *connect the barycenter with the edges of the triangles* $\boldsymbol{X}_i$*, their polar angles (not displayed) are denoted by* ϱ_i*.*

They all are relatively well localized which makes $\langle M_{\rm ap}^3 \rangle$ a local measure of the bispectrum. Furthermore, for equal filter scales only the region around the diagonal of the bispectrum is probed, corresponding to equilateral triangles in Fourier space. When different filter scales are taken into account, other parts further away from the diagonal of $\bar{b}_\kappa^{(m)}$ dominate the integral. Although in this latter case the amplitude of $I^{(m)}$ is lower, the generalized aperture mass, probing the bispectrum for general triangles in Fourier space, contains much more information about the bispectrum and cosmology than the 'diagonal' one. This approach to sample the bispectrum on a large region of Fourier space is similar to a previous study (Takada & Jain 2004), who have used all triangle configurations of the convergence bispectrum in order to predict tight constraints on cosmological parameters from cosmic shear. In contrast to their work, I will later use moments of the aperture mass statistics (which are direct weak lensing observables) as real-space probes of the convergence power spectrum and bispectrum (Sect. 5.3).

2.3.3 Interrelations

As in the second-order case (Sect. 2.2.3) there exist relations between $\langle M_{\rm ap}^3 \rangle$ and the correlation function. These relations were first derived by Jarvis et al. (2004) for the skewness $\langle M_{\rm ap,d}^3 \rangle$ and later generalized for $\langle M_{\rm ap}^3 \rangle$ in Schneider, Kilbinger, & Lombardi (2005). Both studies were undertaken for the Gaussian filter (2.10) since the corresponding expressions for the polynomial filter (2.9) are extremely cumbersome and of no practical use[1].

Following Jarvis et al. (2004), two of the triangle sides which were introduced in Sect. 2.3.1 are reparametrized as

$$\boldsymbol{y}_1 = \boldsymbol{x}_2 \quad \text{and} \quad \boldsymbol{y}_2 = -\boldsymbol{x}_1. \tag{2.40}$$

The connecting vectors between the barycenter of the triangle and its three corners are

$$\boldsymbol{q}_1 = \frac{2\,\boldsymbol{y}_1 - \boldsymbol{y}_2}{3}, \qquad \boldsymbol{q}_2 = \frac{2\,\boldsymbol{y}_2 - \boldsymbol{y}_1}{3}, \qquad \boldsymbol{q}_3 = -\frac{\boldsymbol{y}_1 + \boldsymbol{y}_2}{3}, \tag{2.41}$$

see Fig. 2.8. Let ϱ_k be the polar angle of $\boldsymbol{q}_k, k = 1, 2, 3$, which can be expressed as the complex number $\boldsymbol{q}_k = q_k \exp(\mathrm{i}\varrho_k)$. The 3PCF expressed in these new coordinates is denoted with $\tilde{\Gamma}^{(q)}$ in a projection to an arbitrary triangle center and is equal to the previously defined one,

$$\tilde{\Gamma}^{(q)}(\boldsymbol{y}_1, \boldsymbol{y}_2) = \tilde{\Gamma}^{(q)}(\boldsymbol{x}_2, -\boldsymbol{x}_1) = \Gamma^{(q)}(\boldsymbol{x}_1, \boldsymbol{x}_2), \tag{2.42}$$

[1] M. Lombardi, priv. comm.

which follows from the transformation rules of the 3PCF, see Sect. 2.3.1.4. Choosing the barycenter projection, one finds after some calculations (Schneider et al. 2005) the two quantities

$$\begin{aligned}\left\langle M^3\right\rangle(\theta_1,\theta_2,\theta_3) &= \frac{S}{24}\int\frac{\mathrm{d}y_1\,y_1}{\Theta^2}\int\frac{\mathrm{d}y_2\,y_2}{\Theta^2}\int_0^{2\pi}\frac{\mathrm{d}\psi}{2\pi}\,\tilde{\Gamma}_{\mathrm{b}}^{(0)}(y_1,y_2,\psi)R_{\boldsymbol{\theta}}^{(0)}(y_1,y_2,\psi),\\ \left\langle M^2M^*\right\rangle(\theta_1,\theta_2;\theta_3) &= \frac{S}{24}\int\frac{\mathrm{d}y_1\,y_1}{\Theta^2}\int\frac{\mathrm{d}y_2\,y_2}{\Theta^2}\int_0^{2\pi}\frac{\mathrm{d}\psi}{2\pi}\,\tilde{\Gamma}_{\mathrm{b}}^{(3)}(y_1,y_2,\psi)R_{\boldsymbol{\theta}}^{(3)}(y_1,y_2,\psi),\end{aligned}\tag{2.43}$$

where the two integral kernels are

$$\begin{aligned}R_{\boldsymbol{\theta}}^{(0)} &= \mathrm{e}^{-Z}\,\frac{q_1^2\,q_2^2\,q_3^2}{\Theta^6}\,f_1^{*2}\,f_2^{*2}\,f_3^{*2},\\ R_{\boldsymbol{\theta}}^{(3)} &= \mathrm{e}^{-Z}\left[\frac{q_1^2\,q_2^2\,q_3^2}{\Theta^6}\,f_1^{*2}f_2^{*2}f_3^2-\frac{8}{3}\frac{Q}{\Theta^4}\,f_1^*f_2^*f_3\,g_3^*\right.\\ &\quad\left.+\frac{8}{9}\left(\frac{Q^2}{q_1^2q_2^2q_3^2\Theta^2}g_3^{*2}+\frac{2\,\theta_1^2\theta_2^2}{\Theta^4}\,\frac{Q}{q_3^2\Theta^2}\,f_1^*f_2^*\right)\right],\end{aligned}\tag{2.44}$$

with

$$\begin{aligned}\Theta^2 &= \sqrt{\frac{\theta_1^2\theta_2^2+\theta_1^2\theta_3^2+\theta_2^2\theta_3^2}{3}},\quad S=\frac{\theta_1^2\theta_2^2\theta_3^2}{\Theta^6},\quad Q=q_1\,q_2\,q_1^2\,\mathrm{e}^{\mathrm{i}(\varrho_1+\varrho_2-2\mathrm{i}\varrho_3)},\\ Z &= \frac{(-\theta_1^2+2\theta_2^2+2\theta_3^2)q_1^2+(2\theta_1^2-\theta_2^2+2\theta_3^2)q_2^2+(2\theta_1^2+2\theta_2^2-\theta_3^2)q_3^2}{6\Theta^4},\end{aligned}\tag{2.45}$$

$$\begin{aligned}f_i &= \frac{\theta_j^2+\theta_k^2}{2\Theta^2}+\frac{(q_j\mathrm{e}^{\mathrm{i}\varrho_i}-q_k\mathrm{e}^{\mathrm{i}\varrho_j})q_i\mathrm{e}^{-\mathrm{i}\varrho_i}}{{q_i}^2}\,\frac{\theta_j^2-\theta_k^2}{6\,\Theta^2},\\ g_i &= \frac{\theta_j^2\theta_k^2}{\Theta^4}+\frac{(q_j\mathrm{e}^{\mathrm{i}\varrho_j}-q_k\mathrm{e}^{\mathrm{i}\varrho_k})q_i\mathrm{e}^{-\mathrm{i}\varrho_i}}{{q_i}^2}\,\frac{\theta_i^2(\theta_k^2-\theta_j^2)}{3\Theta^4},\end{aligned}\qquad (ijk)=\text{symm. perm. of (123)}.\tag{2.46}$$

In Figs. 2.9 and 2.10, exemplary plots of the integral filter functions from the two equations defined above are shown. Similar as in the case of the 2PCF, these functions have infinite support, but decrease exponentially.

The four generalized third-order aperture mass statistics are then given by

$$\begin{aligned}\left\langle M_{\mathrm{ap}}^3\right\rangle(\theta_1,\theta_2,\theta_3) &= \Re\left[\left\langle M^3\right\rangle(\theta_1,\theta_2,\theta_3)+\left\langle M^2M^*\right\rangle(\theta_1,\theta_2;\theta_3)\right.\\ &\quad\left.+\left\langle M^2M^*\right\rangle(\theta_1,\theta_3;\theta_2)+\left\langle M^2M^*\right\rangle(\theta_2,\theta_3;\theta_1)\right]/4\,,\\ \left\langle M_{\mathrm{ap}}^2M_\perp\right\rangle(\theta_1,\theta_2;\theta_3) &= \Im\left[\left\langle M^3\right\rangle(\theta_1,\theta_2,\theta_3)+\left\langle M^2M^*\right\rangle(\theta_1,\theta_3;\theta_2)\right.\\ &\quad\left.+\left\langle M^2M^*\right\rangle(\theta_2,\theta_3;\theta_1)-\left\langle M^2M^*\right\rangle(\theta_1,\theta_2;\theta_3)\right]/4\,,\\ \left\langle M_{\mathrm{ap}}M_\perp^2\right\rangle(\theta_1;\theta_2,\theta_3) &= \Re\left[-\left\langle M^3\right\rangle(\theta_1,\theta_2,\theta_3)+\left\langle M^2M^*\right\rangle(\theta_1,\theta_2;\theta_3)\right.\\ &\quad\left.+\left\langle M^2M^*\right\rangle(\theta_1,\theta_3;\theta_2)-\left\langle M^2M^*\right\rangle(\theta_2,\theta_3;\theta_1)\right]/4\,,\\ \left\langle M_\perp^3\right\rangle(\theta_1,\theta_2,\theta_3) &= \Im\left[-\left\langle M^3\right\rangle(\theta_1,\theta_2,\theta_3)+\left\langle M^2M^*\right\rangle(\theta_1,\theta_2;\theta_3)\right.\\ &\quad\left.+\left\langle M^2M^*\right\rangle(\theta_1,\theta_3;\theta_2)+\left\langle M^2M^*\right\rangle(\theta_2,\theta_3;\theta_1)\right]/4\,.\end{aligned}\tag{2.47}$$

These four distinct aperture statistics have very different physical interpretations. The first quantity $\langle M_{\mathrm{ap}}^3\rangle$ is a measure of the skewness of the E-mode of the shear field, or the gradient

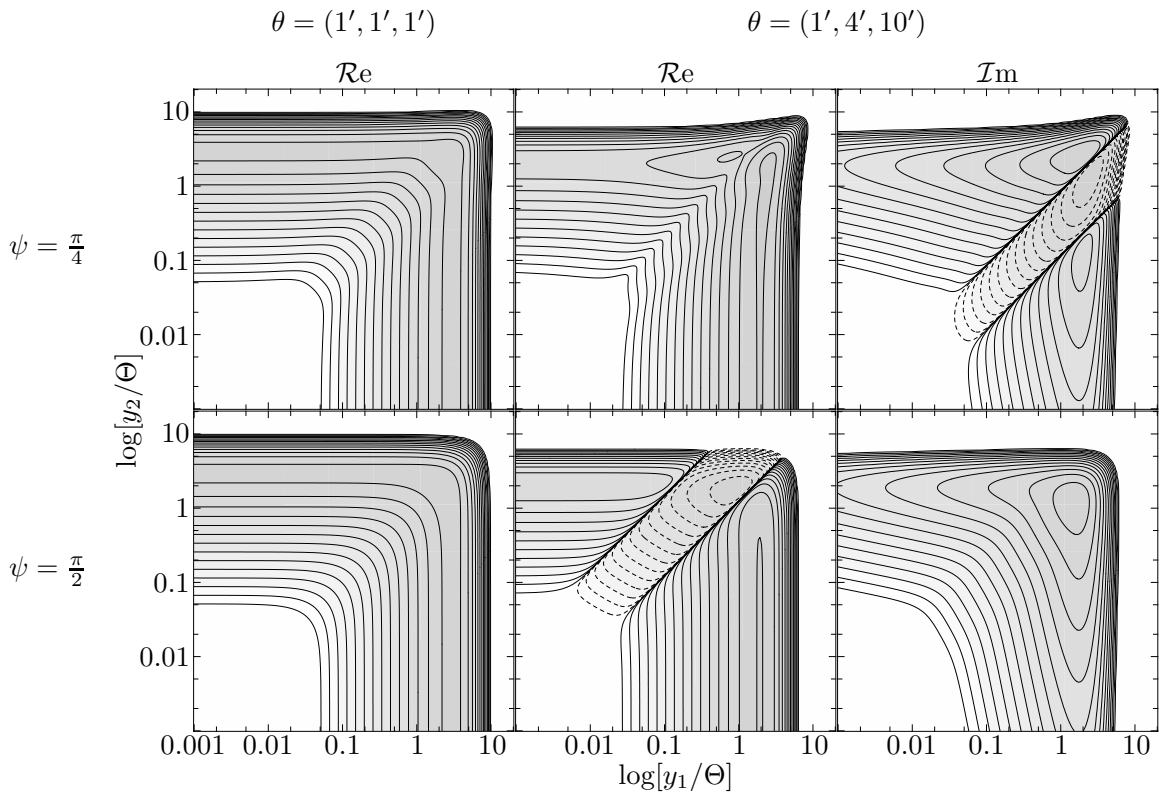

Figure 2.9: *Contours of the integral kernel $R_\theta^{(0)}$ (2.44) as a function of y_1 and y_2 for fixed ψ (upper row: $\psi = \pi/4$, lower row: $\psi = \pi/2$). The left-most of the three columns represents the case where all three aperture radii are equal. The function scales with the aperture radius. Note that the imaginary part vanishes due to symmetry. The two right columns show the real and imaginary part of the integrand for three different filter radii. The contour lines are logarithmically spaced with a factor of 5 between successive lines, starting with 10^{-10}. Dashed lines correspond to negative values.*

part of the corresponding convergence field. This measure is the interesting one in terms of cosmology, because it is generated by the skewness of the large-scale structure, i.e. due to non-linear clustering. It is this quantity which directly probes the bispectrum of cosmic density fluctuations, and thus will provide valuable information about structure formation on large scales and cosmological parameters. The third of the above defined expressions, $\langle M_{\rm ap} M_\perp^2 \rangle$, indicates the presence of a B-mode in the shear field which is correlated with the E-mode. A B-mode is believed to arise mainly from intrinsic alignment of source galaxies and from measurement errors such as imperfect PSF correction, or incorrect shape estimation of galaxies (see Sect. 1.4.3). The second and fourth of the quantities defined in (2.47), which have odd powers of $M_\perp$, are only non-zero if the underlying shear field is not parity-invariant (Schneider 2003). Since parity-violating fields cannot be explained by cosmic effects a significant non-zero measurement of these terms has to originate in measurement systematics.

The interrelations between the 3PCF and the third-order aperture mass allows one, like in the second-order case, to calculate the three-point correlation function from the data, making use of the complete available survey area, and integrate over the 3PCF to obtain the third-order measures for the E- and B-mode.

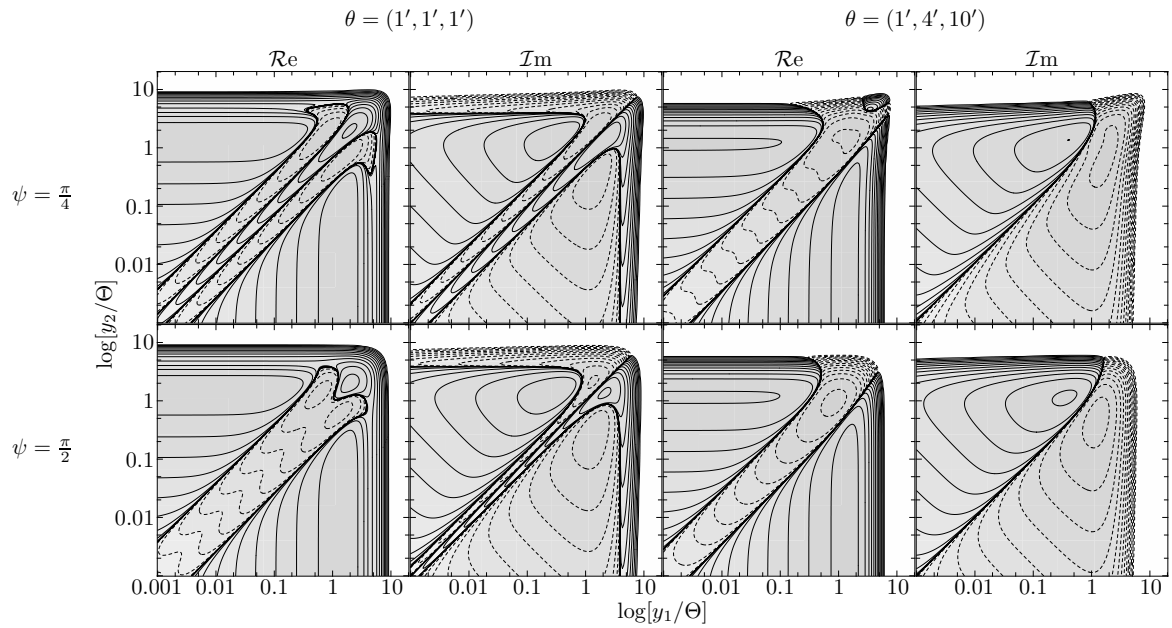

Figure 2.10: *Contours of the integrand kernel $R_\theta^{(3)}$ (2.44). The upper and lower row correspond to fixed values of $\psi = \pi/4$ and $\pi/2$, respectively. In the left two columns, the real and imaginary part of the function is shown for the three aperture radii being equal. The right two panels correspond to three different radii. The contours are the same as in Fig. 2.9.*

2.4 Shear estimators

Second-order shear estimators. In this section I define practical estimators of the various second- and third-order statistics of shear as introduced in the previous section. The shear and its moments (and functions of it) can be estimated from the observed background galaxy ellipticity. Recall the basic equation of weak lensing, $\varepsilon \approx \varepsilon^{\rm s} + \gamma$ (1.78), which states that the observed ellipticity of a galaxy (1.75) equals the sum of intrinsic ellipticity and shear. Assuming random orientations of the background galaxies, the observed ellipticity is an unbiased estimator of the shear, $\langle\varepsilon\rangle = \gamma$. From this, one can construct an unbiased estimator of the 2PCF of shear,

$$\hat{\xi}_\pm(\vartheta) = \frac{1}{N_p(\vartheta)} \sum_{ij} (\varepsilon_{i\rm t}\varepsilon_{j\rm t} \pm \varepsilon_{i\times}\varepsilon_{j\times}) \Delta_\vartheta(|\boldsymbol{\theta}_i - \boldsymbol{\theta}_j|), \tag{2.48}$$

where the 'bin selection' function is

$$\Delta_\vartheta(\theta) = \begin{cases} 1 & \text{for } |\ln\vartheta - \ln\theta| \le \frac{\Delta\ln\vartheta}{2} \\ 0 & \text{else} \end{cases} . \tag{2.49}$$

This is indeed an unbiased estimator of the 2PCF, since from (1.78) one sees that

$$\langle\varepsilon_{i\rm t}\varepsilon_{j\rm t} \pm \varepsilon_{i\times}\varepsilon_{j\times}\rangle = \sigma_\varepsilon^2\delta_{ij} + \xi_\pm(|\boldsymbol{\theta}_i - \boldsymbol{\theta}_j|), \tag{2.50}$$

where σ_ε is the dispersion of the intrinsic ellipticities of the source galaxies. The previous relation is valid under the assumption that the intrinsic orientations of galaxies are uncorrelated and the correlation between shear and intrinsic shape is negligible. Intrinsic alignment

of source galaxies is discussed in Sect. 1.4.3 and supposed to contribute only little to the shear signal. The correlation between the orientation of galaxies and shear has only recently been considered theoretically (Hirata & Seljak 2004). If there exists a correlation between the orientation of a foreground galaxy and its surrounding dark-matter host halo, the shear on a more distant galaxy image produced by this halo can experience a correlation with the intrinsic orientation of the first galaxy. This effect is most effective for projected close pairs of galaxies on the sky, but with quite different physical distances. The amplitude can exceed that of the intrinsic alignment signal depending on the model.

Both of the above mentioned contaminations to the shear signal can be further diminished if (photometric) redshift information for the galaxies are available. Then, close galaxy pairs suspicious of contributing to one of the above mentioned effects can be down-weighted or ignored.

Note that the estimator (2.48) is not optimal in the sense that all galaxies enter the sum with the same weight. When dealing with real data, however, it is convenient to assign individual weights which can depend on the estimated noise of the ellipticity measurements, or the uncertainty of the PSF correction and other factors. For simplicity, in this work all weights are set to unity.

For the dispersion of the aperture mass statistics (2.2, 2.5), an unbiased estimator can be constructed using (2.17),

$$\mathcal{M}_\pm(\theta) = \frac{1}{2\theta^2} \sum_{i=1}^{N(\theta)} \Delta\vartheta_i \vartheta_i \left[\hat{\xi}_+(\vartheta_i) T_+ \left(\frac{\vartheta_i}{\theta} \right) \pm \hat{\xi}_-(\vartheta_i) T_- \left(\frac{\vartheta_i}{\theta} \right) \right]. \tag{2.51}$$

An index has been attached to the bin width $\Delta\vartheta_i$ to account for variable bin widths (e.g. logarithmic bins). Note that for the polynomial filter (2.9) the sum extends to an upper bin index $N(\theta)$ such that $\vartheta_{N(\theta)} = 2\theta$. Using the Gaussian filter (2.10), the sum formally extends to infinity but can be truncated at around $\vartheta_{N(\theta)} \approx 6\theta$. With this definition, $\mathcal{M}_+$ and $\mathcal{M}_-$ are unbiased estimators for $\langle M_{\rm ap}^2 \rangle$ and $\langle M_\perp^2 \rangle$, respectively.

Third-order shear estimators. An unbiased estimator of the natural component $\Gamma^{(q)}, q = 0 \ldots 3$ of the 3PCF (2.24) is

$$\hat{\Gamma}^{(q)}(T_x) = \frac{1}{N_{\rm T}(T_x)} \sum_{ijk} \mathcal{E}^{(q)}(ijk) \Delta_{T_x}(ijk), \tag{2.52}$$

where T_x represents a triangle of points $\boldsymbol{\vartheta}_i, \boldsymbol{\vartheta}_j$ and $\boldsymbol{\vartheta}_k$, uniquely given e.g. by two side lengths x_1, x_2 and the angle ϕ_3 between them. $N_{\rm T}(T_x)$ is the number of triangles within the bin containing T_x and $\mathcal{E}^{(q)}(ijk)$ denotes the $q^{\rm th}$ natural component of the observed product of ellipticities from three galaxies at positions $\boldsymbol{\vartheta}_i, \boldsymbol{\vartheta}_j$ and $\boldsymbol{\vartheta}_k$, according to (2.24), e.g. for $q = 0$

$$\mathcal{E}^{(0)} = (\varepsilon_{\rm ttt} - \varepsilon_{\rm t\times\times} - \varepsilon_{\times\rm t\times} - \varepsilon_{\times\times\rm t}) + {\rm i}(\varepsilon_{\rm tt\times} + \varepsilon_{\rm t\times t} + \varepsilon_{\times\rm tt} - \varepsilon_{\times\times\times}), \tag{2.53}$$

where $\varepsilon_{\tau\sigma\rho} = \varepsilon_{\tau\sigma\rho}(\boldsymbol{\vartheta}_i, \boldsymbol{\vartheta}_j, \boldsymbol{\vartheta}_k) = \varepsilon_\tau(\boldsymbol{\vartheta}_i)\varepsilon_\sigma(\boldsymbol{\vartheta}_j)\varepsilon_\rho(\boldsymbol{\vartheta}_k)$ for $\tau, \sigma, \rho \in \{\text{`t'},\text{`}\times\text{'}\}$. The summation in (2.52) is performed over all possible triples of points $(\boldsymbol{\vartheta}_i, \boldsymbol{\vartheta}_j, \boldsymbol{\vartheta}_k)$ and $\Delta_{T_x}(ijk)$ is the triangle-bin selection function (cf. 2.49) — it is unity if the triangle given by $(\boldsymbol{\vartheta}_i, \boldsymbol{\vartheta}_j, \boldsymbol{\vartheta}_k)$ is in the same bin as T_x and zero otherwise. This estimator is used later (Sect. 2.5.2) in order to calculate the covariance of the 3PCF for the simple case of vanishing cosmic variance, thus, it only includes shot noise from the intrinsic galaxy ellipticities. Using this result, the variance of the aperture mass skewness is inferred and compared with numerical simulations.

2.5 Covariance of the shear estimators

Let x_i be an estimator of some statistics, e.g. of the two-point correlation function ξ_+ at an angular scale θ_i or of the second-order aperture mass $\langle M_{\rm ap}^2(\theta_i)\rangle$ for some aperture radius θ_i. The covariance matrix of this estimator is defined as

$$C(x)_{ij} = \langle x_i x_j\rangle - \langle x_i\rangle \langle x_j\rangle \,. \tag{2.54}$$

In this work, I will also use the cross-covariance between two different statistics x and y,

$$C(x,y)_{ij} = \langle x_i y_j\rangle - \langle x_i\rangle \langle y_j\rangle \,. \tag{2.55}$$

The variance of an estimator x_i is the diagonal of the covariance, $\mathrm{var}(x)_i = \langle x_i^2\rangle - \langle x_i\rangle^2 = C(x)_{ii}$.

In this section, I give analytical expressions for the covariance of various second-order estimators of shear statistics, which are exact for Gaussian fields. For estimators of third-order statistics, the variance is calculated in the simple limiting case of vanishing cosmic variance. The analytical expressions of the covariance for the second-order statistics will be the foundation for the likelihood and Fisher matrix analysis of cosmological parameter constraints in Chapter 3. In Chapter 5 covariance matrices of aperture mass estimators will be calculated where the ensemble average in (2.54) and (2.55) is approximated by the average over several (nearly) independent realizations of ray-tracing simulated shear fields.

2.5.1 Second-order measures

Two-point correlation function. Schneider et al. (2002) calculated the covariance matrices of the estimators defined in the last section, in the case of a Gaussian shear field. The covariance of the correlation function consists of three terms: A pure shot noise term (D), originating from the dispersion of the intrinsic ellipticities of the source galaxies which is present only on the diagonal, a cosmic variance term (V), and a mixed term (M),

$$\begin{aligned}
\mathsf{C}_{++} &= \mathsf{C}(\hat\xi_+,\hat\xi_+) = \mathsf{D} + \mathsf{M}_{++} + \mathsf{V}_{++},\\
\mathsf{C}_{--} &= \mathsf{C}(\hat\xi_-,\hat\xi_-) = \mathsf{D} + \mathsf{M}_{--} + \mathsf{V}_{--},\\
\mathsf{C}_{+-} &= \mathsf{C}(\hat\xi_+,\hat\xi_-) = \qquad\; \mathsf{M}_{+-} + \mathsf{V}_{+-}.
\end{aligned} \tag{2.56}$$

The individual terms involve summations over galaxy positions and are explicitely

$$\begin{aligned}
D(\vartheta_1,\vartheta_2) &= \frac{\sigma_\varepsilon^4}{F}\bar\delta(\vartheta_1-\vartheta_2)N_{\rm p}(\vartheta_1),\\
M_{++}(\vartheta_1,\vartheta_2) &= \frac{2\sigma_\varepsilon^2}{F}\sum_{ijk}\Delta^{ij}_{\vartheta_1}\Delta^{ik}_{\vartheta_2}\xi_+(jk),\\
V_{++}(\vartheta_1,\vartheta_2) &= \frac{1}{F}\sum_{ijkl}\Delta^{ij}_{\vartheta_1}\Delta^{kl}_{\vartheta_2}\Big(\xi_+(il)\xi_+(jk) + \cos[4(\varphi_{il}-\varphi_{jk})]\xi_-(il)\xi_-(jk)\Big),\\
M_{--}(\vartheta_1,\vartheta_2) &= \frac{2\sigma_\varepsilon^2}{F}\sum_{ijk}\Delta^{ij}_{\vartheta_1}\Delta^{ik}_{\vartheta_2}\cos[4(\varphi_{ij}-\varphi_{ik})]\xi_-(jk),\\
V_{--}(\vartheta_1,\vartheta_2) &= \frac{1}{F}\sum_{ijkl}\Delta^{ij}_{\vartheta_1}\Delta^{kl}_{\vartheta_2}\Big(\cos[4(\varphi_{ij}-\varphi_{il}-\varphi_{jk}+\varphi_{kl})]
\end{aligned} \tag{2.57}$$

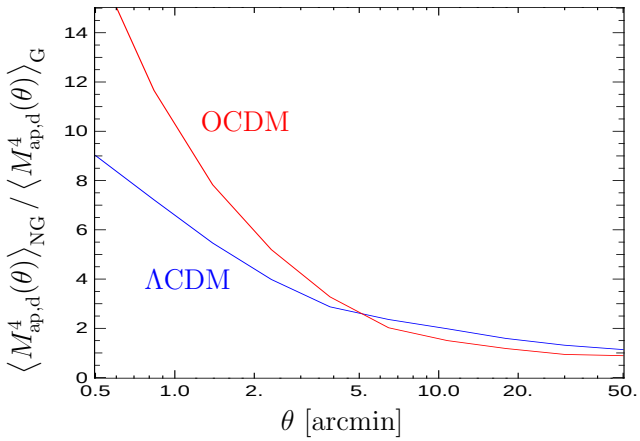

Figure 2.11: *The ratio of the fourth-order aperture mass statistics for the non-Gaussian and the Gaussian case using the OCDM and ΛCDM ray-tracing simulations, see Sect. 5.1. The latter is obtained using the Gaussianized ray-tracing simulations (Sect. 5.2.2).*

$$\times\,\xi_-(il)\xi_-(jk) + \cos[4(\varphi_{ij}-\varphi_{kl})]\xi_+(il)\xi_+(jk)\Big),$$

$$M_{+-}(\vartheta_1,\vartheta_2) = \frac{2\sigma_\varepsilon^2}{F}\sum_{ijk}\Delta^{ij}_{\vartheta_1}\Delta^{ik}_{\vartheta_2}\cos[4(\varphi_{ik}-\varphi_{jk})]\xi_-(jk),$$

$$V_{+-}(\vartheta_1,\vartheta_2) = \frac{2}{F}\sum_{ijkl}\Delta^{ij}_{\vartheta_1}\Delta^{kl}_{\vartheta_2}\cos[4(\varphi_{il}-\varphi_{kl})]\xi_-(il)\xi_+(jk),$$

with $F = N_{\rm p}(\vartheta_1)N_{\rm p}(\vartheta_2)$, $\Delta^{ij}_\vartheta = \Delta_\vartheta(|\boldsymbol{\theta}_i - \boldsymbol{\theta}_j|)$ and $\xi_\pm(ij) = \xi_\pm(|\boldsymbol{\theta}_i - \boldsymbol{\theta}_j|)$ for brevity. $\bar\delta(\vartheta_1 - \vartheta_2)$ is unity if ϑ_1 and ϑ_2 are in the same bin and zero otherwise. σ_ε is the ellipticity dispersion of the galaxies in the absence of shear.

I note here that the expressions for the cosmic variance terms **V** of the covariance only contain the Gaussian or unconnected part of the four-point correlator of shear. On scales below ~ 10 arc minutes the non-Gaussianity of the shear field gets important, see Fig. 2.11 and also Scoccimarro et al. (1999). On scales below 1 arc minute, the shot noise term **D** dominates over **V**, see Fig. 4 of Kilbinger & Schneider (2004), thus under the Gaussian assumption I expect to slightly underestimate the covariance in this transition regime between 1 and 10 arc minutes.

Second-order aperture-mass statistics. The covariance of the aperture mass dispersion is calculated from the covariance of the 2PCF using (2.17), see Schneider et al. (2002). The result is

$$C(\mathcal{M}_\pm;\theta_1,\theta_2) = \frac{1}{4\,\theta_1^2\theta_2^2}\sum_{i=1}^{I_1(\theta_1)}\sum_{j=1}^{I_2(\theta_2)}\Delta\vartheta_i\Delta\vartheta_j\vartheta_i\vartheta_j$$
$$\times\Bigg[T_+\left(\frac{\vartheta_i}{\theta_1}\right)T_+\left(\frac{\vartheta_j}{\theta_2}\right)C_{++}(\vartheta_i,\vartheta_j) + T_-\left(\frac{\vartheta_i}{\theta_1}\right)T_-\left(\frac{\vartheta_j}{\theta_2}\right)C_{--}(\vartheta_i,\vartheta_j)$$
$$\pm\,T_+\left(\frac{\vartheta_i}{\theta_1}\right)T_-\left(\frac{\vartheta_j}{\theta_2}\right)C_{+-}(\vartheta_i,\vartheta_j) \pm T_+\left(\frac{\vartheta_j}{\theta_2}\right)T_-\left(\frac{\vartheta_i}{\theta_1}\right)C_{+-}(\vartheta_j,\vartheta_i)\Bigg]. \qquad (2.58)$$

The upper limits of the sums I_1 and I_2 depend on the aperture radii θ_1 and θ_2, respectively, in the same way as in the equation for the $\langle M_{\rm ap}^2\rangle$-estimator (2.51).

2.5.2 Third-order measures

Three-point correlation function. The covariance of the estimator of the 3PCF $\hat{\Gamma}^{(q)}$ (2.52) consists of four terms, which are proportional to σ_ε^6, σ_ε^4, σ_ε^2 and σ_ε^0, respectively. The first term corresponds to pure ellipticity noise, the last term originates from cosmic variance alone and the two other terms are mixed contributions of both error sources. I consider here the simple case of vanishing cosmic variance, and analytically calculate the variance of the 3PCF and $M_{\rm ap,d}^3$. The only remaining term for the covariance of the 3PCF estimator $\hat{\Gamma}^{(q)}$ (2.52) in this case is the σ_ε^6-term — it reads

$$C(\hat{\Gamma}^{(q)},\hat{\Gamma}^{(p)};T_x,T_y) = \frac{1}{N_{\rm T}(T_x)N_{\rm T}(T_y)} \times \sum_{ijklmn} \left\langle \mathcal{E}^{(q,{\rm s})}(ijk)(\mathcal{E}^{(p,{\rm s})})^*(lmn)\right\rangle \Delta_{T_x}(ijk)\Delta_{T_y}(lmn), \tag{2.59}$$

where the superscript 's' indicates intrinsic ('source') ellipticity. The term in angular brackets is non-zero only if the two triangles given by $(\boldsymbol{\vartheta}_i,\boldsymbol{\vartheta}_j,\boldsymbol{\vartheta}_k)$ and $(\boldsymbol{\vartheta}_l,\boldsymbol{\vartheta}_m,\boldsymbol{\vartheta}_n)$ are identical (under the assumption that different galaxies are intrinsically uncorrelated), and factorizes into a sum of products of three two-point terms, each of the form $\langle\varepsilon_\tau(\boldsymbol{\vartheta}_i)\varepsilon_{\tau'}(\boldsymbol{\vartheta}_l)\rangle\delta_{il}\delta_{\tau\tau'}$. With $\langle\varepsilon_{\rm t}\varepsilon_{\rm t}\rangle = \langle\varepsilon_\times\varepsilon_\times\rangle = \sigma_\varepsilon^2/2$ and $\langle\varepsilon_{\rm t}\varepsilon_\times\rangle = 0$, the term in angular brackets becomes $8\cdot[\sigma_\varepsilon^2/2]^3 = \sigma_\varepsilon^6$. The sum reduces to a triple sum over $\Delta_{T_x}(ijk)$ which is just the number of triangles in the respective bin. Finally, one gets

$$C(\hat{\Gamma}^{(q)},\hat{\Gamma}^{(p)};T_x,T_y) = \frac{\sigma_\varepsilon^6}{N_{\rm T}(T_x)}\bar{\delta}(T_x,T_y), \tag{2.60}$$

where $\bar{\delta}(T_x,T_y)$ is zero if the two triangles T_x and T_y are in different bins, and unity otherwise. This result will be used in the next section in order to calculate the variance of the aperture mass skewness.

Third-order aperture mass statistics. The covariance of the generalized third-order aperture mass $M_{\rm ap}^3(\theta_i,\theta_j,\theta_k)$ depends on six scalar quantities, namely the 2×3 filter scales involved. In order to obtain a two-dimensional matrix, I relabel all non-degenerate combinations of filter triplets ($\langle M_{\rm ap}^3\rangle$ is invariant under permutations of its arguments) with a single index. The resulting $\langle M_{\rm ap}^3(\theta_1,\theta_2,\theta_3)\rangle$-vector is organized such that $(\theta_1,\theta_2,\theta_3)$ is in lexical order, I further demand that $\theta_1\le\theta_2\le\theta_3$. Note that the labeling order does not play a role in the later analysis. For a number of N distinct filter scales, there are $\binom{N+2}{3} = N(N+1)(N+2)/6$ different combinations.

I define the covariance matrix $\mathsf{C}(M_{\rm ap}^3)$ for the generalized third-order aperture mass statistics, and $\mathsf{C}(M_{\rm ap,d}^3)$ for the skewness of $M_{\rm ap}$, which is a function of only one filter scale, $M_{\rm ap,d}^3(\theta) = M_{\rm ap}^3(\theta,\theta,\theta)$.

The covariance of $M_{\rm ap}^3$ can be obtained by integrating over the covariance of the 3PCF, in analogy to the second-order case. This can be done analytically for the simple exemplary case of vanishing cosmic variance, of all six aperture radii being equal (this corresponds to the

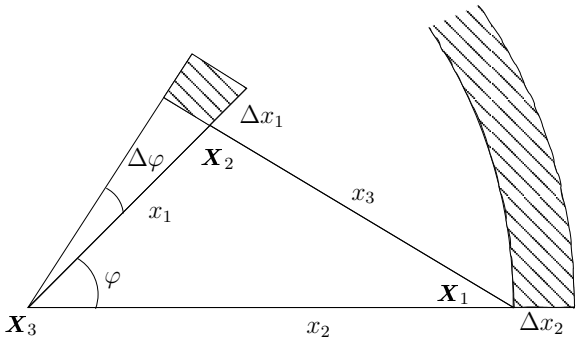

Figure 2.12: *Sketch for the determination of the number of galaxy triplets N_T for a triangle $T_x = (x_1, x_2, \varphi)$ within a bin given by $(\Delta x_1, \Delta x_2, \Delta\varphi)$. $N_\mathrm{T}(T_x)$ is the number of positions $\boldsymbol{X}_3$ times the number of galaxies in the annulus at $\boldsymbol{X}_1$ times the galaxy number in a small area at $\boldsymbol{X}_2$.*

variance of $M^3_\mathrm{ap,d}$), and in the absence of a B-mode, in which case $\langle M^3_\mathrm{ap}\rangle = M^3$ (see eq. 2.47). Then

$$\mathrm{var}(M^3_\mathrm{ap,d};\theta) = \frac{1}{(48\pi\theta^4)^2}\int \mathrm{d}x_1\, x_1 \int \mathrm{d}x_2\, x_2 \int \mathrm{d}\varphi \int \mathrm{d}y_1\, y_1 \int \mathrm{d}y_2\, y_2 \int \mathrm{d}\psi \\ \times C(\hat\Gamma^{(0)}_\mathrm{b}, \hat\Gamma^{(0)}_\mathrm{b}; T_x, T_y) R^{(0)}_\theta(x_1, x_2, \varphi) R^{(0)}_\theta(y_1, y_2, \psi). \tag{2.61}$$

Here, the triangles T_x and T_y are given by (x_1, x_2, φ) and (y_1, y_2, ψ), respectively. Inserting (2.60) yields the function $\bar\delta$ in the integrand, which can be used to eliminate three integrals since for any function f

$$\int \mathrm{d}y_1 \int \mathrm{d}y_2 \int \mathrm{d}\psi f(y_1, y_2, \psi)\bar\delta(T_x, T_y) = f(x_1, x_2, \varphi)\Delta x_1 \Delta x_2 \Delta\varphi, \tag{2.62}$$

where Δx_1, Δx_2, $\Delta\varphi$ are given by the bin size in which the triangle T_x is situated.

For simplicity, I assume that boundary effects due to the finite field size A can be neglected. The number of galaxy pairs at the triangle points $\boldsymbol{X}_3$ and $\boldsymbol{X}_1$ with distance $\in [x_2; x_2 + \Delta x_2]$ is the number of galaxies N times the area of an annulus multiplied by the galaxy number density n, $N \cdot (2\pi x_2 \Delta x_2\, n)$. To get the number of galaxy triples, one has to multiply this by the number of galaxies in a small area at position X_2, which is $x_1 \Delta x_1 \Delta\varphi\, n$, and one gets $N_\mathrm{T}(T_x) = 2\pi A n^3 x_1 \Delta x_1 x_2 \Delta x_2 \Delta\varphi$, see also Fig. 2.12. Thus,

$$\mathrm{var}(M^3_\mathrm{ap,d};\theta) = \frac{\sigma_\varepsilon^6}{2\pi A n^3}\frac{1}{(48\pi\theta^4)^2}\int \mathrm{d}x_1 \int \mathrm{d}x_2 \int \mathrm{d}\varphi\, R^2_\theta(x_1, x_2, \varphi). \tag{2.63}$$

Solving the integral, one gets

$$\mathrm{var}(M^3_\mathrm{ap,d},\theta) = \frac{11}{15552\,\pi^2}\frac{\sigma_\varepsilon^6}{A n^3\,\theta^4} \\ = 10^{-16}\left(\frac{\sigma_\varepsilon}{0.3}\right)^6\left(\frac{A}{9\,\mathrm{deg}^2}\right)^{-1}\left(\frac{n}{25\,\mathrm{arcmin}^{-2}}\right)^{-3}\left(\frac{\theta}{\mathrm{arcmin}}\right)^{-4}. \tag{2.64}$$

Note that the variance of M^2_ap (Schneider et al. 2002) has the same dependence on the observed area A, but is only quadratically inverse as a function of both n and θ. I compare this analytical result with numerical simulations in Fig. 5.3.

Chapter 3

Study of the geometry of cosmic shear surveys

The second-order cosmic shear statistics as defined in Chapter 2 are used extensively in this part of the thesis and their properties and dependence on cosmology are studied. Estimators of these statistics and their covariance are employed to predict constraints on cosmological parameters from a cosmic shear survey. With the primary goal to minimize the errors on cosmological parameters, I address the question concerning the influence of the survey design on the determination of various cosmological parameters. In particular, I study the survey geometry which comprehends the distribution of lines of sight and the angular range of the survey. The measurement of cosmic shear with a sufficiently high precision to constrain cosmological parameters requires many independent lines of sight to keep the sampling or "cosmic variance" low. On the other hand, it is important to measure the shear on a large range of angular scales. In order to design a cosmic shear survey under the inevitable constraint of limited telescope time, one has to carefully choose the geometry of the survey and find a trade-off between a low-cosmic-variance survey containing many independent but sparsely sampled lines of sight and fewer but densely covered sky regions providing a larger signal-to-noise over the measured angular range. I compare different survey settings with respect to their capability of constraining cosmological parameters. To that end, I simulate cosmic shear surveys for a variety of realistic survey strategies and noise characteristics and calculate the corresponding covariance matrices of second-order shear statistics by Monte-Carlo-integrating the analytical expressions (2.56 – 2.58). The parameters of the simulated surveys are described in Sect. 3.1. I use various combinations of second-order statistics (Sect. 3.2) and investigate their ability to constrain cosmological parameters of non-linear models of large-scale structure formation (Sect. 1.2.7). Various techniques such as likelihood analysis (Sect. 3.3), the Fisher information matrix (Sect. 3.4) and Karhunen-Loève (KL) eigenmode analysis are applied in order to quantitatively determine the accuracy with which cosmological parameters can be obtained from cosmic shear measurements. Moreover, the degeneracies between parameters are studied. I review the KL technique and introduce its application to second-order cosmic shear statistics in Sect. 3.5 and present numerical results in Sect. 3.6.

3.1 Survey characteristics

I consider several different survey strategies for the analysis presented in this chapter. Firstly, individual images are distributed randomly but non-overlapping within circular patches on the sky, where 'image' means one individual field of view of size $13' \times 13'$. This field of view corresponds to medium-sized wide-field imaging cameras, e.g. VIMOS at the 8.2m ESO-VLT telescope UT3 (Melipal). The random placement of images accounts for the fact that for realistic surveys, bright stars and foreground galaxies should be avoided and therefore, sparse sampling of a sky region is necessary. It is assumed that the shear correlation can be

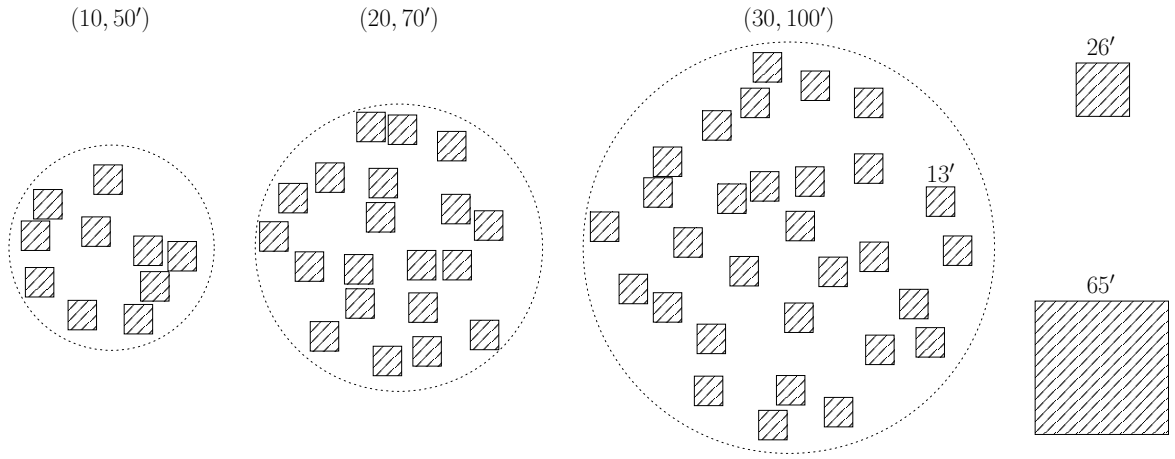

Figure 3.1: *Realizations of circular patches (N, R) where N is the number of images per patch and R the patch radius. The squares represent the individual images with a side length of 13 arc minutes. For comparison, two fields of size $26' \times 26'$ and $65' \times 65'$, respectively, are shown on the right side.*

measured across image boundaries, up to scales as large as the patch diameter. The individual patches are assumed to be separated by at least several degrees. On these scales, the shear correlation is virtually zero, so different patches can be considered as uncorrelated; they probe statistically independent parts of the large-scale structure.

A survey consists of P patches of radius R, each patch containing N images. The total number of images, $n = PN = 300$ is kept fixed for all surveys, corresponding to 14.1 square degrees for the entire survey. For N being a factor of n the values 10, 20, 30, 50, and 60 are considered, corresponding to geometries with $P = 30$, 15, 10, 6 and 5 patches, respectively. Throughout, I denote these patch geometries by the two numbers (N, R). An illustration of some patches is given in Fig. 3.1.

The patch geometries are compared to a configuration which consists of 300 single uncorrelated images, where 'uncorrelated' means separated by at least several degrees as before. This configuration has the smallest possible cosmic variance, but the shear correlation can be only measured up to $\sqrt{2} \cdot 13$ arc minutes. In the following, I denote this survey design with $300 \cdot 13'$.

Two more survey strategies of single uncorrelated fields are used later for the KL eigenmode analysis of a cosmic shear survey, see Sect. 3.6. The first survey consists of 75 uncorrelated $26' \times 26'$ square images — typical fields of view for current wide-field cameras like WFI on the ESO 2.2-m telescope. The second survey represents 12 independent $65' \times 65'$-fields, corresponding to new-generation wide-field cameras like MegaCam/CFHT. I refer to these surveys with $75 \cdot 26'^2$ and $12 \cdot 65'^2$, respectively.

If not indicated otherwise, the number density of source galaxies is $n_{\rm gal} = 30\,{\rm arcmin}^{-2}$. This number density of high-redshift galaxies which are usable for weak lensing shape measurements can be achieved with high-quality ground-based imaging data from a 4 m-class telescope. The source galaxy ellipticity dispersion is $\sigma_\varepsilon = 0.3$, if not stated otherwise.

3.2 The input data

For the analysis of survey strategies I use the following four combinations of second-order shear statistics as input data $\boldsymbol{x}$:

1. The two-point correlation function $\xi_+(\theta_i)$ measured at various angular scales θ_i with its covariance matrix C_{++}.

2. The 2PCF $\xi_-(\theta_i)$ with its covariance C_{--}.

3. The combination of the two 2PCF ξ_+ and ξ_- to the joint data vector $\xi_{\rm tot} = (\xi_+(\theta_1), \ldots, \xi_+(\theta_N), \xi_-(\theta_1), \ldots, \xi_-(\theta_N))$, with the joint covariance matrix, written as the block matrix
$$\mathsf{C}_{\rm tot} = \begin{pmatrix} \mathsf{C}_{++} & \mathsf{C}_{+-} \\ \\ \mathsf{C}_{+-}^{\rm t} & \mathsf{C}_{--} \end{pmatrix}. \tag{3.1}$$

4. The aperture mass statistics $\langle M_{\rm ap}^2(\theta_i)\rangle$ corresponding to the polynomial filter (2.9) for various aperture radii θ_i and its covariance $\mathsf{C}(\mathcal{M}_+)$.

The 2PCF is calculated from the convergence power spectrum using eq. (2.12); the aperture mass dispersion is obtained from the 2PCF, see eq. (2.17). The covariance matrices (2.57) are calculated using a Monte-Carlo integration. To this end, galaxy positions are simulated for a realistic survey geometry and number density — these positions are used to perform the summation in (2.57). In order not to introduce artificial Poisson noise, the galaxies are not distributed randomly but subrandomly onto the survey fields, see Chapter 7 of Press et al. (1992). Because the number of operations increases with the galaxy number to the fourth power, it is not feasible, except for a very small survey area, to sum over all galaxy positions. Instead, a random subsample of galaxy positions is used (for more details see Kilbinger & Schneider 2004).

The covariance matrices depend on the survey strategy in various ways. First of all, the sampling of angular scales is encoded in the covariance: The better a particular scale is sampled by some survey geometry and the more galaxy pairs contribute to the estimate of the correlation function, the smaller is the noise for this angular scale. Secondly and related to the previous point, the relative contribution to the noise coming from cosmic variance and intrinsic ellipticity dispersion depends on the survey, in particular, on the number of independent directions on the sky. In general, a survey with a dense angular sampling will have a larger cosmic variance than a survey with the same total area but sparse sampling of a larger range of angular scales or of more independent sky regions.

For the likelihood and Fisher matrix analyses (Sects. 3.3 and 3.4), I use 20 angular logarithmic bins for the 2PCF, the smallest separation between two galaxies considered being 0.2 arc minutes. Since the calculation of the covariance is very time-consuming, I have chosen a rather small number of angular bins. The largest angular distance is determined by the survey setting — for a patch geometry, it is chosen to be equal to the patch diameter R. The joint 2PCF data vector $\xi_{\rm tot}$ has 40 entries, corresponding to the two times 20 angular bins for ξ_+ and ξ_-. The aperture mass dispersion $\langle M_{\rm ap}^2\rangle$ is calculated for 18 different radii.

For the Karhunen-Loève analysis (Sect. 3.6), I discard the largest bin for the 2PCF. This is to avoid very large angular separations with small numbers of galaxy pairs which caused the

results to be unstable when the binning was varied (Sect. 3.6.6). Therefore, ξ_+ and ξ_- are sampled for 19 bins and $\xi_{\rm tot}$ for 38 angular separations. For the aperture mass dispersion, I take into account 14 filter scales.

3.3 Likelihood analysis of survey strategies

In order to compare different weak lensing survey strategies regarding their potential to determine cosmological parameters, a measure of the precision is needed with which the parameters can be constrained. I choose a likelihood function for an n-dimensional multivariate Gaussian distribution,

$$\mathcal{L}(\boldsymbol{x};\boldsymbol{p}) = \frac{1}{(2\pi)^n\sqrt{\det \mathsf{C}(\boldsymbol{p})}} \exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}(\boldsymbol{p}))^{\rm t}\mathsf{C}^{-1}(\boldsymbol{p})(\boldsymbol{x}-\boldsymbol{\mu}(\boldsymbol{p}))\right], \tag{3.2}$$

where $\boldsymbol{\mu}$ is the mean of the data vector $\boldsymbol{x}$, $\langle\boldsymbol{x}\rangle = \boldsymbol{\mu}$ and C is the covariance, $\mathsf{C} = \langle(\boldsymbol{x}-\boldsymbol{\mu})(\boldsymbol{x}-\boldsymbol{\mu})^{\rm t}\rangle$. Both the mean and the covariance can depend on some model parameter $\boldsymbol{p} = (p_1, p_2, \ldots, p_m)$. The assumption that the data is drawn from a Gaussian probability distribution is justified since non-Gaussian contributions are small (see the comment in Sect. 2.5.1).

For the likelihood analysis (and the Fisher matrix study in the next section) I assume the covariance to be constant and not to depend on $\boldsymbol{p}$. This is a valid approximation; as shown in Sect. 3.6.5, the contribution of the parameter-dependent covariance term to the Fisher matrix (3.6) is small in comparison to the parameter-dependent mean term. Next and not least important, the numerical calculation of the covariance is very time-consuming; were it non-constant, an adequate sampling of the likelihood in parameter space would not be feasible.

A 'figure-of-merit' is defined as

$$\chi^2(\boldsymbol{p}) = (\boldsymbol{x}-\boldsymbol{\mu}(\boldsymbol{p}))^{\rm t}\mathsf{C}^{-1}(\boldsymbol{x}-\boldsymbol{\mu}(\boldsymbol{p})), \tag{3.3}$$

which in the case of constant covariance is — up to an additive constant — equal to $-2\ln\mathcal{L}$.

Since the likelihood analysis of cosmic shear survey strategies has been discussed in Kilbinger (2002) and Kilbinger & Schneider (2004), I will only give a short summary here. In this analysis, I vary two parameters at a time while keeping the others fixed. The parameters under consideration are $\Omega_{\rm m}, \sigma_8, \Gamma$ and the source redshift parameter z_0.

In all cases, $\xi_{\rm tot}$ yields the best constraints on the cosmological parameters, closely followed by ξ_+, which dominates the contribution to $\xi_{\rm tot}$. Next in the list is the aperture mass statistics, the worst constraints come from ξ_- alone.

I compare the likelihood contours in the resulting two-dimensional parameter space for the individual survey strategies (see Figs. 5 and 6 of Kilbinger & Schneider 2004). The differences between the patch geometries are rather small, the uncorrelated images survey yield the best parameter constraints for $\xi_{\rm tot}$ and ξ_+. This implies that the way of sampling is not a dominant factor, and that the measurement of the shear on large angular scales can be left out in favor of a small cosmic variance, as is the case of the $300\cdot13'$ survey.

The statistics ξ_- and $\langle M_{\rm ap}\rangle$ show a different behavior to ξ_+. They are more sensitive to large angular scales, thus they provide far better constraints for patch geometries than for the $300\cdot13'$ survey with its lack of large-scale information.

This simple case of the likelihood analysis in the presence of two unknown parameters is just an showcase example to stress the differences in the survey strategies and the behavior of the second-order statistics. A more thorough and realistic study has to include more parameters. Since a full likelihood analysis on a higher-dimensional parameter space is extremely time-consuming, I constrain myself to a local neighborhood around the maximum of the likelihood. To this purpose the Fisher matrix is introduced in the following section.

3.4 Fisher matrix analysis of survey strategies

The Fisher matrix (Kendall & Stuart 1969; Tegmark et al. 1997) is defined as

$$F_{\alpha\beta} = \left\langle \frac{\partial^2[-\ln\mathcal{L}]}{\partial p_\alpha \partial p_\beta} \right\rangle = \left(\frac{\partial^2[-\ln\mathcal{L}]}{\partial p_\alpha \partial p_\beta} \right)_{\boldsymbol{p}=\boldsymbol{p}_0}, \tag{3.4}$$

where the likelihood $\mathcal{L}$ depends on the model parameter vector $\boldsymbol{p} = (p_1, \ldots, p_m)$ and $\boldsymbol{p}_0$ denotes the "true" parameter value. The Fisher matrix is the expectation value of the Hessian matrix of $(-\ln\mathcal{L})$ at $\boldsymbol{p} = \boldsymbol{p}_0$, which coincides on average with $\mathcal{L}$'s maximum if the maximum likelihood estimator of p_0 is unbiased (Tegmark et al. 1997) — thus it is a local measure of how fast $\mathcal{L}$ falls off from the maximum.

The smallest possible variance σ of any unbiased estimator of some parameter p_α is given by the Cramér-Rao inequality

$$\sigma(p_\alpha) \geq \sqrt{(F^{-1})_{\alpha\alpha}}\,; \tag{3.5}$$

the expression on the right-hand side is called the *minimum variance bound* (MVB). If only one parameter is considered, (3.5) simplifies to $\sigma(p_\alpha) \geq 1/\sqrt{F_{\alpha\alpha}}$.

In the case where the data is drawn from a multi-variate Gaussian probability distribution (3.2), the general expression for the Fisher matrix is

$$F_{\alpha\beta} = \frac{1}{2}\mathrm{tr}[\mathsf{C}^{-1}\mathsf{C}_{,\alpha}\mathsf{C}^{-1}\mathsf{C}_{,\beta} + \mathsf{C}^{-1}\mathsf{M}_{\alpha\beta}], \tag{3.6}$$

where $\mathsf{M}_{\alpha\beta} = \boldsymbol{\mu}_{,\alpha}\boldsymbol{\mu}_{,\beta}^{\mathrm{t}} + \boldsymbol{\mu}_{,\beta}\boldsymbol{\mu}_{,\alpha}^{\mathrm{t}}$ and the comma notation as abbreviation for partial derivatives is introduced, i.e. $\mathsf{C}_{,\alpha} = \partial\mathsf{C}/\partial p_\alpha$.

In the case of a constant covariance matrix, $\mathsf{C}_{,\alpha} = 0$ (see also previous section), eq. (3.6) simplifies to

$$F_{\alpha\beta} = \boldsymbol{\mu}_{,\alpha}^{\mathrm{t}}\mathsf{C}^{-1}\boldsymbol{\mu}_{,\beta}. \tag{3.7}$$

In the case of a simple survey geometry, e.g. a single large field of view, where border effects of the survey can be neglected, the covariance for all shear estimators scales with A^{-1} where A is the observed area. Thus, the MVB is approximately proportional to $1/\sqrt{A}$. The derivatives $\boldsymbol{\mu}_{,\alpha}$ of the shear statistics are calculated throughout this work based on the theoretical models for the power- and the bispectrum as described in Sect. 1.4.2. The numerical derivatives are obtained using polynomial extrapolation of finite differences, see Chapter 5.7 of Press et al. (1992). For the KL analysis, in particular in Sect. 3.5.2, also the derivatives of the covariance are needed. These are easily calculated from eq. (2.57), by differentiating ξ_+ and ξ_- in these expressions. The shot-noise term D, being independent of cosmology, does not contribute, and the cosmic variance term V, being quadratic in the correlation function, has a larger influence than the mixed term M.

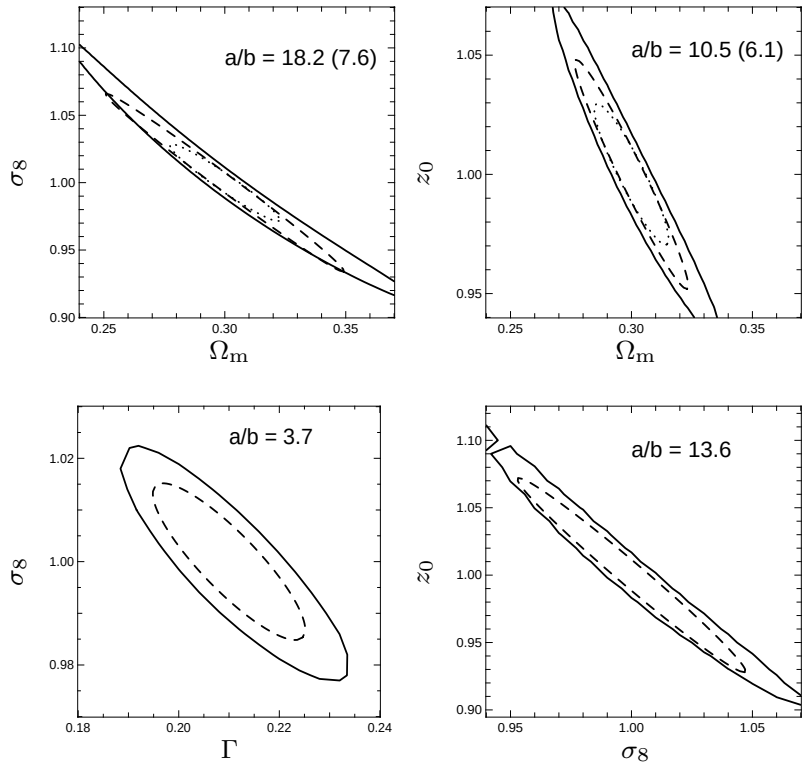

Figure 3.2: *1σ-likelihood contours (solid lines) using the log-likelihood (3.3) compared with the 1σ-error ellipse from the Fisher matrix (3.6), for the combined 2PCF ξ_{tot}: The dashed ellipses are for a flat Universe (as it is the case for the likelihood contours), the dotted ones are for $\Omega_\Lambda = 0.7$. a/b is the axis ratio of the ellipses (the case $\Omega_\Lambda = 0.7$ is in parentheses). The configuration $300 \cdot 13'$.*

3.4.1 Comparison of the Fisher matrix with the likelihood function

I compare the minimum variance bound (3.5) with the 1σ-levels of the likelihood function (3.2) for the combined 2PCF ξ_{tot}. Only two parameters are varied, the others are kept fixed, in which case the Fisher information matrix reduces to a 2×2 matrix, which can be visualized as 1σ-error ellipses. This comparison is shown in Fig. 3.2. As expected from the Cramér-Rao inequality, the likelihood contours are larger than the 1σ-ellipse from the Fisher matrix. The orientation of the Fisher error ellipse coincides with the likelihood shape, i.e. the direction of the minimal and maximal degeneracy of parameters is recovered. The larger the degeneracy between two parameters, the larger the deviation between the local approximation by the Fisher matrix and the likelihood function. In the case of a flat Universe where the sum of Ω_m and Ω_Λ is restricted to unity, the degeneracy is much larger than for a fixed cosmological constant $\Omega_\Lambda = 0.7$. I checked the correctness of the local approximation of the likelihood at the 0.1σ-level, and found an excellent match between the likelihood contours and the Fisher error ellipses.

3.4.2 More parameters

After having considered only two varying parameters, I calculate the minimum variance bound for three and more parameters out of $(\Omega_m, \sigma_8, \Gamma, \Omega_\Lambda, n_s)$ simultaneously. In general, the aper-

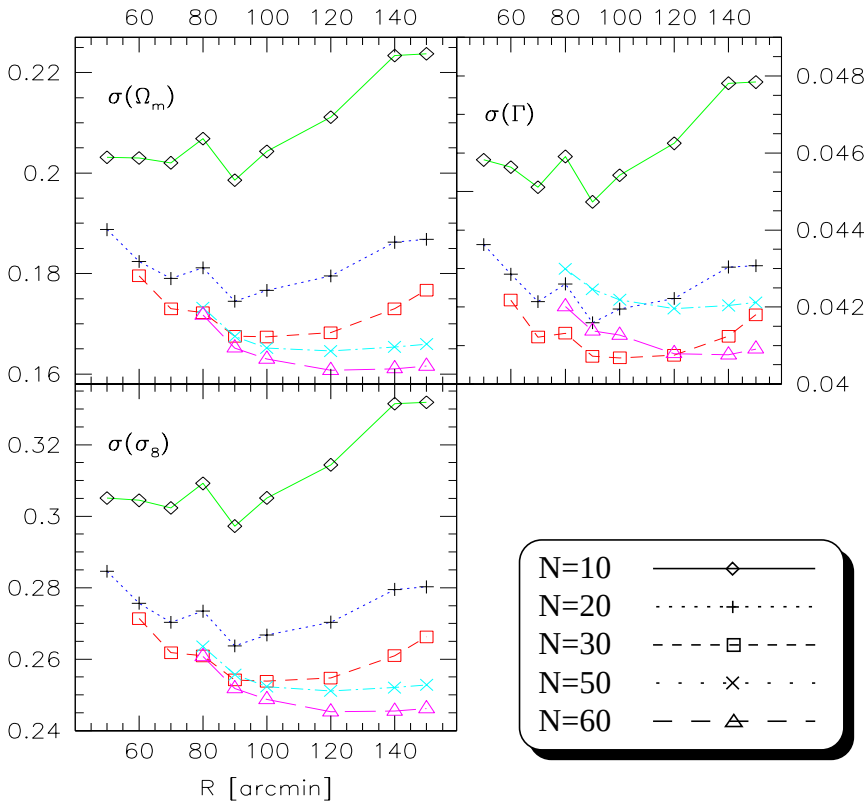

Figure 3.3: *The minimum variance bound for the parameters Ω_m, Γ and σ_8, for the $\langle M_{ap}^2 \rangle$-statistics. All hidden parameters are kept fixed and a flat Universe is assumed. Each point in the plot represents a patch geometry with N images in patches of radius R. The MVB for the $300 \cdot 13'$-survey is given in Table 3.1.*

ture mass gives less tight constraints than the combined two-point correlation functions (Kilbinger & Schneider 2004). However, this difference gets smaller the more parameters are included. Furthermore, with the aperture mass statistics and its ability to separate the E- from the B-mode of the shear field one can contol systematic measurement errors (Sect. 1.4.3) and therefore, I will focus on this statistics for the remaining part of the Fisher matrix analysis.

Figures 3.3 - 3.5 show the MVB for a different number of free cosmological parameters for the individual patch geometries. The fiducial parameters correspond to model 1 (Table B.1). In the cases where Ω_Λ is not a free parameter, the prior is a flat Universe ($\Omega_\Lambda = 1 - \Omega_m$).

When taking into account three or more parameters, the uncorrelated image configuration gives very poor constraints on these parameters. The MBV for this survey strategy is in most cases more than twice the value of the MBV for the least optimal patch geometry. The reason is that the lack of medium- and large-scale information dramatically raises the degeneracy between parameters. The advantage of having a small cosmic variance cannot compensate for this missing out of the large angular scales for this survey. The more parameters are to be determined from the data the stronger is the degradation of the errors.

When Ω_Λ is added as a free parameter (compare Fig. 3.3 with Fig. 3.4), the variance of the shape parameter increases by more than a factor of two, whereas the variances of Ω_m and

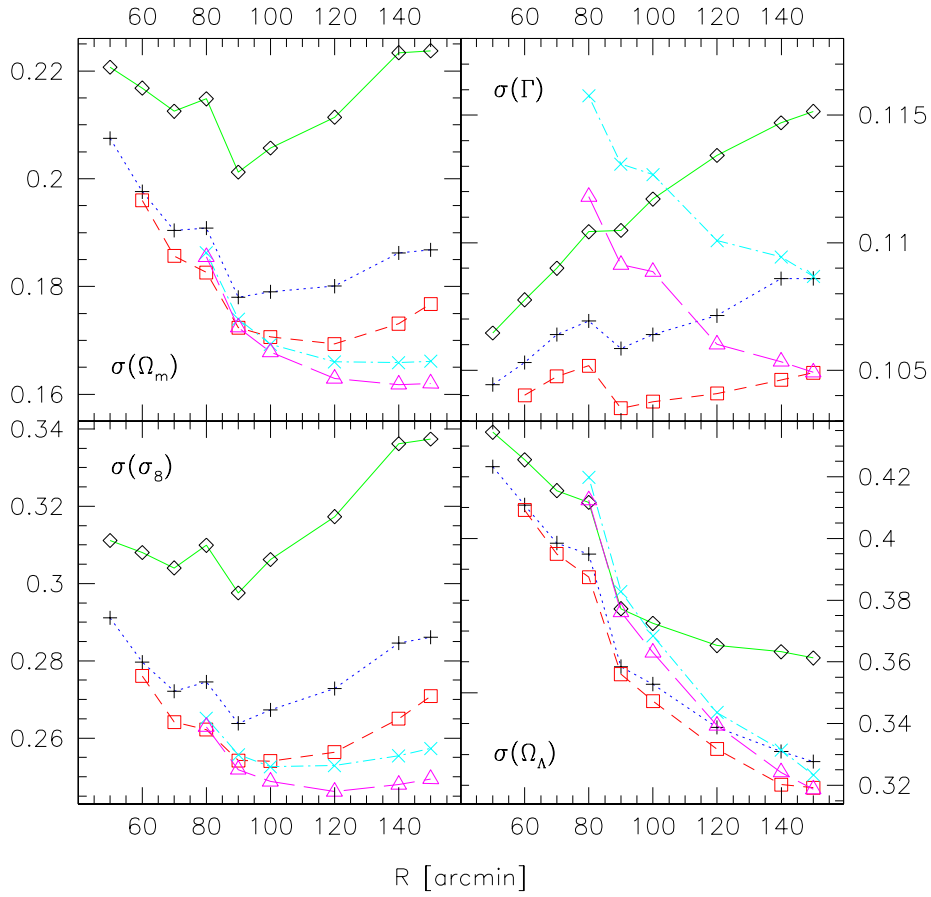

Figure 3.4: *The minimum variance for the parameters $\Omega_\mathrm{m}, \Gamma, \sigma_8$ and Ω_Λ using the $\langle M_\mathrm{ap}^2 \rangle$-statistics. See Fig. 3.3 for more details. The MVB for the $300 \cdot 13'$-survey is given in Table 3.1.*

σ_8 are only slightly deteriorated. For large patches, the variances of the three parameters $\Omega_\mathrm{m}, \sigma_8$ and Γ are just rescaled, whereas for small patches, the change is more complex. The constraint on the cosmological constant is very poor, confirming the results of Bernardeau et al. (1997).

When the spectral index n_s is added (Fig. 3.5), the minimum variance bound of Ω_m and σ_8 again increase in the same way as before, when Ω_Λ was added to the analysis, although this time by a greater amount. The variance on Γ changes completely, taking a similar functional form on R and N as the variance on n_s. The reason for this is that both parameters determine the shape of the power spectrum in about the same way and, therefore, the dependence of cosmic shear on these two parameters is very similar (see also Sect. 3.6.3). By contrast, Ω_m and σ_8 influence mainly the amplitude of the power spectrum.

For each data point in Figs. 3.3 - 3.5, only one realization for each of the P patches is used, corresponding to one survey with N images in $P = 300/N$ patches. Due to the (quasi-)random positions of the images in each patch there is some scatter in the resulting covariance and consequently, in the Fisher matrix and the MVBs which are shown in these figures. In order to quantify this scatter, I create two new surveys by bootstrapping from a larger number of

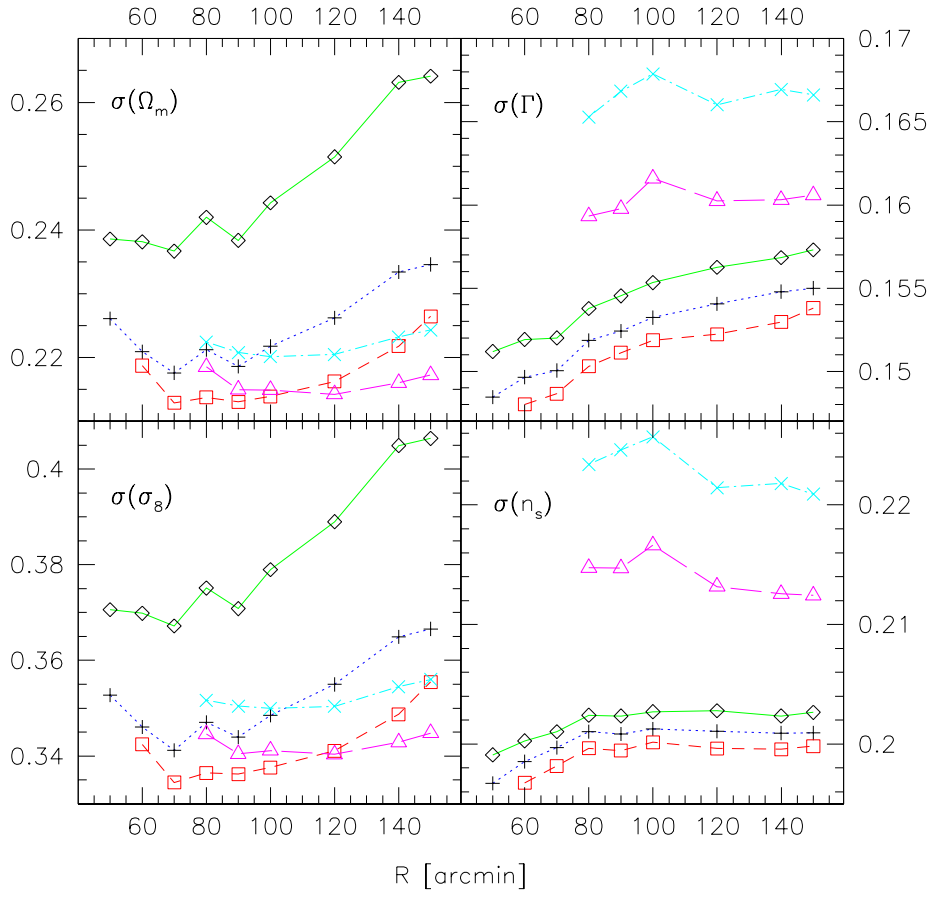

Figure 3.5: *The minimum variance for the parameters $\Omega_\mathrm{m}, \Gamma, \sigma_8$ and n_s using the $\langle M_\mathrm{ap}^2 \rangle$-statistics. See Fig. 3.3 for more details. The MVB for the $300 \cdot 13'$-survey is given in Table 3.1.*

additional realizations for two of the patch geometries. I calculate the rms of the MVB from the bootstrapping sample and find the scattering of the minimum variance bound to be at a level of about one percent.

From Fig. 3.3, the best geometry is a $(60, 120')$-patch survey. Considering Fig. 3.5, a configuration with $N = 30$ and small R yields the best minimum variance bounds. For both cases, a survey with 30 images per patch and a patch radius of around 100 arc minutes seems to be a good choice. The patch radius has only a small influence on the minimum variance bound, more important is the number of images per patch and therefore the number of independent patches.

The difference in the minimum variance bound between individual patch geometries make up to 25 percent. Thus, the conclusion of the Fisher matrix analysis presented in this section is that by an appropriate choice of the survey strategy, an improvement for the determination of cosmological parameters of about 25 percent can be made for each parameter.

Table 3.1: *The minimum variance for several combinations of parameters, for the 300 · 13′ configuration consisting of 300 uncorrelated images. In each row, those parameters which have an entry are assumed to be determined from the data, the other parameters are fixed. The counterpart of the three rows for the patch geometries are the Figs. 3.3 - 3.5.*

Ω_m	σ_8	Γ	Ω_Λ	n_s
0.53	0.76	0.09		
0.53	0.76	0.14	0.64	
0.53	0.77	0.16		0.20

3.5 Karhunen-Loève (KL) eigenmodes technique

Modern cosmological experiments yield a huge number of data that have to be processed and analyzed. Therefore, it is essential to study data compression techniques which can efficiently reduce the data vector to a lower-dimensional space. The effect of such a compression on the Fisher information matrix can be seen by analysing the Karhunen-Loève (KL) eigenmodes (Karhunen 1946; Loève 1948) associated with the Fisher matrix. These techniques are already in use in cosmic microwave background studies, see e.g. Bond (1995) for COBE/FIRAS[1] data and Tegmark et al. (1997) for a general review of the KL eigenmode analysis. In the context of galaxy surveys, Hamilton et al. (2000) studied PSCz[2] with KL techniques and Szalay et al. (2003) used this method in the context of SDSS. Watkins et al. (2002) employed these techniques to analyze the peculiar velocity surveys for removal of small-scale, non-linear modes from large-scale linear modes which retain cosmological information. Recent studies by Huterer & White (2005) have underlined its importance for filtering unwanted non-cosmological information at small angular scales.

The data vector in weak lensing surveys which is used to extract cosmological information is usually the shear correlation at various angular separations, estimated from the shapes of the observed background galaxies. Thus, it represents an already compressed data set for which likelihood analyses are feasible although further compression can speed up the analysis. Moreover, for third-order statistics, the number of measured triangle configurations can be very large even in a binned way and and further compression of the data might be necessary to place constraints in a high-dimensional cosmological parameter space. Therefore, the study of KL methods in the context of cosmic shear will be very useful for future weak lensing surveys.

Being not only an effective tool for data compression, the KL eigenmode analysis can also help to understand the specific linear combination of angular scales and the redshift range that contribute most significantly to an estimator. A KL eigenmode analysis of weak lensing observables in redshift space has already been done in the context of cosmic shear by Heavens (2003), where the distribution of observed ellipticities was directly used as data vector. I extend these previous analyses to the second-order shear statistics $\xi_\pm$, ξ_tot and $\langle M_\mathrm{ap}^2 \rangle$ (Sect. 3.2) as functions of projected angular scales and for realistic survey geometries. Before I present the results of this KL study in Sect. 3.6, the basic ideas of data compression and its effect on the Fisher matrix are reviewed, following Tegmark et al. (1997).

[1] Far Infrared Absolute Spectrophotometer
[2] Point Source Catalogue redshift

3.5.1 Data compression and the Fisher matrix

A general linear data compression can be written as

$$\tilde{\boldsymbol{x}} = \mathsf{T}\boldsymbol{x}, \tag{3.8}$$

where from the n-dimensional data vector $\boldsymbol{x}$ a new $\tilde{n}$-dimensional data set $\tilde{\boldsymbol{x}}$ is constructed via the $\tilde{n} \times n$ matrix T. This means that the expectation value $\langle\boldsymbol{x}\rangle$ of the original data vector $\boldsymbol{x}$ transforms to $\langle\tilde{\boldsymbol{x}}\rangle = \mathsf{T}\langle\boldsymbol{x}\rangle$ and hence the covariance $\mathsf{C} = \langle\boldsymbol{x}\boldsymbol{x}^{\mathrm{t}}\rangle - \langle\boldsymbol{x}\rangle\langle\boldsymbol{x}\rangle^{\mathrm{t}}$ changes accordingly to $\tilde{\mathsf{C}} = \mathsf{T}\mathsf{C}\mathsf{T}^{\mathrm{t}}$. An example for a data compression is the dispersion of the aperture mass statistics $\langle M_{\mathrm{ap}}^2\rangle$ (Sect. 2.2.2) which is a linearly compressed version of the correlation functions ξ_+ and ξ_-. The compression matrix T depends in this case on the functions T_+ and T_- (2.17).

The Fisher matrix F (3.6) corresponding to the original data vector $\boldsymbol{x}$ changes to $\tilde{\mathsf{F}}$, given by

$$\tilde{F}_{\alpha\beta} = \frac{1}{2}\mathrm{tr}\left[(\mathsf{T}\mathsf{C}\mathsf{T}^{\mathrm{t}})^{-1}(\mathsf{T}\mathsf{C}_{,\alpha}\mathsf{T}^{\mathrm{t}})(\mathsf{T}\mathsf{C}\mathsf{T}^{\mathrm{t}})^{-1}(\mathsf{T}\mathsf{C}_{,\beta}\mathsf{T}^{\mathrm{t}}) + (\mathsf{T}\mathsf{C}\mathsf{T}^{\mathrm{t}})^{-1}(\mathsf{T}\mathsf{M}_{\alpha\beta}\mathsf{T}^{\mathrm{t}})\right]. \tag{3.9}$$

Clearly, when the dimensionality n and $\tilde{n}$ of the original and the transformed data vectors are the same and T has full rank, the transformation is a similarity transformation and the Fisher matrix remains unchanged, $\tilde{\mathsf{F}} = \mathsf{F}$.

Usually, the goal of a KL analysis is to find a compression matrix T for $\tilde{n} < n$ with a minimal loss of information corresponding to a minimal increase of the resulting errors on cosmological parameters. Regarding only one parameter p_α with the minimum variance bound (3.5) as a lower bound of its error bar, one seeks to maximize the Fisher matrix element $\tilde{F}_{\alpha\alpha}$.

Following TTH, I consider the two simple cases of a constant mean $\boldsymbol{\mu}$ and a constant covariance C, where 'constant' means with respect to the parameter vector $\boldsymbol{p}$. The optimization is done first for one parameter and later for several parameters simultaneously with the aid of a singular value decomposition (SVD).

For cosmic shear, the assumption of a constant covariance seems quite natural, since in most cases of weak lensing observations the covariance is not calculated analytically but extracted directly from the data (e.g. van Waerbeke et al. 2002) and thus it is constructed to be constant and not dependent on cosmological parameters. However, the situation is different for the mean, since with a model parameter independent mean it is not possible to even define a likelihood function. And although within the Fisher matrix formalism, a constant mean can be considered in a consistent way, this mean first has to be found using e.g. the maximum likelihood estimator. But the distinction into a constant mean and a constant covariance case is of rather technical nature, and the general case of parameter-dependent mean and covariance can be considered by combining the two, see Sect. 3.5.4. Thus, a strategy to obtain cosmological parameters and their minimum variance bound from data would be first to define a likelihood function $\mathcal{L}$ using a parameter dependent mean (and parameter dependent covariance if desired) and to find the maximum likelihood parameter $\boldsymbol{p}_0$. Secondly, the Fisher matrix is calculated by evaluating the second derivatives of $\mathcal{L}$ at the point $\boldsymbol{p} = \boldsymbol{p}_0$. A KL analysis of this Fisher matrix resulting in an optimal data compression can be undertaken by considering the two cases of constant mean and covariance independently, and then combining both cases. Using the compressed data vector and its covariance, a detailed likelihood analysis is now possible because of the reduced dimensionality after compression.

Note that in this case the "mean" is a second-order statistics (the shear correlation), and the covariance is of fourth order (although reduced to only depend on second-order because of the assumption of a Gaussian shear field). I therefore choose a different approach than e.g. Heavens (2003) who directly use the galaxy ellipticities as data vector. In their case, the (zero) mean is of first order, and the covariance contains the information about the power spectrum.

3.5.2 Constant mean

In the case where the mean $\boldsymbol{\mu}$ is known and constant, the second term in (3.6) vanishes. To begin with, one considers a compression matrix which consists of a single row vector, $\mathsf{T} = \boldsymbol{b}^{\mathrm{t}}$. The diagonal element of the compressed Fisher matrix (3.9) which determines the error bar of the parameter p_α becomes

$$\tilde{F}_{\alpha\alpha} = \frac{1}{2}\left(\frac{\boldsymbol{b}^{\mathrm{t}}\mathsf{C}_{,\alpha}\boldsymbol{b}}{\boldsymbol{b}^{\mathrm{t}}\mathsf{C}\boldsymbol{b}}\right). \tag{3.10}$$

The task is now to find a compression $\boldsymbol{b}^{\mathrm{t}}$ which maximizes this expression. Without loss of generality the compression vector $\boldsymbol{b}$ can be normalized such that $\boldsymbol{b}^{\mathrm{t}}\mathsf{C}\boldsymbol{b} = 1$. The maximization problem can then be solved by maximizing the numerator of (3.10) with the side constraint of the denominator equating to unity. Introducing a Lagrangian multiplier, this is equivalent to maximizing

$$\boldsymbol{b}^{\mathrm{t}}\mathsf{C}_{,\alpha}\boldsymbol{B} - \lambda\boldsymbol{b}^{\mathrm{t}}\mathsf{C}\boldsymbol{b}. \tag{3.11}$$

Differentiating with respect to $\boldsymbol{b}$ and setting the result to zero, one obtains the generalized eigenvalue problem

$$\mathsf{C}_{,\alpha}\boldsymbol{b} = \lambda\mathsf{C}\boldsymbol{b}. \tag{3.12}$$

Since C is symmetric and positive definite there exists an invertible matrix L with $\mathsf{C} = \mathsf{L}\mathsf{L}^{\mathrm{t}}$. Using this *Cholesky decomposition* (e.g. Press et al. 1992), the previous equation can be reduced to an ordinary eigenvalue problem,

$$(\mathsf{L}^{\mathrm{t}}\mathsf{C}_{,\alpha}\mathsf{L}^{-\mathrm{t}})\mathsf{L}^{\mathrm{t}}\boldsymbol{b} = \lambda\mathsf{L}^{\mathrm{t}}\boldsymbol{b}. \tag{3.13}$$

Solving this equation for all n orthogonal eigenvectors $\mathsf{L}^{\mathrm{t}}\boldsymbol{b}_k, k = 1\ldots n$ results in n real eigenvalues λ_k. A compression matrix T is then constructed containing as row vectors the first $\tilde{n}$ eigenvectors which have been sorted by the absolute value of their corresponding eigenvalues. These KL eigenmodes satisfy an orthogonality relation and the new data vector $\tilde{\boldsymbol{x}}$ is *statistically orthogonal*, $\langle\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}^{\mathrm{t}}\rangle = \mathbb{1}$. Hence, $\tilde{\boldsymbol{x}}$ is a vector of independent unit variance random variables with a diagonal covariance matrix $\tilde{\mathsf{C}}$.

Eigenmodes with high eigenvalues or low rank (row) numbers contain more information about the parameter p_α than ones with small (absolute) eigenvalues and high rank numbers, which contain almost no additional information. As will be shown in Sect. 3.6, a compression by a factor of nearly two is typically achieved for the independent analysis of various parameters which yields about the same MVB as for the uncompressed case.

The matrix T constructed in this way represents a set of eigenvectors, rank-ordered according to their signal-to-noise ratio: An individual eigenmode $\boldsymbol{b}_k$ contributes to the measurement error for the parameter p_α as $\delta p_\alpha = 1/|\lambda_k|$. Thus, the signal-to-noise ratio for this mode therefore is $p_\alpha(\delta p_\alpha)^{-1} = p_\alpha|\lambda_k|$.

The new Fisher matrix diagonal element $\tilde{F}_{\alpha\alpha}$ corresponding to a KL compression with $\tilde{n}$ eigenmodes is simply (TTH)

$$\tilde{F}_{\alpha\alpha} = \frac{1}{2}\sum_{k=1}^{\tilde{n}} \lambda_k^2. \tag{3.14}$$

In this analysis (Sect. 3.6), I repeat the data compression with different mode numbers $\tilde{n}$ and plot for various cosmological parameters p_α the associated error $\Delta p_\alpha = 1/(\tilde{F}_{\alpha\alpha})^{1/2}$ as a function of $\tilde{n}$, see Fig. 3.6. Clearly, the error is a decreasing function of the mode number $\tilde{n}$. However, it reaches a constant plateau for some $\tilde{n}_0 < n$, thus the original error for the uncompressed case is recovered before all modes are used for the parameter estimation. I find in most of the cases that a plateau is reached for $\tilde{n}_0 \lesssim n/2$, thus a compression factor of nearly two is possible without any loss of information. Note that the compression factor depends of course on the original number of bins — I comment on the binning in Sect. 3.6.6.

Since in the case of a constant mean, the Fisher matrix contains products of the inverse covariance and the (derivative of the) covariance, it is independent of the survey area, and only very slightly sensitive to the survey geometry.

3.5.3 Constant covariance

When C is independent of p_α, the first term in (3.9) vanishes and one wishes to maximize the quantity

$$\tilde{F}_{\alpha\alpha} = \frac{\boldsymbol{b}^{\mathrm{t}}\mathsf{M}_{\alpha\alpha}\boldsymbol{b}}{\boldsymbol{b}^{\mathrm{t}}\mathsf{C}\boldsymbol{b}}. \tag{3.15}$$

Proceeding as in the case of a constant mean, one introduces a Lagrangian multiplier and finds the eigenproblem

$$(\mathsf{L}^{-1}\mathsf{M}_{\alpha\alpha}\mathsf{L}^{-\mathrm{t}})\mathsf{L}^{\mathrm{t}}\boldsymbol{b} = \lambda\mathsf{L}^{\mathrm{t}}\boldsymbol{b} \tag{3.16}$$

to be equivalent to the maximization problem. However, the matrix $\mathsf{M}_{\alpha\alpha} = 2\boldsymbol{\mu}_{,\alpha}\boldsymbol{\mu}_{,\alpha}^{\mathrm{t}}$ is merely of rank one, since any two rows (and columns) of $\mathsf{M}_{\alpha\alpha}$ are linearly dependent and proportional to $\sum_j(\mu_{,\alpha})_j$. Therefore, the left-hand side of the previous equation, $2(\mathsf{L}^{-1}\boldsymbol{\mu}_{,\alpha})(\boldsymbol{\mu}_{,\alpha}^{\mathrm{t}}\boldsymbol{b})$, maps $\boldsymbol{b}$ onto a one-dimensional subspace of $\mathbb{R}$ spanned by $\mathsf{L}^{-1}\boldsymbol{\mu}_{,\alpha}$ and the only non-trivial eigenvector is $\mathsf{L}^{\mathrm{t}}\boldsymbol{b}_0 = \mathsf{L}^{-1}\boldsymbol{\mu}_{,\alpha}$ with eigenvalue $\lambda_0 = 2|\mathsf{L}^{-1}\boldsymbol{\mu}_{,\alpha}| = \mathrm{tr}\left[\mathsf{C}^{-1}\mathsf{M}_{\alpha\alpha}\right]$.

Inserting this eigenvector $\boldsymbol{b}_0 = \mathsf{C}^{-1}\boldsymbol{\mu}_{,\alpha}$ into the compression matrix T (3.9), one finds that the modified Fisher matrix is the same as the original one, $\tilde{F}_{\alpha\beta} = F_{\alpha\beta} = \boldsymbol{\mu}_{,\alpha}^{\mathrm{t}}\mathsf{C}^{-1}\boldsymbol{\mu}_{,\beta}$. Thus, in contrast to the constant mean case, there exists only one mode and this mode contains all information.

If geometrical effects of the survey are neglected, the covariance is proportional to one over the survey area A. Thus, in the case of constant covariance the Fisher matrix is proportional to A, in contrast to the constant mean case, where F is independent of A. For reasonably large cosmic shear surveys, the second term in (3.6) is therefore dominant over the first one. In Sect. 3.6.5 this statement will be quantified further.

3.5.4 General case

TTH describe how the general case (when neither mean nor covariance is constant) can be treated efficiently, by simply adding the one eigenmode from the constant covariance case to the $\tilde{n}$ modes from the constant mean analysis. However, as mentioned in the previous section,

the constant covariance eigenmode is (for reasonably large survey areas) dominant over all the other modes and contains the bulk part of the information about cosmological parameters. This is also true for the likelihood: In Sect. 3.6.7, it is shown that the constant mean modes contribute negligibly to the likelihood of the parameters Ω_m and σ_8 in comparison with the constant covariance modes.

3.5.5 Joint parameter estimation

For the independent estimation of a parameter p_α, the KL method is devised to minimize the associated error bar $\Delta p_\alpha = (\tilde{F}_{\alpha\alpha})^{-1/2}$ in a suitably rotated basis. However, for the case of joint parameter estimation, the objects to be minimized are $\Delta p_\alpha = (\mathsf{F}^{-1})^{1/2}_{\alpha\alpha}$ which is a more demanding optimization problem. Therefore, I adopt an alternative approximate technique described in TTH as follows. The individual compression matrices T_α optimized for the independent estimation of the parameters p_α are arranged into a new matrix after multiplying each row (corresponding to an eigenvector) with its corresponding eigenvalue. This can be written as the $n \times nm$-dimensional matrix $\mathsf{T} = (\mathbf{\Lambda}_1\mathsf{T}_1^\mathrm{t}, \ldots, \mathbf{\Lambda}_m\mathsf{T}_m^\mathrm{t})$ where $\mathbf{\Lambda}_\alpha = \mathrm{diag}(\lambda_{\alpha 1}, \ldots, \lambda_{\alpha n})$ is a diagonal matrix containing the eigenvalues corresponding to the individual eigenanalysis of the α^th parameter. T will contain a lot of redundant information, because of unavoidable near-degeneracies of parameters measured from cosmic shear, e.g. between Ω_m and σ_8 or between Γ and n_s. Thus, the individual compression matrices contain linearly (nearly) dependent eigenvectors.

In order to separate useful from redundant information, this new matrix T is factorized using a singular value decomposition (SVD), $\mathsf{T} = \mathsf{U}\mathbf{\Sigma}\mathsf{V}^\mathrm{t}$ (Press et al. 1992, Sect. 2.6) according to the following scheme,

$$ {}_n\Big\{ \overbrace{\Big(\mathbf{\Lambda}_1\mathsf{T}_1^\mathrm{t} \,\Big|\, \cdots \,\Big|\, \mathbf{\Lambda}_m\mathsf{T}_m^\mathrm{t} \Big)}^{nm} = \overbrace{\Big(\mathsf{U} \Big)}^{n} \cdot \overbrace{\Big(\mathbf{\Sigma} \Big)}^{n} \cdot \overbrace{\Big(\qquad \mathsf{V}^\mathrm{t} \qquad \Big)}^{nm} \Big\}_n , \tag{3.17} $$

where the $n \times n$-dimensional matrix U is orthogonal, $\mathsf{U}^\mathrm{t}\mathsf{U} = \mathsf{U}\mathsf{U}^\mathrm{t} = \mathbb{1}$, and the $n \times nm$-matrix V^t is column-orthogonal, $\mathsf{V}^\mathrm{t}\mathsf{V} = \mathbb{1}$. $\mathbf{\Sigma} = \mathrm{diag}(\sigma_i)$ is a diagonal matrix containing the *singular values*, which can be interpreted as generalized eigenvalues. Such a decomposition can be found for any matrix T uniquely up to simultaneous permutations of the columns of $\mathsf{U}, \mathbf{\Sigma}$ and V^t, and, if two or more singular values σ_i and σ_j are equal, up to linear combinations of the corresponding columns i and j.

After sorting the columns of U^t according to the modulus of the corresponding singular value, in a similar manner to the way is was done for the single parameter compression matrices T_α, the final compression matrix is $\mathsf{T}_\mathrm{joint} = \mathsf{U}^\mathrm{t}$. The columns of U with corresponding non-zero singular values form an orthonormal set of basis vectors of the same space that is spanned by all initial compression vectors of the single parameter analysis. Modes with high singular values contain the bulk information about the cosmological parameters, whereas column vectors of U corresponding to vanishing or very small singular values capture redundant or almost redundant information. Inserting $\mathsf{T}_\mathrm{joint}$ into (3.9), the Fisher matrix after the final compression is easily calculated. Based on the amplitude of the singular values, one can choose a final mode number $\tilde{n}_\mathrm{joint} < n$ fixing the compression factor. In the same way as for the single parameter estimation the error bars Δp_α can be plotted as a function of the compression factor and the saturation plateau can be used to determine the final mode number, see Fig. 3.6. In the analysis presented here, I find that the result does not depend on

whether the individual T_α contain all $\tilde{n}$ modes or whether modes with small eigenvalues have been discarded before the construction of T. Thus, the SVD is a robust method to separate useful from redundant information in a given data set.

Also for the constant covariance case (Sect. 3.5.3), a joint parameter estimation analysis can be done. The m eigenvectors resulting from the m individual parameter analyses are combined to the $n \times m$-matrix T_0, after having multiplied them by their respective eigenvalues. The singular value decomposition of T_0 will result in m singular values, which yield information of the degree of degeneracy between the parameters. I will discuss results for this analysis in Sect. 3.6.

3.5.6 Window functions

The components $\tilde{x}_i$ of the new data vector contain pairwise independent and uncorrelated information about cosmology, which in this case is contained in the convergence power spectrum. Using the KL eigenmode technique, one can study in detail how the power spectrum is sampled by the various second-order statistics in order to yield uncorrelated data points. By comparing modes corresponding to high and low eigenvalues respectively, one can see which scales carry much and which ones carry little of the cosmological information.

For any second-order statistics, the dependence on the power spectrum is encoded in a window function. For the two components of the shear 2PCF ξ_+ and ξ_- these are the broad-band Bessel functions J_0 and J_4, respectively, see eq. (2.12). The filter function corresponding to the aperture mass statistics $\langle M_\mathrm{ap}^2\rangle$ is the narrow-peaked function $[24\,\mathrm{J}_4(\eta)/\eta^2]^2$, see eq. (2.17).

For each component $\tilde{x}_i$ of the new data vector I define a new window function W_i which is a linear combination of the original filter functions associated with different smoothing scales θ_j. In the case of the 2PCF, the new data vector is

$$\tilde{\xi}_i^\pm = T_{ij}^\pm \xi_\pm(\theta_j) = \frac{1}{2\pi}\int_0^\infty \mathrm{d}\ell\,\ell\,P_\kappa(\ell) W_i^\pm(\ell); \quad W_i^\pm(\ell) = \sum_j T_{ij}^\pm \mathrm{J}_{0,4}(\ell\theta_j). \tag{3.18}$$

Similarly, for $\langle\tilde{M}_\mathrm{ap}^2\rangle$ one obtains the new window functions

$$W_i^\mathrm{E}(\ell) = \sum_j T_{ij}^\mathrm{E}[24\,\mathrm{J}_4(\ell\theta_j)/(\ell\theta_j)^2]^2. \tag{3.19}$$

In the case of the combined vector $\tilde{\xi}^\mathrm{tot} = (\tilde{\xi}^+, \tilde{\xi}^-)$, the window functions W_i^tot are linear combinations of W_i^+ and W_i^-.

In case of a constant mean, the rows of the matrix T contain the $\tilde{n}$ transposed eigenvectors $\boldsymbol{b}$, see Sect. 3.5.2. For the constant covariance case, there is only one eigenmode, in this case $\mathsf{T}_0 = \boldsymbol{\mu}_{,\alpha}^\mathrm{t}\mathsf{C}^{-\mathrm{t}}$ if the compression is optimized for the parameter p_α. I denote all quantities corresponding to the constant covariance case with a subscript '0', e.g. $\tilde{\xi}_0^+$, W_0^+ etc.

The objective here is to study these window functions for different survey geometries and noise characteristics. This will give insights about how the convergence power spectrum is sampled in order to constrain cosmological parameters with cosmic shear, and how different Fourier modes are probed by a given survey strategy.

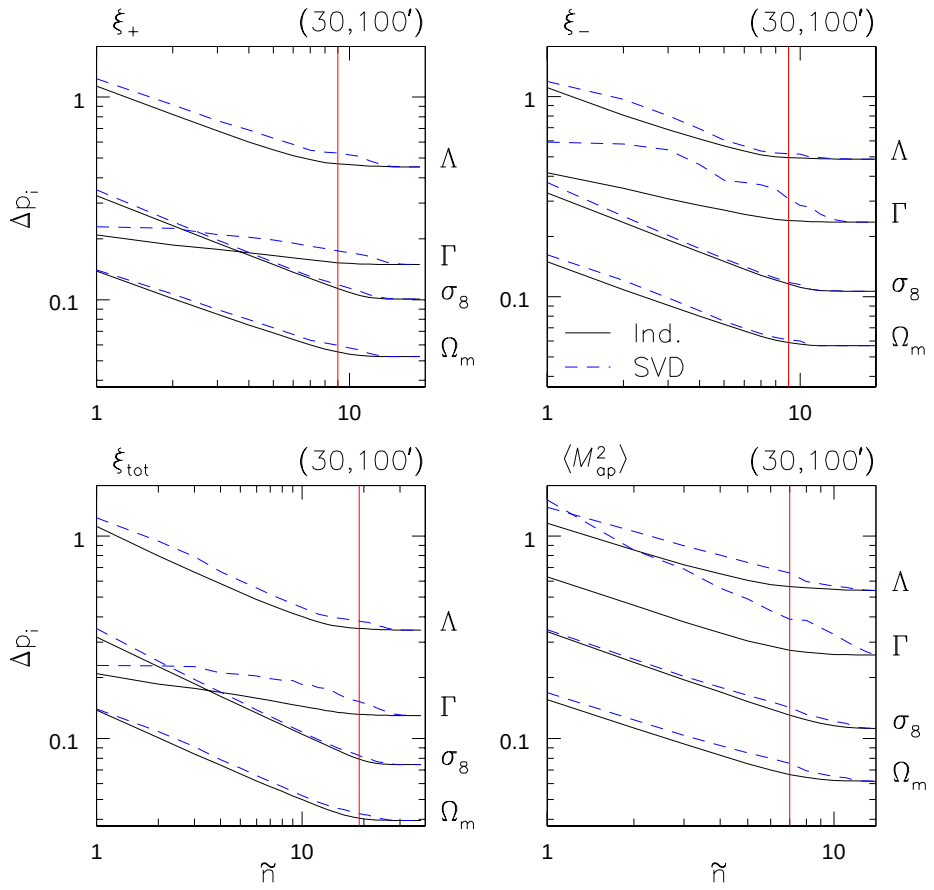

Figure 3.6: *$\Delta p_\alpha = (\tilde{F}_{\alpha\alpha})^{-1/2}$ as a function of the mode number $\tilde{n}$ is plotted, according to (3.14) for the constant mean case and various cosmological parameters. Solid lines correspond to the independent parameter estimation (Sect. 3.5.2), dashed lines represent the joint parameter analysis (Sect. 3.5.5). The four panels correspond to the shear statistics ξ_+, ξ_-, their combination $\xi_{\rm tot}$ and $\langle M_{\rm ap}^2\rangle$ (see Sect. 3.2) The survey geometry is $(30, 100')$. The maximum mode number n equals the number of bins, which is 38 for $\xi_{\rm tot}$, 19 for both ξ_+ and ξ_- and 14 for $\langle M_{\rm ap}^2\rangle$.*

3.6 KL eigenmode analysis

In this section, the impact of various survey strategies, noise structures and the effect of binning are considered in the determination of the KL eigenmodes. I separately study the constant mean case (see Sect. 3.5.2) and the constant variance case (see Sect. 3.5.3). Eigenmodes associated with the independent analysis of a single parameter and the joint analysis of several parameters (see Sect. 3.5.5) are investigated.

3.6.1 Error bars

Constant mean For the case of a constant mean (Sect. 3.5.2), the MVB $\Delta p_\alpha = (\tilde{F}_{\alpha\alpha})^{-1/2} = (2/\sum_{k=1}^{\tilde{n}} \lambda_k^2)^{1/2}$ is plotted in Fig. 3.6 as a function of the compression mode number $\tilde{n}$. Similar to the results of TTH for the case of CMB anisotropies, the errors for the single parameter estimation show a power-law decrease until some mode number $\tilde{n}_0$, at which a saturation

plateau is reached. For comparison, the error bars for the individual parameters using the SVD modes, optimized for the joint parameters simultaneously, are plotted. These errors are only slightly larger than the ones for the single analysis and the plateau is established for a larger compression numbers. In the case of $\Omega_\mathrm{m}, \sigma_8, \beta$ and z_0, the corresponding MVBs also show a power-law decrease. For Γ, n_s and Ω_Λ, the shape is more complicated. Note that for the joint estimation the individual error bars Δp_α are also obtained from the reciprocal of the corresponding diagonal element $\tilde{F}_{\alpha\alpha}$ and not by inverting the full Fisher matrix, since in both the above described cases, a single parameter is determined from the Fisher matrix.

Similar to the likelihood analysis of cosmic shear surveys (Sect. 3.3), the joint use of ξ_+ and ξ_- yields the smallest error bars on cosmological parameters, followed by ξ_+. However, ξ_- gives better constraints than $\langle M_\mathrm{ap}^2 \rangle$ and thus behaves opposite to the case of the likelihood analysis. This is not a contradiction since the likelihood analysis was performed for two parameters simultaneously in contrast to the KL analysis presented here. As stated in Sect. 3.4.2, the efficacy of the aperture mass dispersion with respect to the 2PCF is improved with increasing number of parameters considered simultaneously. Moreover, in the case of constant covariance, $\langle M_\mathrm{ap}^2 \rangle$ yields better constraints than ξ_-, see Sect. 3.6.5.

The saturation limit is reached for approximately the same mode number for each of the individual cosmological parameters. Irrespective of the survey strategy, a little less than the first half of eigenmodes contain virtually all information about each individual parameter, thus a compression factor of nearly two is possible without increasing the error.

Constant covariance I also conduct a joint parameter estimation for the case of a constant covariance (Sect. 3.5.3). The complete information of all seven cosmological parameters is encoded in the $m = 7$ individual eigenmodes. However, the result of the SVD shows that for $\Omega_\mathrm{m}, \sigma_8, z_0$ and β already the first singular mode carries basically all of the information on these parameters. For Γ, n_s and Ω_Λ, the saturation of the error is reached after two modes. Apparently, the first mode picks up most of the information about the first group of highly-degenerate parameters, the second one completes the information about the parameters from the second group. This picture is consistent with the correlation matrix of the parameters, as will be discussed in Sect. 3.6.3.

3.6.2 Window functions for the compressed eigenmodes

Constant mean In Fig. 3.7 the rank-ordered window functions $W_i(\ell)$ associated with the individual determination of the parameters Ω_m and Γ, respectively, are plotted. As an example, the first three window functions containing most of the information are compared with two higher-order modes which contain less or negligible information. In Fig. 3.8, the first three window functions from the joint parameter analysis are shown, multiplied by the corresponding singular values in order to display their relative contribution of the corresponding mode, since the magnitude of the singular value encodes the information content.

Because the window functions for $\tilde{\xi}_i^+$ are linear combinations of $\mathrm{J}_0(\ell\theta_j)$, there is always an extended tail which takes contributions from very large scales. ξ_+ and consequently its compressed version $\tilde{\xi}^+$ are both sensitive to large-scale power which is not covered by the survey. The window functions for $\langle \tilde{M}_\mathrm{ap}^2 \rangle$ are very localized and each eigenmode samples only a small ℓ-region. For all statistics, small scales are noise dominated and contribute with only small amplitude to higher-order, low signal-to-noise KL modes, see also Fig. 3.11.

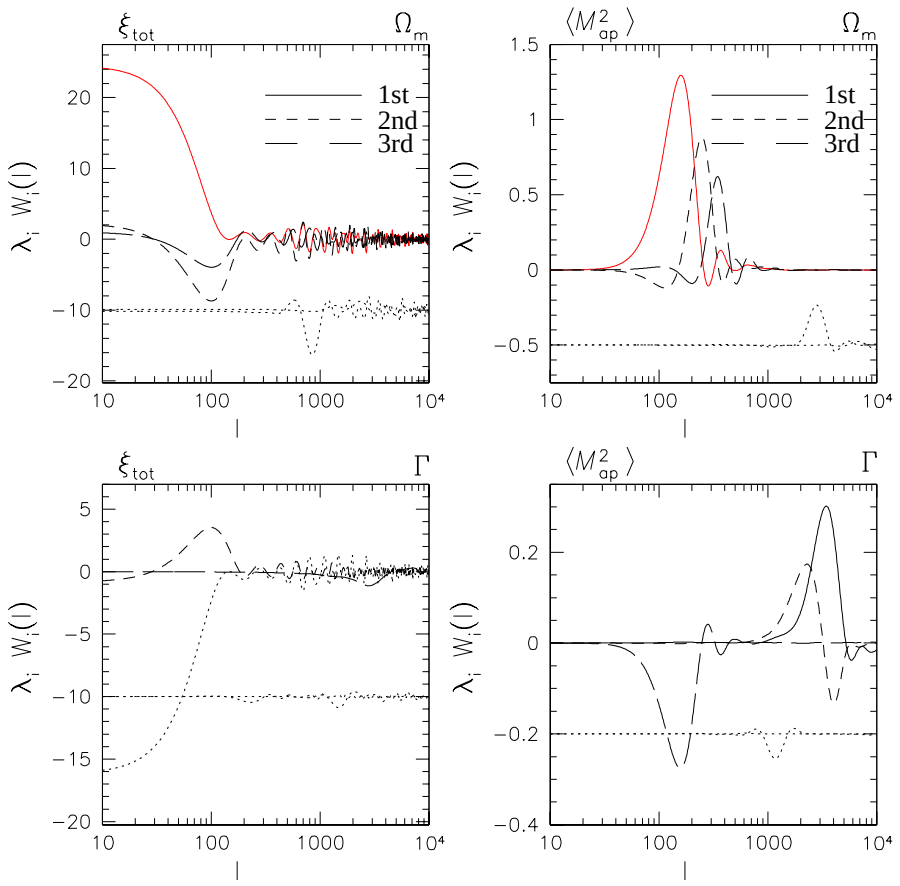

Figure 3.7: *Window functions $W_i(\ell)$ multiplied with the corresponding eigenvalues λ_i for Ω_m and Γ, for the constant mean case. The first three as well as to two (shifted) higher-order eigenmodes ($i = 14, 19$ for ξ_tot and $i = 10, 14$ for $\langle M_\mathrm{ap}^2 \rangle$) are plotted. The survey strategy is a $(30, 100')$ patch geometry (see Sect. 3.1). Note that the window functions for the two higher-order modes have been displaced from zero to increase the visibility.*

The fact that the information content is larger for small ℓ and decreases with increasing ℓ can be understood when considering the simple case in which the covariance matrices and their derivatives are diagonal. Then, the eigenproblem matrix $\mathbf{A} = \mathbf{L}^\mathrm{t}\mathbf{C}_{,\alpha}\mathbf{L}^{-\mathrm{t}}$ (3.13) is also diagonal with $A_{kk} = (C_{,\alpha})_{kk}/C_{kk}$ being the eigenvalues. The diagonals of both $\mathbf{C}$ and $\mathbf{C}_{,\alpha}$ are decreasing functions of the angular scale (except for small bumps due to geometrical effects from the survey). Since the decrease of $\mathbf{C}_{,\alpha}$ is in general shallower than $\mathbf{C}$, the largest eigenvalues occur on the largest angular scales.

Constant covariance For the constant covariance case, the window functions W_0 are much broader since there is only one eigenmode containing all information and taking contributions from all angular scales (Fig. 3.9). Although these are still linear combinations of Bessel functions as described above, the individual contributions blend together to produce a composite window which peaks at a median angular scale determined by the survey geometry and where the signal dominates over the noise both from the intrinsic ellipticity dispersion of galaxies at small angular scales and from the finite sky coverage at larger angular scales.

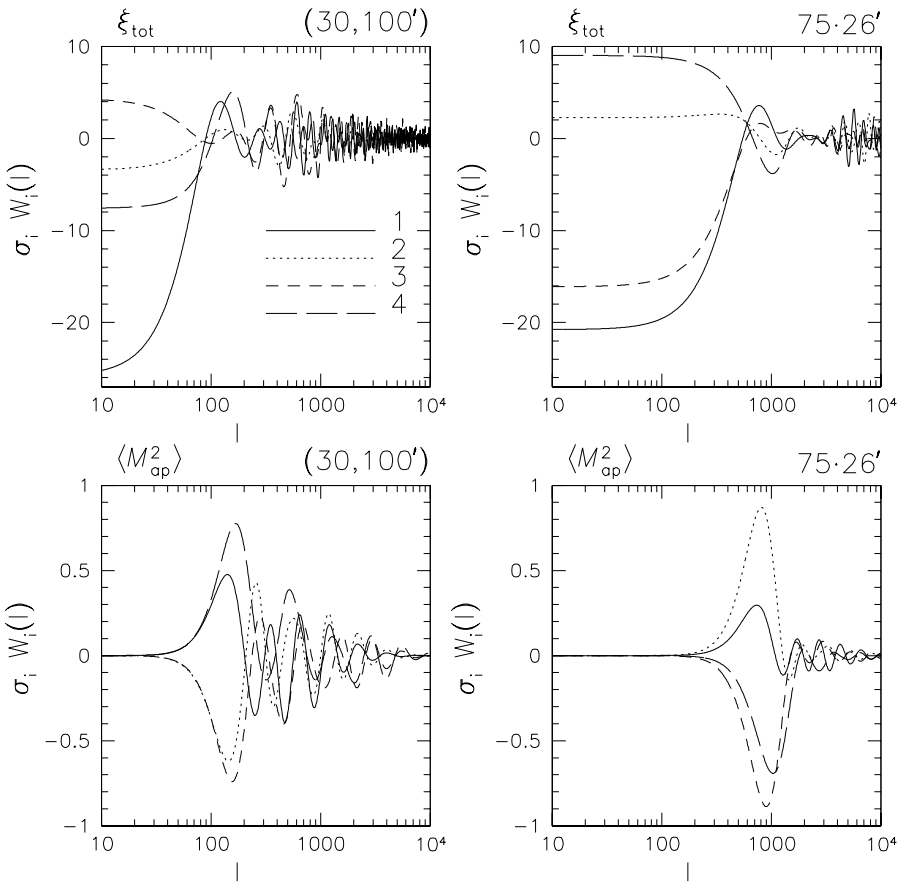

Figure 3.8: *Window functions $\sigma_i W_i$ for the joint analysis of all parameters using SVD in the case of constant mean. The first four singular modes multiplied by the corresponding singular values are displayed. Window functions corresponding to the two surveys $(30, 100')$ and $75 \cdot 26'$ and to the two statistics $\xi_{\rm tot}$ and $\langle M_{\rm ap}^2\rangle$ are contrasted.*

The strong degeneracy among various parameters is reflected in the behavior of their associated window functions: for all four statistics, degenerate parameters have very similar curves (see next section for a discussion of the different near-degeneracies). The Γ-, $n_{\rm s}$ and Ω_Λ-filter functions have a zero-transition – the low ℓ-plateau has opposite sign than the peak. All three parameters influence the slope of the power spectrum – an estimate of a tilt in the power spectrum is obtained by adding up power on large and small scales with opposite sign.

Figure 3.9 displays for the dominant contibution to the Fisher matrix (second term in eq. 3.6) how individual Fourier modes are sampled and combined in an "optimal" way to constrain cosmological parameters as indicated. It shows on which scales the convergence power spectrum has the largest influence on the determination of cosmological parameters, depending on the survey strategy (see also Sect. 3.6.4). It will be interesting in future work to apply these techniques presented here to third-order statistics of cosmic shear, allowing one to quantify the sampling of the convergence bispectrum to extract cosmological information. It will be possible to study the scales of interest for a combined analysis of the power and the bispectrum which will reduce the near-degeneracies between parameters.

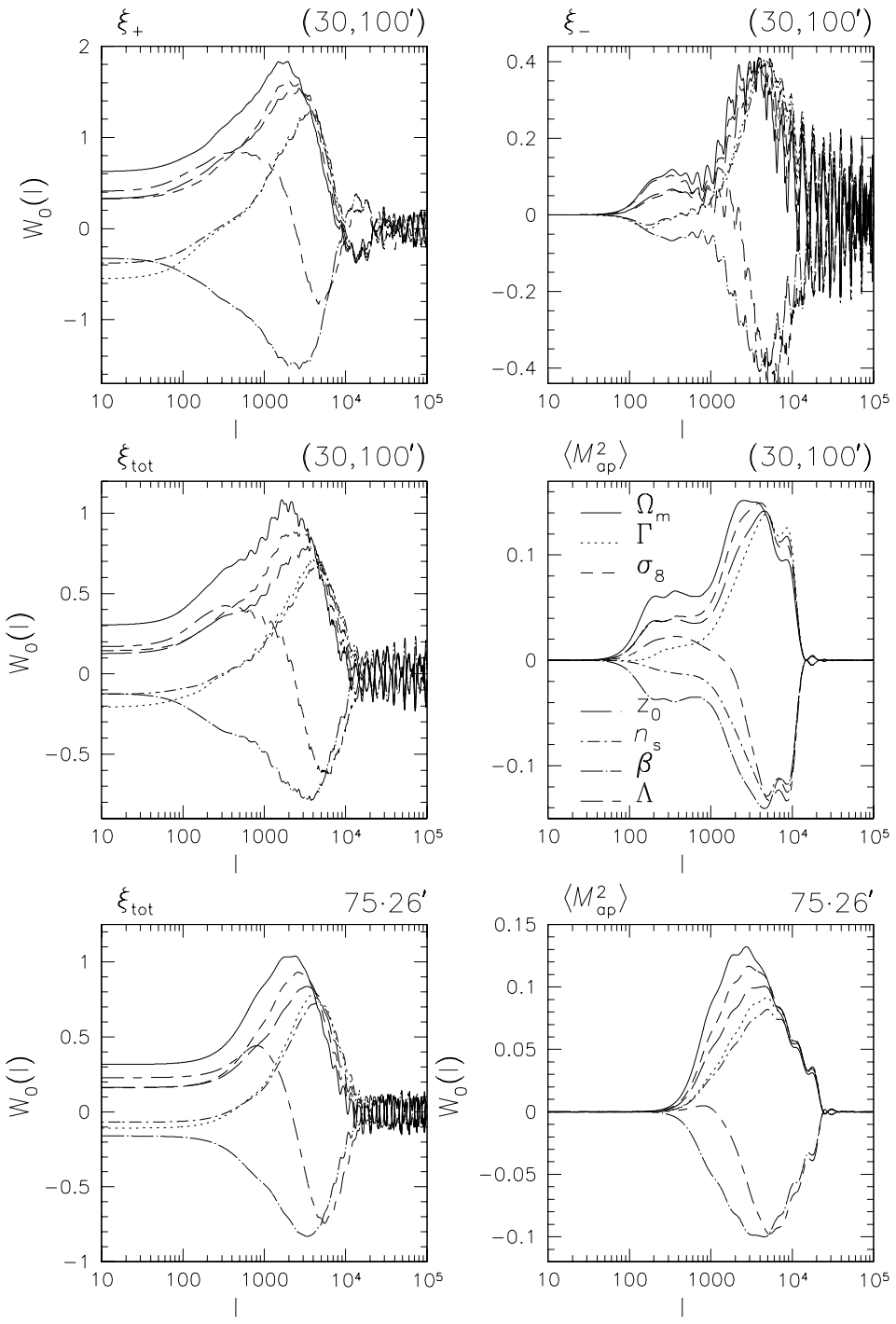

Figure 3.9: *The window function W_0 for the constant covariance case associated with various parameters as indicated in the middle right panel. In the case of the $(30, 100')$ survey, all four shear statistics are shown (upper two rows). For comparison, window functions corresponding to $\xi_{\rm tot}$ and $\langle M_{\rm ap}^2 \rangle$ are plotted for the $75 \cdot 26'$ strategy.*

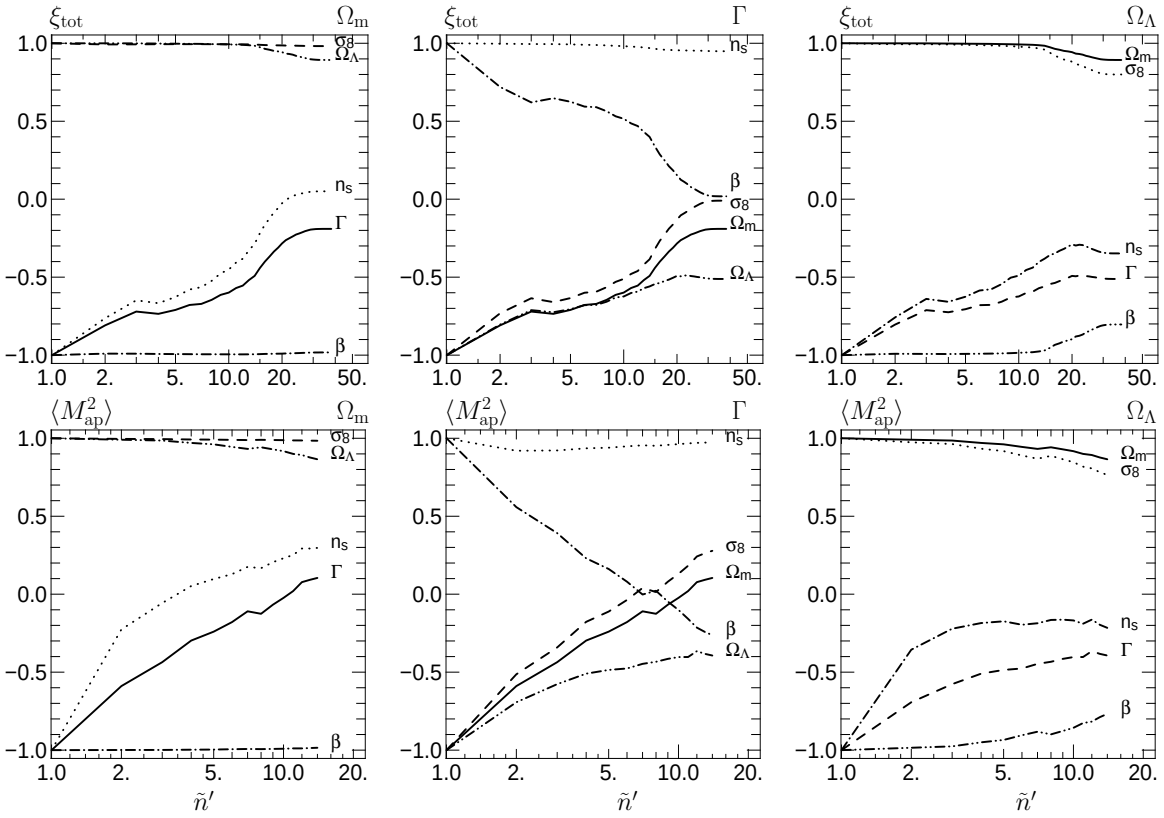

Figure 3.10: *The correlation of $\Omega_{\rm m}$ (left), Γ (middle) and Ω_Λ (right panels) with the other parameters, as a function of the mode number $\tilde{n}'$ for the joint parameter estimation. The upper and lower row corresponds to $\xi_{\rm tot}$ and $\langle M_{\rm ap}^2\rangle$, respectively. The correlation coefficients with z_0 are not plotted since it is virtually identical with the ones for σ_8.*

3.6.3 Correlation between cosmological parameters

Constant mean The KL eigenmode analysis seeks to maximize the diagonal elements of the Fisher matrix and therefore tries to minimize the error corresponding to the MVB (3.5). However, also the off-diagonal elements of the Fisher matrix contain important information — they determine the level of degeneracy between cosmological parameters. The correlation coefficient of the inverse Fisher matrix

$$r_{\alpha\beta} = \frac{F^{-1}_{\alpha\beta}}{\sqrt{F^{-1}_{\alpha\alpha}F^{-1}_{\beta\beta}}} \tag{3.20}$$

is a measure of the correlation between the $\alpha^{\rm th}$ and $\beta^{\rm th}$ parameter. For $\alpha \neq \beta$, it can vary between -1 and 1. In the two-dimensional case, $r_{12} = r_{21} = 0$ corresponds to an error ellipse with major and minor axes parallel to the coordinate (parameter) axes – the probability distribution of the parameters factorizes. For $r_{12} \to 1$, the ellipse degenerates to a line.

In Fig. 3.10 the correlation coefficient of the Fisher matrix as a function of the compression mode number $\tilde{n}$ is plotted. For very small mode numbers, the Fisher matrix becomes singular and all parameters are completely degenerate. The correlation decreases for increasing mode

numbers; however, for some parameter combinations, there is very little decrease, corresponding to a near-degeneracy even if all information is included. One group of highly-degenerate parameters are $\Omega_\mathrm{m}, \sigma_8, z_0$ and β, another near-degenerate pair is Γ and n_s. Between the first and the second group, the correlation in some cases changes sign when going from high to small mode number. The first group roughly influences the convergence power spectrum in modifying its amplitude whereas the second group tilts the power spectrum, see Fig. 1.9. Although the cosmological constant Ω_Λ influences the shape of the power spectrum, it is mainly degenerate with Ω_m, less with σ_8, z_0 and β and not so much with Γ and n_s.

In general, no or a very late plateau is formed for the correlation coefficient. Thus, the off-diagonal elements of the Fisher matrix keep on evolving even after the diagonal elements have reached the saturation limit. Redundant modes which do not carry any information regarding the Fisher diagonal are nevertheless important and help to reduce parameter degeneracies.

Constant covariance I also calculate the correlation coefficient $r_{\alpha\beta}$ for the constant covariance case, where the SVD produces only $m = 7$ singular modes. The correlation coefficient forms a plateau when as few as two of the singular modes are included, as it is the case for the diagonal elements of the Fisher matrix (Sect. 3.6.1). However, even though F does not change much when three modes and more are added, this is not true for the inverse Fisher matrix since F is quite ill-conditioned and small changes have a large impact on F^{-1}. For a numerically stable determination the of MVBs from the inverse Fisher matrix, one needs as many modes as number of parameters.

3.6.4 Survey strategy

Constant mean The overall shape of the constant mean window functions is similar for different survey geometries (Fig. 3.8). When two surveys sample different scales, their respective window functions are shifted accordingly. For example, the low-ℓ-tail of the $\tilde{\xi}^+$-filter is much broader for the surveys consisting of single uncorrelated images than for the patch geometries and drop at much larger ℓ corresponding to smaller scales. Further, for some cosmological parameters the peaks of the $\langle\tilde{M}_\mathrm{ap}^2\rangle$-window functions are shifted towards higher ℓ; the peak positions are plotted as a function of the mode number $\tilde{n}$ in Fig. 3.11, for Ω_m and Γ. For Ω_m (and all other parameters from the same degeneracy-group, see Sect. 3.6.3) this shift is clearly visible, whereas for Γ and n_s the peaks seem to be randomly distributed. For the latter two parameters, there is no preference of large scales, apparently all scales contribute to the Fisher matrix.

Constant covariance The window functions W_0 for the constant covariance also case depend on how different scales are sampled by the survey (see Figs. 3.9). For patch geometries, the filter functions are generally broader towards small ℓ. In the case of $\langle\tilde{M}_\mathrm{ap}^2\rangle$, a low-$\ell$-plateau is formed, which is absent for the uncorrelated images-surveys. A similar effect is also present for the window functions W_0 in the joint parameter estimation case (Fig. 3.12). From this figure, it is also clear that only the first two modes carry significant information about cosmology. The reason is that there are basically two groups of parameters entering the convergence power spectrum, and the informational content can be described by two eigenmodes.

In both cases (constant mean and covariance), the differences between patch strategy and single uncorrelated image survey are present most prominently for low ℓ, because the main difference between these two kinds of settings is the sampling of large scales for the patches.

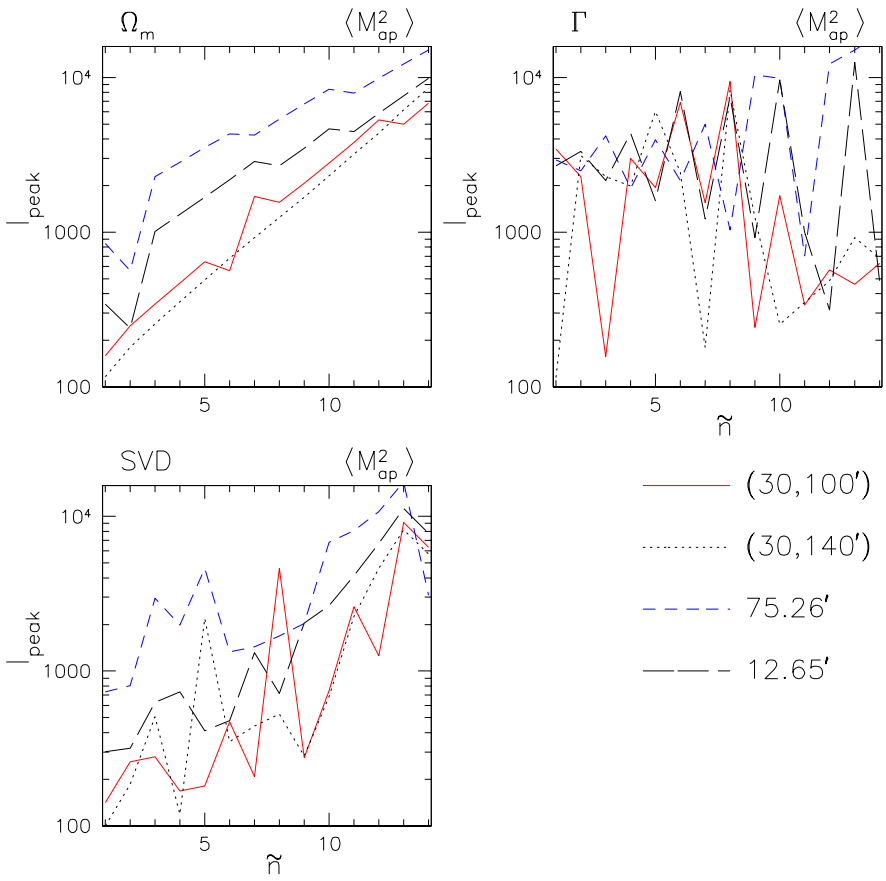

Figure 3.11: *The peak position of the $\tilde{n}^{\rm th}$ $\langle\tilde{M}_{\rm ap}^2\rangle$-windows as a function of the corresponding eigenmode number $\tilde{n}$, for single analysis of the parameters $\Omega_{\rm m}$ (upper left panel) and Γ (upper right), and for the joint estimation (lower panel). $\theta_{\rm peak} = 5/\ell_{\rm peak} = 17'(1000/\ell_{\rm peak})$ is the corresponding peak in real space.*

3.6.5 Noise levels

In this section, the dependence of the Fisher matrix for different survey characteristics and noise levels is quantified. The calculations are done for a $(30, 100')$-patch survey geometry. I vary the survey area A between 1.4 and 14 square degrees, by adding more and more independent, uncorrelated patches to the survey, each patch containing 30 images. For the variation of the other parameters (source ellipticity dispersion σ_ε, number of background galaxies $n_{\rm gal}$, number of bins n and redshift parameter z_0), a single patch was used to calculate the Fisher matrix.

Constant mean If the mean is constant the Fisher matrix is independent of the area of the survey. Except for the binning (see next section), the sensitivity to the survey characteristics is weaker than for the constant covariance case. For the fiducial values of $\sigma_\varepsilon = 0.3, n_{\rm gal} = 30\,{\rm arcmin}^{-2}$, $n = 20$ and $z_0 = 1$, the area where both terms in eq. (3.6) are equal varies between 0.06 (for $n_{\rm s}$) and 0.4 ($\Omega_{\rm m}$) square degrees in the case of $\xi_{\rm tot}$. For the aperture mass statistics, this area of equal contribution of both terms to the Fisher matrix is roughly a factor

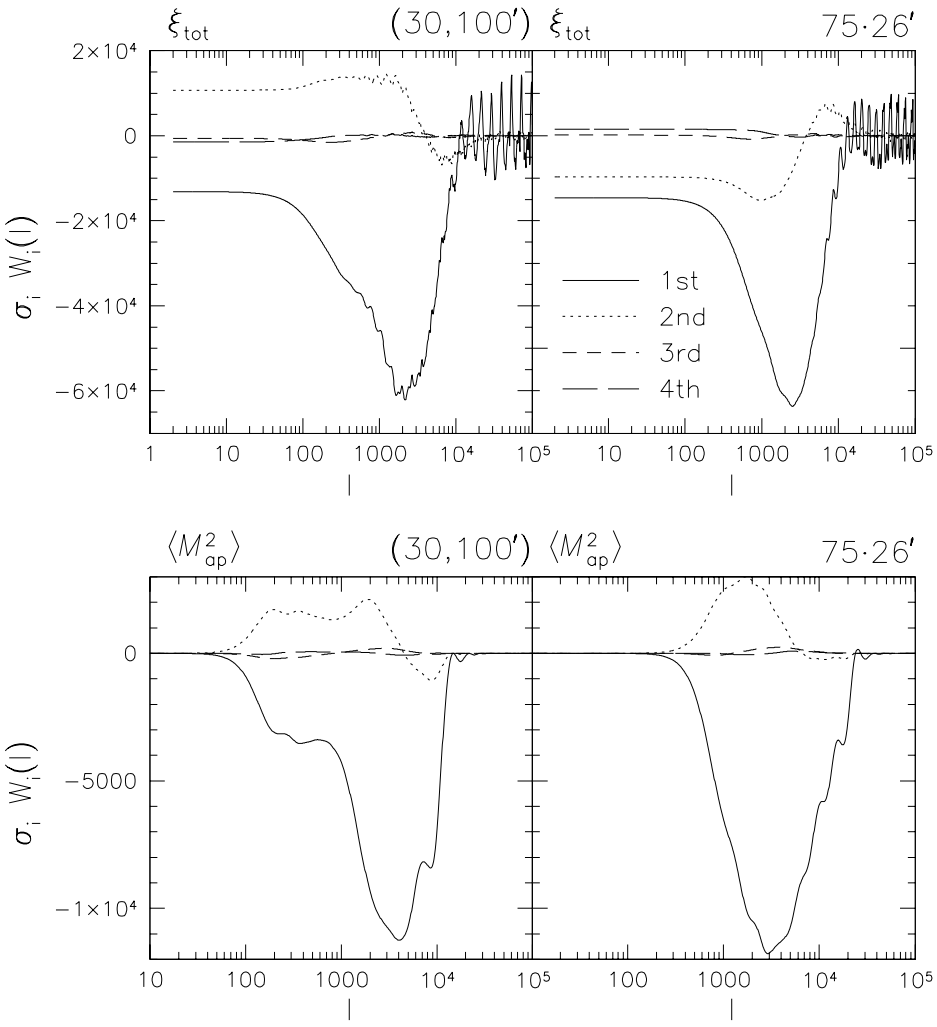

Figure 3.12: *The first four window functions multiplied by the singular values, $\sigma_i W_i$, for the joint analysis of constant covariance, corresponding to two survey strategies as indicated.*

of 4 smaller. Thus, only for very small survey areas, the constant mean term is important.

Constant covariance In the constant covariance case, I write the MVB as

$$\Delta p_\alpha = F_{\alpha\alpha}^{-1/2} = k \left(\frac{A}{10\,\mathrm{sq\,deg}^2}\right)^{-0.5} \left(\frac{\sigma_\varepsilon}{0.3}\right)^{\mu} \left(\frac{n_\mathrm{gal}}{30\,\mathrm{arcmin}^{-2}}\right)^{\nu} \left(\frac{n}{20}\right)^{\eta} z_0^{\lambda}, \qquad (3.21)$$

where the constant k and the power-law indices μ, ν, η and λ are given in Table 3.2 for each parameter α.

Note that although the MVB for $\langle M_\mathrm{ap}^2 \rangle$ is smaller than the one for the 2PCF, this is not true when more than one parameter is considered. In that case, the MVB is given by the inverse of (a submatrix of) the Fisher matrix. The off-diagonal terms are typically larger for $\langle M_\mathrm{ap}^2 \rangle$ which decreases the diagonal of the inverse matrix.

Table 3.2: *The coefficients describing the MVB for the constant covariance case as function of various survey characteristics, for different cosmological parameters, see eq. (3.21).*

	$\xi_{\rm tot}$					$\langle M_{\rm ap}^2 \rangle$				
Param.	k	μ	ν	η	λ	k	μ	ν	η	λ
$\Omega_{\rm m}$	0.0074	−0.77	0.40	−0.03	0.52	0.0073	−0.90	0.51	−0.02	0.70
Γ	0.0109	−1.28	0.71	−0.08	1.03	0.0085	−1.20	0.67	−0.05	1.03
σ_8	0.0098	−0.96	0.52	−0.05	0.85	0.0086	−1.03	0.58	−0.05	1.00
z_0	0.0157	−1.08	0.59	−0.06	−0.59	0.0131	−1.11	0.62	−0.04	−0.49
$n_{\rm s}$	0.0237	−1.34	0.76	−0.09	1.04	0.0184	−1.23	0.68	−0.04	1.00
β	0.0206	−1.09	0.59	−0.06	0.36	0.0172	−1.11	0.62	−0.04	0.45
Ω_Λ	0.0498	−1.25	0.70	−0.09	0.50	0.0424	−1.23	0.68	−0.06	0.24

3.6.6 Binning

The number of angular bins is varied between 20 and 50 in order to quantify the effect of binning on the results. For the constant mean case, the attainable compression factor of almost two and the power-law decrease of the Fisher matrix diagonal are unaffected by the binning. However, and as a consequence of this, the saturation level decreases with increasing bin and therefore total mode number. I find the height of the plateau to weakly depend on the bin number, as $n^{-\eta}$ where η is different for different statistics and parameters, and ranges between 0.18 and 0.3. The plateau is expected to be stationary for very large bin numbers, since it is not possible to increase the cosmological information by refining the binning ad infinitum. The weak dependence on n of the plateau shows that this saturation limit is not yet reached by the number of bins considered here.

The Fisher matrix in the case of constant covariance is only very weakly dependent on the bin number, as shown in the previous section (see Table 3.2). The diagonal elements of the Fisher matrix for both cases (constant mean/covariance) are shown in Fig. 3.13.

The binning has a small influence on the correlation coefficient. Between the two degeneracy groups $(\Omega_{\rm m}, \sigma_8, z_0, \beta$ and $\Gamma, n_{\rm s})$ r gets smaller with increasing bin number; for n varying from 20 to 50, the change is typically smaller than 0.25 when all modes are included. The correlation within a group basically does not change.

3.6.7 Consequences for the likelihood

The KL data compression is optimized in the context of the Fisher information matrix, in the sense that the MVBs are minimized. How will more realistic error bars on parameters behave when obtained from the KL eigenmodes? In this section, the KL studies presented in the previous sections are extended to the likelihood function, and an exemplary case is analyzed. For the compressed data vector $\tilde{\boldsymbol{x}}$ and its covariance, $\tilde{\mathsf{C}}$, the modified figure-of-merit $\tilde{\chi}^2$ (cf. 3.3) is defined as

$$\tilde{\chi}^2(\boldsymbol{p}) = (\tilde{\boldsymbol{x}} - \tilde{\boldsymbol{\mu}}(\boldsymbol{p}))^{\rm t} \tilde{\mathsf{C}}^{-1} (\tilde{\boldsymbol{x}} - \tilde{\boldsymbol{\mu}}(\boldsymbol{p})) \tag{3.22}$$

with $\tilde{\boldsymbol{\mu}} = \mathsf{T}\boldsymbol{\mu}$. For a $(30, 100')$ survey geometry, $\tilde{\chi}^2$ is obtained for $\Omega_{\rm m}$ and σ_8 with all other parameters being fixed, assuming a flat Universe. Confidence levels of $\tilde{\chi}^2$ are plotted

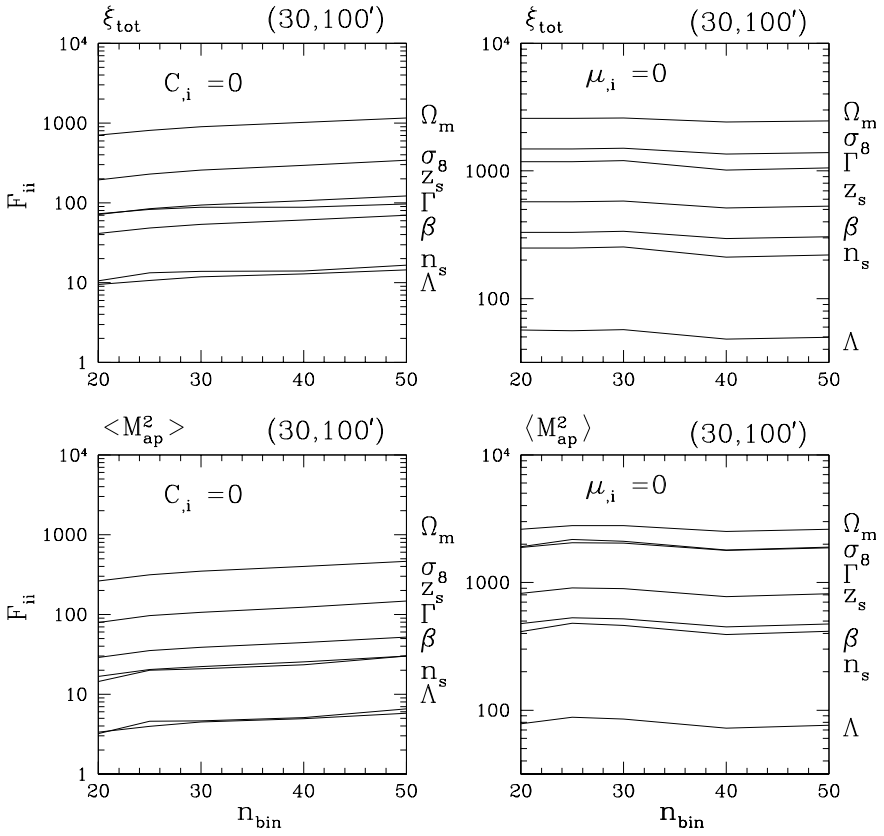

Figure 3.13: *Effect of the bin number n on the Fisher matrix in the constant mean case for various estimators as indicated in the panels.*

in Fig. 3.14 for the uncompressed data vector $\xi_{\rm tot}$ (case A) and five combinations of KL eigenmodes (B - F) as follows.

First, the two constant covariance modes (one for $\Omega_{\rm m}$ and σ_8, respectively) are combined and a SVD is performed (Sect. 3.5.5). Using the first singular mode (B), the degeneracy direction is very badly constrained but orthogonal to that direction the likelihood corresponding to the uncompressed data is very well recovered. The first singular mode completely constrains a combination of $\Omega_{\rm m}$ and σ_8 whose functional dependence determines the degeneracy "valley". The second singular mode (C) adds further information along the near-degeneracy direction. Note that all information about the two parameters are compressed into the two singular modes and therefore the confidence levels (C) are the same than for (A).

Next, the constant mean modes are considered. The 2×38 modes corresponding to the individual optimization for $\Omega_{\rm m}$ and σ_8, respectively, are combined and a SVD is performed as before, yielding 38 singular modes in this case. According to Fig. 3.6, one expects that about half the modes will recover most of the original likelihood. However, the first 20 and 25 modes (D and E) do not constrain $\Omega_{\rm m}$ and σ_8 very well. Only if about 30 modes are taken into account (not shown) most of the original information is sampled by the compressed data.

Finally, the general case is considered and the two constant covariance modes are combined with 30 of the constant mean modes (F). This combination does not yield an improvement

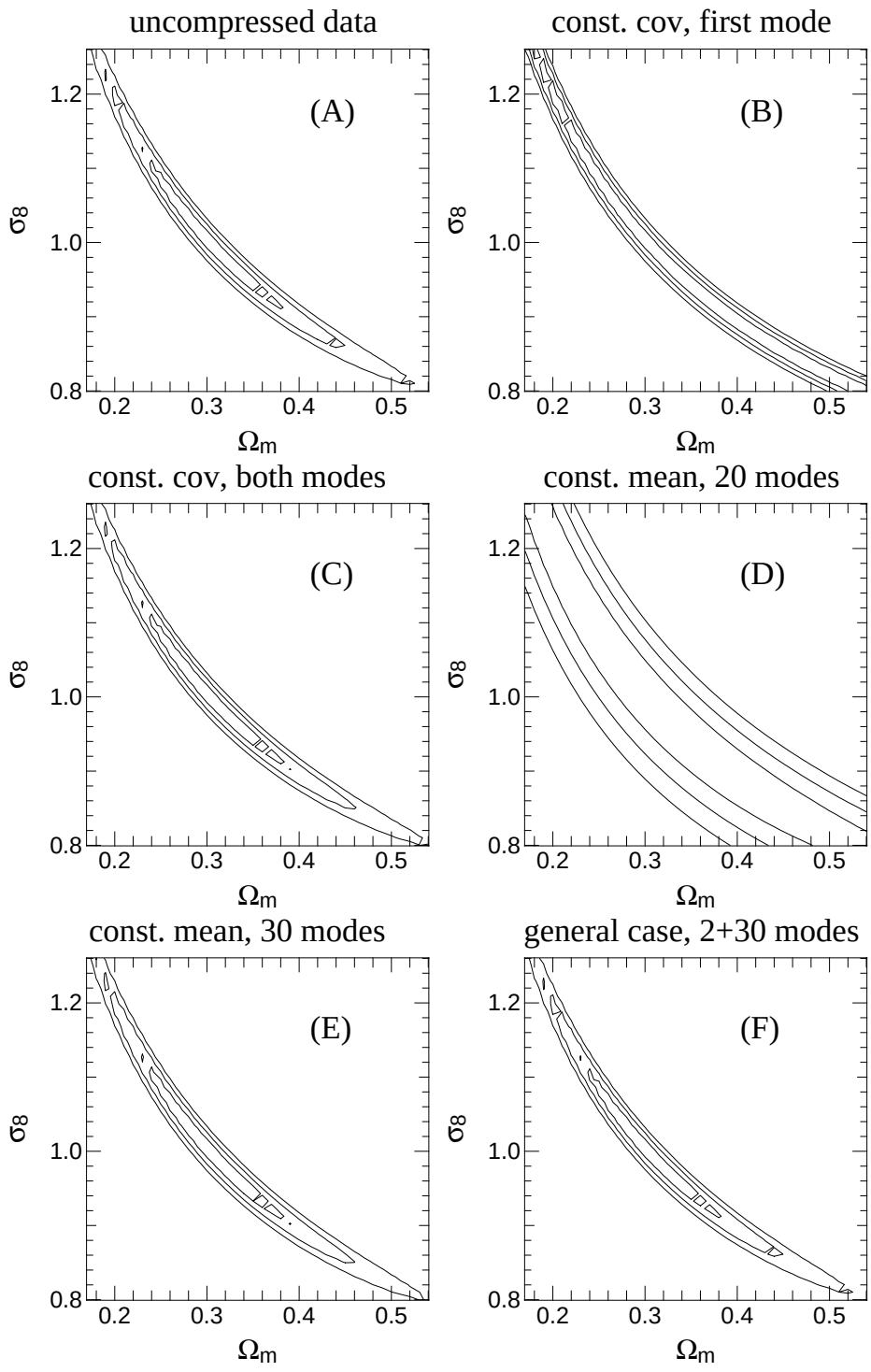

Figure 3.14: *Confidence contours (1, 2 and 3σ) from the figure-or-merit $\tilde{\chi}^2$ (3.22) in the $\Omega_{\rm m}$-σ_8-plane. All other parameters are fixed and a flat Universe is assumed. The six panels correspond to the uncompressed combined data vector $\xi_{\rm tot}$ (A) and to various combinations of KL eigenmodes (B-F) as described in the text.*

with respect to (C). Therefore, the conclusion obtained from the Fisher matrix analysis which states that the constant covariance modes carry negligible information with respect to the constant mean modes is also valid for the more general case of likelihood analysis.

3.7 Summary and conclusions of the survey strategy analysis

I numerically calculate the covariance matrices (2.57 - 2.58) for the second-order estimators of cosmic shear $\xi_\pm$ and $\langle M_{\rm ap}^2 \rangle$, which were derived in Schneider et al. (2002), via a Monte-Carlo integration technique. Galaxy positions are simulated corresponding to various cosmic shear survey geometries of 14 square degree area. These surveys consist of a total of 300 images of size $13' \times 13'$ which are randomly distributed in patches on the sky. A number of (semi-)random patch configurations are compared to a survey consisting of 300 completely uncorrelated images. I summarize the results of the maximum likelihood analyses presented in Kilbinger (2002) and Kilbinger & Schneider (2004), before I perform a detailed analysis using the Fisher information matrix for three and four cosmological parameters out of $(\Omega_{\rm m}, \sigma_8, \Gamma, \Omega_\Lambda, n_s)$. The more parameters are assumed to be determined from the data, the more important becomes large-scale information in order to resolve the near parameter degeneracies. Using the combined ξ_+ and ξ_-, some of the patch geometries yield tighter constraints than the uncorrelated image configuration. The aperture mass is best applied to patches with $N = 30$ images, the results are nearly independent of the patch radius in most cases.

The differences between the individual patch geometries make up to 25 percent for the minimum variance bound on several parameters. Thus, a 25 percent improvement on the determination on cosmological parameters can be obtained solely by choosing an appropriate geometry for a future cosmic shear survey.

For the KL eigenmode analysis, two different scenarios are considered. In one case, the mean of second-order shear estimators is used to constrain the cosmological parameters while the covariance is constant and independent of cosmology. The second case uses the covariance of these estimators to constrain cosmological parameters assuming a constant mean. I study the information content of KL eigenmodes in both cases for various parameters. For the constant mean case, there are several eigenmodes among which the information is distributed; however, there is only one eigenmode associated with the constant covariance case which contains the complete information. The resulting error bars are inversely proportional to the square root of the survey area for the constant covariance case, and independent of the survey area if the mean is constant. Thus, for reasonable sky coverage (more than about 0.4 square degrees) the first case dominates over the second and the bulk part of the cosmological information is collected in only one mode per parameter.

From the results of the KL eigenmode analysis I find that a compression factor almost of two can be achieved in most cases without information loss, if the cosmological parameters are determined using the covariance. KL analysis provides rank-ordered, uncorrelated eigenmodes and corresponding window functions in the case of single parameter estimation. These sets of window functions provide cleanest measures of the projected power spectrum for a specific survey strategy. The first eigenmodes contain most of the information that one can extract and directly use for maximum-likelihood studies to constrain cosmological parameters. Typically, the error bars on cosmological parameters are smallest when the joint correlation function $\xi_{\rm tot}$ is used. ξ_+ gives better constraints than ξ_- which is in turn better than the aperture mass statistics $\langle M_{\rm ap}^2 \rangle$.

I provide both an independent study of different parameters as well as a joint analysis of all parameters for each of the two cases of constant mean and constant covariance. The parameters which are near-degenerate have similar KL eigenfunctions. Among the five cosmological parameters I have considered, (Ω_m, σ_8) and (Γ, n_s) show similar levels of degeneracy and have very similar eigenmodes. Moreover, the parameters characterizing the source galaxy redshift distribution (z_0 and β, see eq. 1.90) are degenerate to a high level with Ω_m and σ_8.

The eigenmodes associated with Ω_m and σ_8 mainly focus on larger angular scales where cosmological information is not contaminated by noise on small scales. On the other hand, Γ and n_s measure the shape of the projected matter power spectrum which is easier to determine when information on smaller angular scales is also available. Therefore the eigenmode windows associated with these two parameters tend to take more contributions from smaller scales.

In contrast, the eigenmodes associated with the constant covariance case take contributions from virtually all angular scales probed by the survey and reach a maximum roughly at medium angular scales where the signal dominates over both shot noise and scatter due to finite sky coverage. The window functions associated with the combined analysis of (ξ_+, ξ_-) are dominated by contributions from ξ_+, whereas the $\langle M_\mathrm{ap}^2 \rangle$-filter, although more localized, is not very different from that of ξ_-.

Higher-order KL modes do not carry any useful information and approximately half of the modes contain all of the information content of the Fisher matrix for a constant mean. However, the off-diagonal terms of the Fisher matrix which encode the cross-correlation among various parameters are also a function of the number of modes that are included to reconstruct the Fisher matrix. I find that these terms keep evolving even after the diagonal terms have already reached a saturation limit. For very small numbers of modes the reconstructed Fisher matrix becomes singular, since the information sampled in only the few eigenmodes is too little in order to put constraints on more than one parameter, thus different cosmological parameters become completely degenerate.

The joint analysis of all seven parameters being estimated from the mean only (constant covariance) shows that only two eigenmodes (out of seven) are needed to constrain the parameters and to lift their near-degeneracies as far as possible. This is because the parameters can be arranged into two groups according to their degeneracy between each other. One group consisting of $\Omega_\mathrm{m}, \sigma_8, z_0$ and, to a smaller extend, Ω_Λ, the other group being composed of Γ and n_s. Two independent modes are sufficient to comprise most of the information about the two degeneracy groups.

The KL eigenmodes are also used in various combinations for an exemplary case of likelihood analysis involving Ω_m and σ_8. In the case of constant covariance, the degeneracy direction between these two parameters is encoded in the first of two eigenmodes. The constant mean modes do not contribute significantly to the likelihood in comparison to the constant covariance modes as in the case of the Fisher matrix analysis.

In the analysis involving KL eigenmodes I have not included and modeled any systematic measurement errors. I have assumed that the errors are dominated by Poisson noise and intrinsic ellipticity noise at small scales and cosmic variance at larger angular scales. The covariance of the second-order shear statistics include all these noise sources, which are exact if the shear is a Gaussian field. Non-Gaussianity leads to an under-estimation of the noise on angular scales between 1 and 10 arc minutes (see Sect. 2.5.1).

Chapter 4

Third-order shear statistics from theoretical predictions and ray-tracing simulations

In this chapter two different third-order statistics of cosmic shear, the three-point correlation function $\Gamma^{(q)}$, $q = 0, 1, 2, 3$ and the generalized third-order aperture mass $\langle M_{\rm ap}^3 \rangle$ are determined from theoretical predictions and numerical ray-tracing simulations separately, the latter based on CDM N-body simulations. A comparison is undertaken in order to assess the accuracy of non-linear models of the bispectrum in the framework of cosmic shear. I use results from quasi-linear perturbation theory (PT, Sect. 1.2.4) and hyper-extended perturbation theory (HEPT, Sect. 1.2.7.2) in order to model the convergence bispectrum (1.92 - 1.94). PT describes the large-scale structure accurately only on very large scales in the linear regime. Since the HEPT fitting formulae have been obtained using N-body simulations, they are expected to match the ray-tracing simulations reasonably well. Apart from the comparison between theory and simulations, some symmetry properties established in Sect. 2.3.1.4 of the 3PCF are verified numerically and predictions regarding the amplitude and the shape of the 3PCF as a function of various triangle parameters are investigated. Furthermore, the third-order aperture mass statistics and its dependence on cosmological parameters is examined.

4.1 The three-point correlation function

Since it is not possible to display the 3PCF as a function of all its three arguments characterizing a triangle, the following analysis is carried out by considering a special set of triangles for which the 3PCF is plotted as a function of only one varying parameter. For the theoretical model, the natural components of the 3PCF (2.24) in the ortho-center projection are calculated from the bispectrum using (2.33) and (2.34). From the OCDM ray-tracing simulations (model 3, Table B.1), the eight parity-modes of the 3PCF (2.23) are obtained by averaging over triangles with the same configuration (within some bin). The natural components are obtained from the parity-modes and vice-versa using (2.24) and (2.25).

4.1.1 The 3PCF from the ray-tracing simulations

In order to calculate the parity-modes of the 3PCF from the ray-tracing simulations, a brute-force approach is adopted. For every triple of data points, where each data point is a representation of the two Cartesian shear components γ_1 and γ_2 given at an angular position on a grid, the eight parity-modes are calculated, and the results are stored in a bin corresponding to the triangle formed by the three positions. This is a very time-consuming method: For N being the number of data points, the computing time of this algorithm is proportional to N^3. In contrast, a tree code can reduce the computation cost to $\mathcal{O}(N \log N)$ (Jarvis et al. 2004)

or even $\mathcal{O}(N)$ (Zhang & Pen 2005). However, the brute-force method is chosen for several reasons. First of all, in this section only a small number of special triangles (e.g. isosceles triangles with two sides measuring 1 arc minute) is considered for which the 3PCF is compared with theoretical predictions. Therefore, the number of N^3 combinations for a general triangle is vastly reduced and so is the computation time. Secondly, the ray-tracing data is given on a grid in which case the advantage of a tree-code over the brute-force approach is less pronounced. Only for real data, where the galaxies are clustered and the observed field is blanked by gaps and masked regions because of bright stars or telescope reflections, a tree-code algorithm shows its full superiority. Moreover, the brute-force code makes extensive use of the fact that the shear positions are given on a regular grid: The distances between points are calculated using integer operations and a particular triangle of points is translated over the grid in both directions without the need of recalculating the angles for the projection of the 3PCF onto the triangle center. Thus for each triangle position on the grid, the 3PCF can then be obtained by a small number of multiplications and additions of the shear at the corresponding triangle points and the pre-calculated angles.

I obtain the 3PCF and its variance using the mean and the rms over seven OCDM ray-tracing simulations, each field with an area of 10.5 square degrees. Since the individual fields represent virtually independent realizations of the simulated large-scale structure, the variance corresponds to the sampling or cosmic variance. Note that the error bars in Figs. 4.4 - 4.6 are not independent but highly correlated.

For triangles, with side lengths x_1 and x_2 larger than about 200 times the pixel size, the computation time is quite long after all. In order to reduce the calculation time for these cases, the triangles are not shifted to every grid point. Instead random numbers uniformly distributed between 1 and $r_{\rm step}$ are chosen for the steps between successive shifts. The number $r_{\rm step}$ is typically of the order of x_1 and x_2 in units of the grid width. This sparse sampling of the shear data on the grid points will not introduce any bias to the mean of the 3PCF, nor to the variance, since cosmic variance is independent of the number density of sample points (for real data this is the number density of background galaxies $n_{\rm gal}$). Note however, that if the number of sampling points is too small, this estimate of the cosmic variance can get very noisy.

4.1.2 The 3PCF from theoretical non-linear models

On numerically integrating the convergence bispectrum $\bar{b}_\kappa$, the natural components of the 3PCF are obtained using eqs. (2.33) and (2.34) (for the latter equation, cyclic permutations of the indices yield $\Gamma^{(2)}$ and $\Gamma^{(3)}$). Since the comparison is made with simulations for only one or two cosmological models, the bispectrum for these models is pre-computed and tabulated values are stored, resulting in the number of integrations to reduce from four to three.

The three-dimensional integral is calculated numerically using a Gaussian quadrature routine. The upper integration limits are set to $\ell_{\rm lim} = 5 \cdot 10^5$ and the integrand is evaluated at the roots of Gauss-Legendre polynomials for the two ℓ-integrations and equidistantly for φ. The number of sampling points chosen for the ℓ-integrations is 300 and for the angular φ-integration is 75, representing a compromise between accuracy and computation time.

On increasing the upper limit $\ell_{\rm lim}$, the number of sampling points and the sampling rate of the bispectrum, the results of the numerical integration are changed by small amounts, see Fig. 4.1. The "fluctuations" between different sets of the above integration parameters are of the order of up to about 10^{-8}, and approximately represent the absolute precision

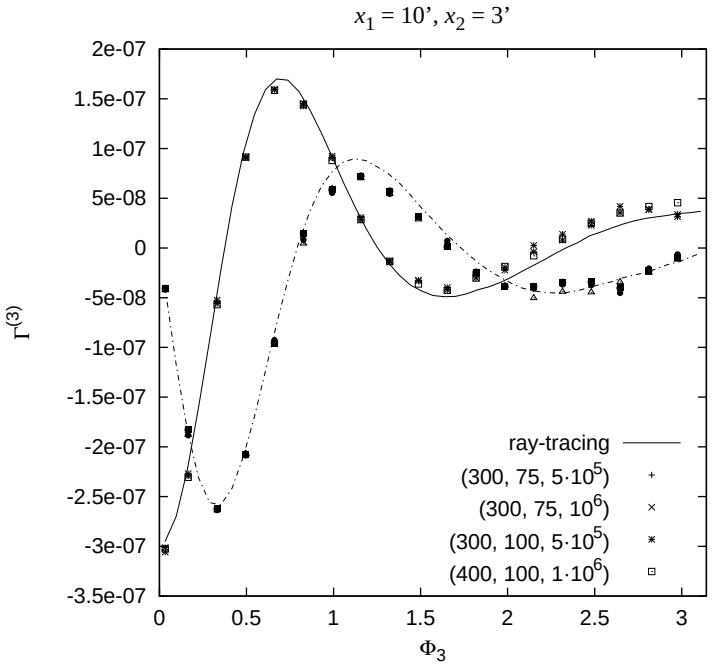

Figure 4.1: *$\Gamma^{(3)}(x_1, x_2, \phi_3)$ (solid line: real part, dashed: imaginary part) for triangles with $x_1 = 10', x_2 = 3'$, with different numerical integration parameters. The three numbers in brackets denote the values for the number of sampling points for the ℓ-integrations, for the φ-integration, and the cutoff for the ℓ-integration, respectively.*

of the numerical integration. For the typical magnitude of the 3PCF in the angular range between 1 and 20 arc minutes, this corresponds to a relative precision of 1 to 10 percent. This choice of numerical parameters is sufficient for carrying out a qualitative comparison between the theoretical prediction of HEPT and ray-tracing simulations. Note also that the HEPT bispectrum of the density contrast fits N-body simulations with a precision of only up to 15%.

4.1.3 Comparison between theory and simulations

Figs. 4.3 to 4.6 show the 3PCF for triangles with fixed x_1 and x_2 and the angle between these two sides ϕ_3 in the range $[0; \pi]$. The 3PCF for $\pi \leq \phi_3 \leq 2\pi$ can be obtained by performing a mirror transformation with respect to the side x_2 on a triangle given by (x_1, x_2, ϕ_3) which will result in the triangle $(x_1, x_2, 2\pi - \phi_3)$, see Fig. 4.2. This is a parity transformation, and thus

$$\begin{aligned}
\gamma_{\mu\nu\lambda}(x_1, x_2, 2\pi - \phi_3) &= \quad \gamma_{\mu\nu\lambda}(x_1, x_2, \phi_3) && \text{(parity-even)}, \\
\gamma_{\mu\nu\lambda}(x_1, x_2, 2\pi - \phi_3) &= -\gamma_{\mu\nu\lambda}(x_1, x_2, \phi_3) && \text{(parity-odd)}, \\
\Gamma^{(q)}(x_1, x_2, 2\pi - \phi_3) &= \quad \left(\Gamma^{(q)}\right)^*(x_1, x_2, \phi_3) && \text{for } \; q = 0 \ldots 3.
\end{aligned} \tag{4.1}$$

The theoretical model reproduces the results from the ray-tracing simulations very well for most triangle configurations and on all angular scales under consideration. Only for very degenerate triangles, where one of the inner angles is close to zero, there is a mismatch between theory and simulations. This is due to the inacurracy of the theoretical model on very small scales and to discreteness effects of the ray-tracing fields, which are most pronounced for nearly collapsed triangles for which one of the side lengths is close to the pixel size and reliable estimates of the 3PCF cannot be made.

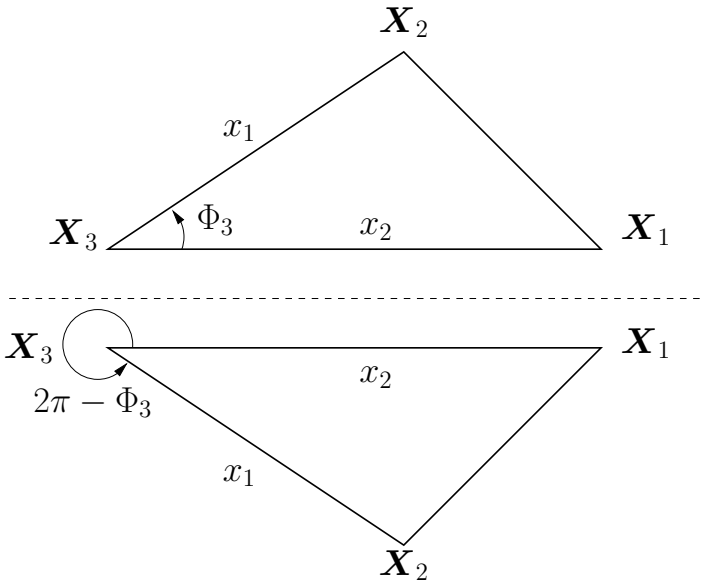

Figure 4.2: *Mirror transformation of a triangle with respect to the side x_2. The cyclic triangle is reflected into an anti-cyclic one, the angle ϕ_3 changes into $2\pi - \phi_3$.*

4.1.4 Verification of the symmetry properties of the 3PCF

The most important conclusion about the 3PCF that can be drawn from Figs. 4.5 and 4.6 is that for a general triangle without special symmetry properties, none of the four complex components (or equivalently the eight real parity-components) of the 3PCF or their linear combinations vanish, and all components have about the same amplitude range. The parity-odd modes are not subdominant as might be suggested by only looking at isosceles triangles (Figs. 4.3 and 4.4), but have the same order of magnitude than the parity-even modes. This is in contrast to the second-order case, where the two quantities $\langle\gamma_\mathrm{t}\gamma_\times\rangle$ and $\langle\gamma_\times\gamma_\mathrm{t}\rangle$ vanish identically because of parity symmetry. Consequently, in future weak lensing surveys, all eight independent components of the 3PCF have to be measured in order not to discard any cosmological information.

The symmetry properties of the 3PCF are verified using both the theoretical model and the ray-tracing simulations. For isosceles triangles with $x_1 = x_2$, the two parity-even modes $\gamma_{\mathrm{t}\times\times}$ and $\gamma_{\times\mathrm{t}\times}$ are found to be equal and the parity-odd components vanishing $\gamma_{\mathrm{tt}\times} = \gamma_{\times\times\times} = \gamma_{\mathrm{t}\times\mathrm{t}} + \gamma_{\times\mathrm{tt}} = 0$. For the complex natural components, $\Gamma^{(0)}$ and $\Gamma^{(3)}$ are equal and real, and $\Im\Gamma^{(1)} = -\Im\Gamma^{(2)}$, as was already stated in Sect. 2.3.1.4.

4.1.5 Other properties of the 3PCF

The amplitude of the 3PCF components is typically of the order 10^{-7} to 10^{-6} for the angular range between $1'$ and $10'$. This is about two orders of magnitudes below the 2PCF signal. Even for small angular scales, where non-Gaussian contributions are important, the skewness of cosmic shear is much smaller than its variance. This makes the measurement of third-order shear functions very challenging. The 3PCF signal decreases with increasing triangle size because the correlation of the large-scale mass distribution decreases with separation and thus the shear signal diminishes, too. The 3PCF shows cyclic patterns which reflect the periodicity of the triangle when ϕ_3 is varied between 0 and 2π. On top of that, some of the components show higher-frequency sinusoidal oscillations.

The equilateral triangle with its maximum symmetry is a special case for the 3PCF as

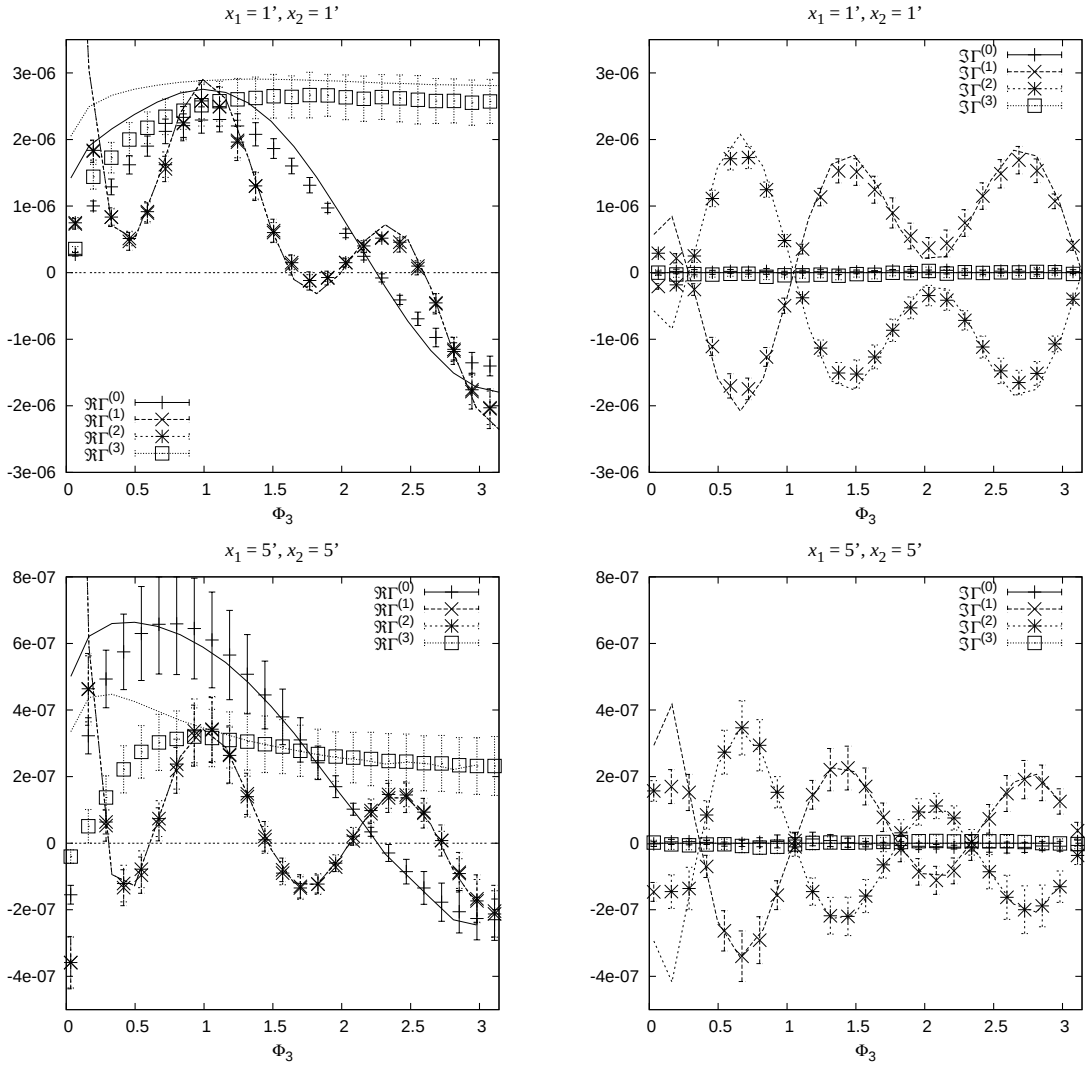

Figure 4.3: *The real (left panels) and imaginary (right panels) part of the four natural components of the 3PCF $\Gamma^{(q)}(x_1, x_2, \phi_3), q = 0 \ldots 3$ (2.24) for iscosceles triangles with $x_1 = x_2 = 1'$ (top) and $x_1 = x_2 = 5'$ (bottom), as a function of the angle ϕ_3 between x_1 and x_2. The lines correspond to the theoretical HEPT predictions, the points with error bars are obtained from the ray-tracing simulations where the errors represent the rms over seven realizations.*

can be seen in the plots of isosceles triangles (Figs. 4.3 and 4.4). Some of the parity-even modes have local extrema and the parity-odd modes have zero-crossings for $\phi_3 = \pi/3$. This behavior shows up for all angular scales between 1 and 20 arc minutes and thus seems to be related to the symmetry of equilateral triangle. In contrast, the zero-crossings of the parity-mode for $\pi/2 < \phi_3 < \pi$ are only present for scales larger than about 2 arc minutes. These features therefore are not caused by the triangle's symmetry but are based on properties of the convergence bispectrum and its scale dependence.

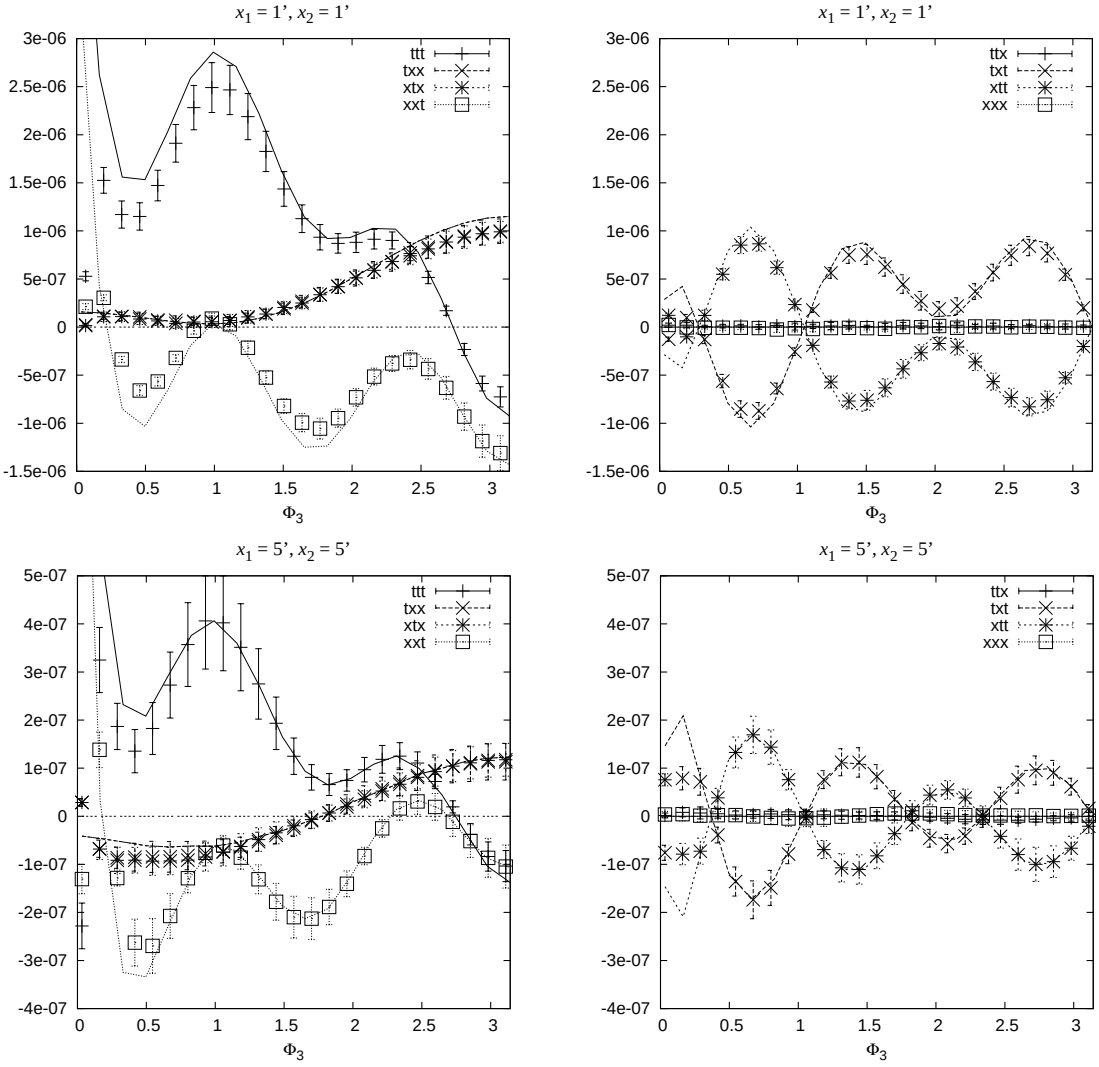

Figure 4.4: *The parity-even (left panels) and parity-odd (right panels)components of the 3PCF (2.23) for iscosceles triangles with $x_1 = x_2 = 1'$ (top) and $x_1 = x_2 = 5'$ (bottom), as a function of the angle ϕ_3 between x_1 and x_2. The lines correspond to the theoretical HEPT predictions, the points with error bars are obtained from the ray-tracing simulations where the errors represent the rms over seven realizations.*

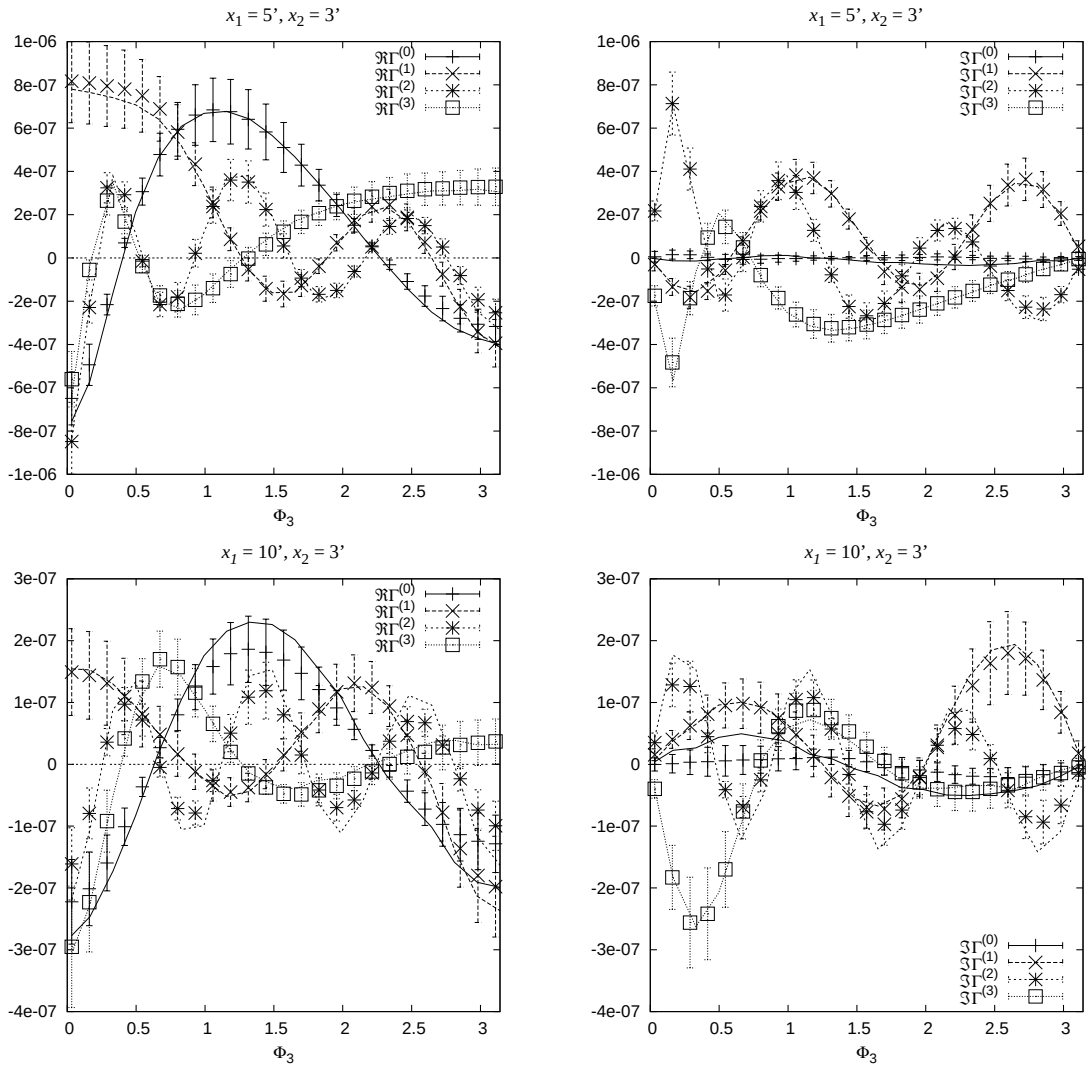

Figure 4.5: *The real (left panels) and imaginary (right panels) part of the four complex components of the 3PCF $\Gamma^{(q)}$ (2.24) for triangles with $x_1 = 5', x_2 = 3'$ (top) and $x_1 = 10', x_2 = 3'$ (bottom), as a function of the angle ϕ_3 between x_1 and x_2. The lines correspond to the theoretical HEPT predictions, the points with error bars are obtained from the ray-tracing simulations where the errors represent the rms over seven realizations.*

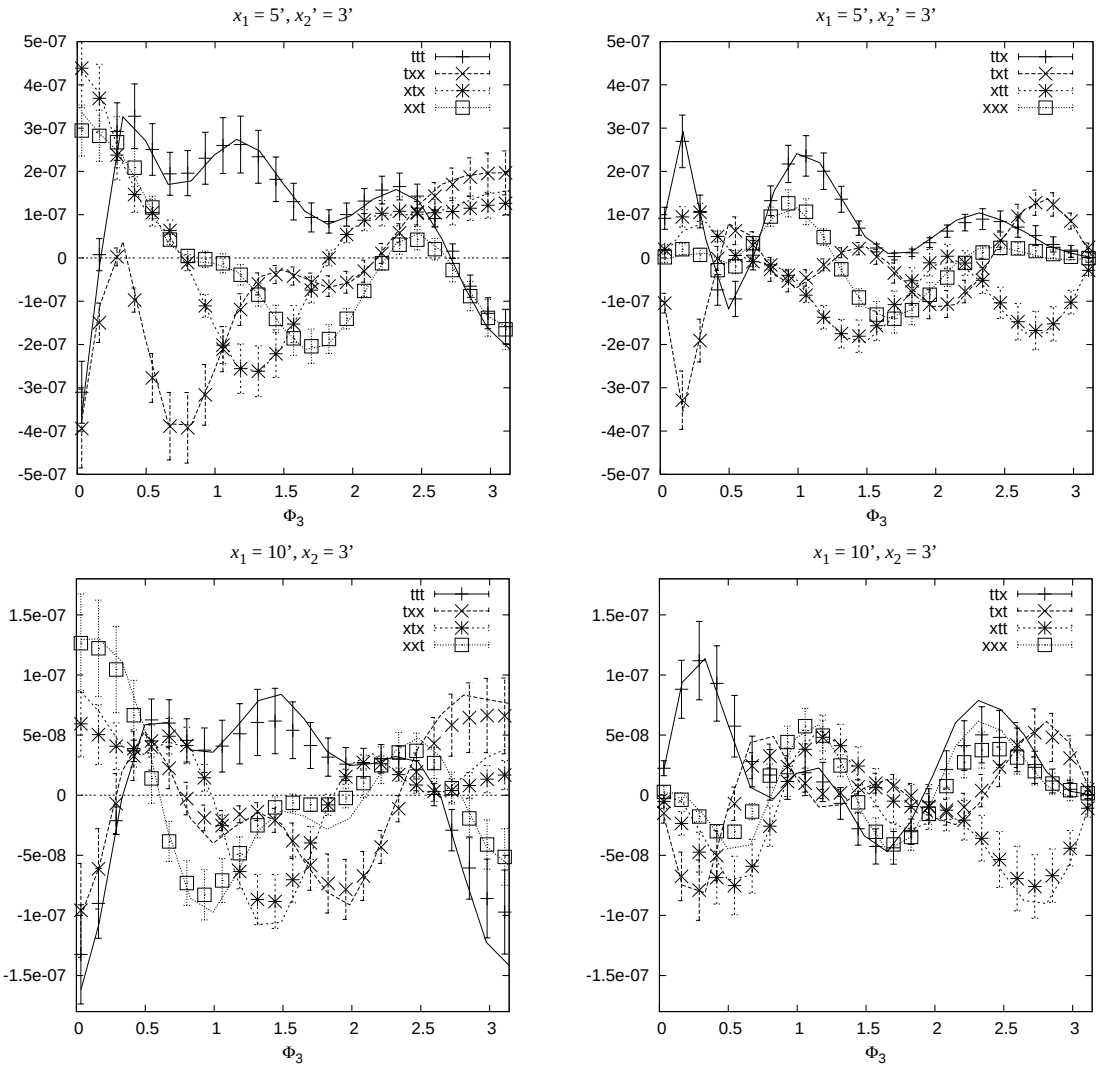

Figure 4.6: *The parity-even (left panels) and parity-odd (right panels) components of the 3PCF (2.23) for triangles with $x_1 = 5', x_2 = 3'$ (top) and $x_1 = 10', x_2 = 3'$ (bottom), as a function of the angle ϕ_3 between x_1 and x_2. The lines correspond to the theoretical HEPT predictions, the points with error bars are obtained from the ray-tracing simulations where the errors represent the rms over seven realizations.*

4.2 Aperture mass statistics

Figs. 4.7 and 4.8 show $\langle M_{\rm ap}^2 \rangle$ and $\langle M_{\rm ap}^3 \rangle$ from the ΛCDM simulations (model 2, Table B.1) and the theoretical predictions using both quasi-linear perturbation theory and non-linear models. The latter reproduce the results from the simulations reasonably well for angular scales above ~ 1 arc minute, whereas PT largely underestimates the shear signal for scales smaller than about 30 arc minutes. In obtaining the aperture mass, the Gaussian filter with infinite support as described by eq. (2.10) is used, therefore the largest aperture which can be put onto the field without being too close to the border is for $\theta_{\rm max} = a/6 = 34'$, where $a = 204.8'$ is the field size.

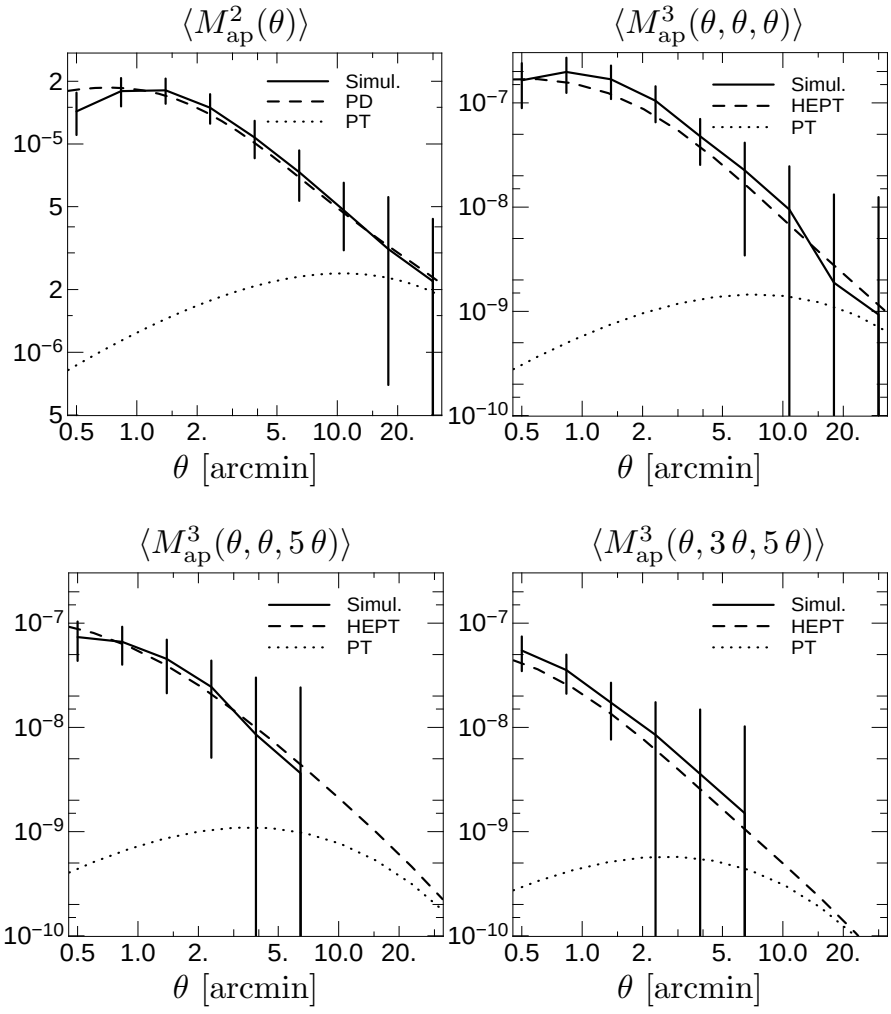

Figure 4.7: *$\langle M_{\rm ap}^2 \rangle$ and $\langle M_{\rm ap}^3 \rangle$ from the 36 ΛCDM simulations (solid lines) as compared to theoretical predictions (dashed and dotted lines). The error bars are the rms values from the 36 fields. The curves for $\langle M_{\rm ap}^3 \rangle$ are calculated from the simulations for aperture radii smaller than one sixth of the field size. PD = Peacock & Dodds (1996) and HEPT = Scoccimarro & Couchman (2001) denote non-linear models of the power spectrum and the bispectrum, respectively (Sect. 1.2.7). PT = (quasi-)linear perturbation theory (Sect. 1.2.4).*

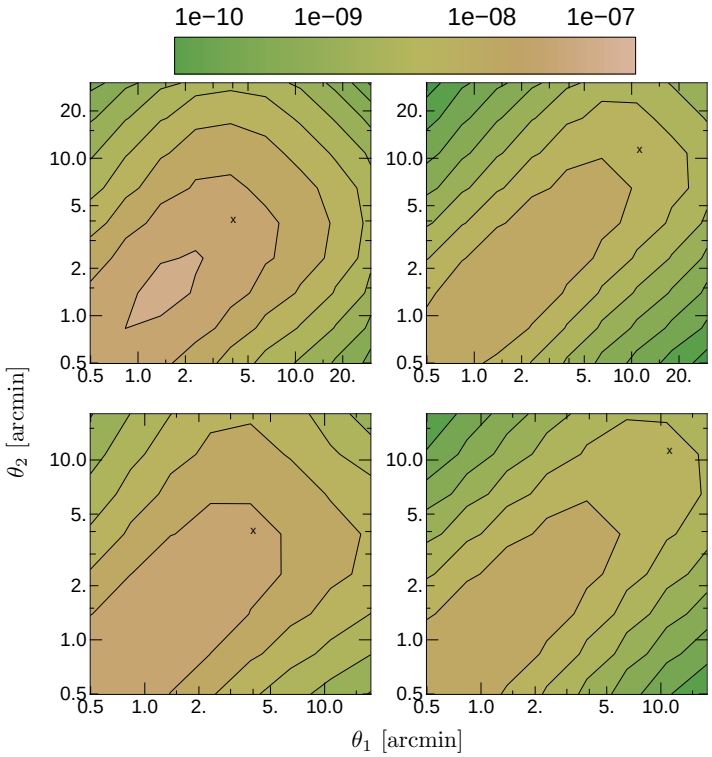

Figure 4.8: *Contours of* $\langle M_{\rm ap}^3(\theta_1,\theta_2,\theta_3)\rangle$ *from simulations (upper row) and from the HEPT model (lower row). In each panel* θ_3 *is fixed to the value indicated by the cross, corresponding to 3.87 arc minutes (left column) and 10.77 arc minutes (right column), respectively.*

A major part of this thesis (Chapter 5) is dedicated to the study of weak lensing aperture mass measurements in order to constrain cosmological parameters. It is therefore instructive to show the dependence of the aperture mass on various cosmological parameters, and to compare its second- and third-order moments. The more different the dependencies are for the second- and third-order statistics, the better will be the improvement on the parameter constraints on combining both statistics. In Figs. 4.9 and 4.10 the logarithmic derivatives of the aperture mass statistics with respect to cosmological parameters used in this work (see also App. B) are shown as calculated in HEPT.

For all parameters shown in Fig. 4.9, the curves are quite featureless and their similarity is due to the near-degeneracies between the parameters. For example, one sees that the ratios $(\partial\langle M_{\rm ap}^2\rangle/\partial n_{\rm s})/(\partial\langle M_{\rm ap}^2\rangle/\partial\Gamma) \approx (\partial\langle M_{\rm ap,d}^3\rangle/\partial n_{\rm s})/(\partial\langle M_{\rm ap,d}^3\rangle/\partial\Gamma)$ are roughly equal and constant as a function of the aperture radius θ. Therefore, the two parameters Γ and $n_{\rm s}$ are expected to have the same near-degeneracy for both statistics.

The ratio of derivatives with respect to $\Omega_{\rm m}$ and σ_8 are slowly increasing functions of θ, with significant differences between $\langle M_{\rm ap}^2\rangle$ and $\langle M_{\rm ap,d}^3\rangle$. From that one can infer that the reduced skewness $s_3 = \langle M_{\rm ap,d}^3\rangle/\langle M_{\rm ap}^2\rangle^2$ breaks the $\Omega_{\rm m}$ - σ_8 degeneracy of second-order cosmic shear statistics. From Fig. 4.9 one sees that $\partial\ln\langle M_{\rm ap,d}^3\rangle/\partial\sigma_8 \approx 2\partial\ln\langle M_{\rm ap}^2\rangle/\partial\sigma_8$, so $\partial s_3/\partial\sigma_8 \approx 0$ — s_3 is indeed nearly independent of σ_8, as predicted from quasi-linear perturbation theory (Bernardeau et al. 1997; Schneider et al. 1998).

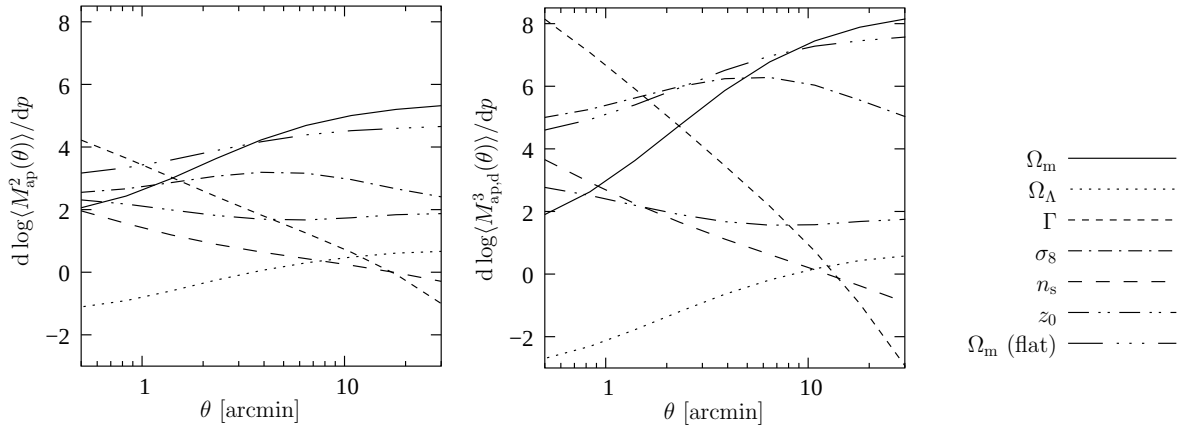

Figure 4.9: *Logarithmic derivatives of* $\langle M_{\rm ap}^2 \rangle$ *and* $\langle M_{\rm ap,d}^3 \rangle$ *with respect to some cosmological parameters as indicated in the figure legend.*

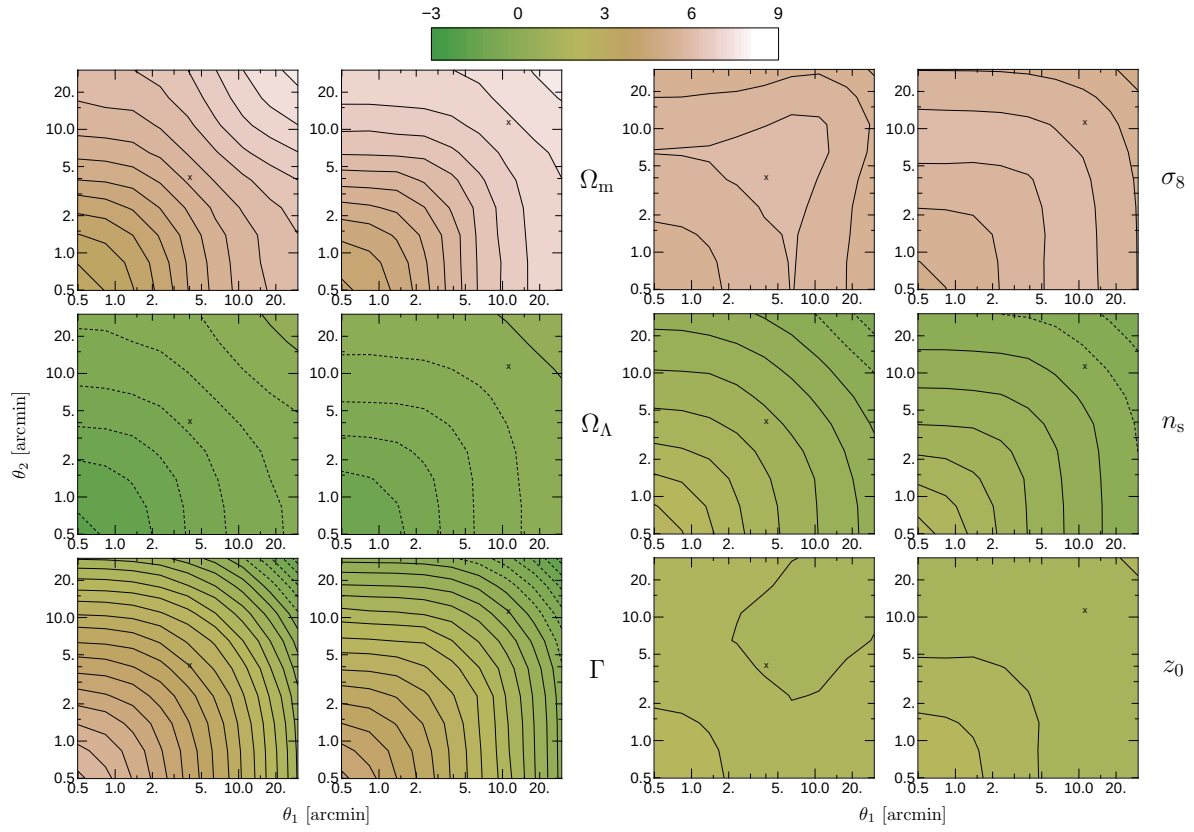

Figure 4.10: *Contours of* $\mathrm{d}\ln\left\langle M_{\rm ap}^3(\theta_1,\theta_2,\theta_3)\right\rangle/\mathrm{d}p$ *from the HEPT model, with* $p = \Omega_{\rm m}, \Omega_\Lambda, \Gamma, \sigma_8, n_{\rm s}$ *and* z_0 *as indicated. In each panel* θ_3 *is fixed to the value indicated by the cross, corresponding to 3.87 (left column) and 10.77 arc minutes (right column), respectively.*

4.3 Visualization of the 3PCF

The apparent difficulty to visualize a polar quantity with eight components depending on three scalar variables is an obstacle in developing an intuitive feeling for the 3PCF of shear. For the second-order case, it is much more simple — there are only two (non-trivial) real components of the two-point correlation function depending on a single scalar which is the modulus of the angular separation of the two points in question. In contrast, the 3PCF is a much more complex function, and the understanding of its behavior and its properties will be valuable for any further studies using this statistics.

Whereas in Sect. 4.1.4 the components of the 3PCF were plotted for special triangles with only one of the arguments varying, now the objective is to display the shear pattern of the 3PCF for more general triangles, which is inspired by Bernardeau et al. (2003). In their work, a special smoothed projection of the 3PCF was introduced and calculated from theoretical predictions of the HEPT model. Following Bernardeau et al. (2003), I generate plots of the shear pattern of the 3PCF from the OCDM ray-tracing simulations (App. B) as follows: The side x_3 of a triangle is kept fix, and the points $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are assigned the coordinates $\boldsymbol{X}_1 = (-x_3/2, 0)^{\mathrm{t}}$ and $\boldsymbol{X}_2 = (x_3/2, 0)^{\mathrm{t}}$. Every point (y_1, y_2) on a grid in the $\mathbb{R}^2$-plane is identified with the third triangle point $\boldsymbol{X}_3$. The natural components $\Gamma^{(q)}$ (2.24) are calculated for the triangle $(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{X}_3)$, and a 'shear stick' is drawn at position $\boldsymbol{X}_3 = (y_1, y_2)$, its length and orientation indicating the absolute value and phase φ of the polar $\Gamma^{(q)}_{\mathrm{cart}} = |\Gamma^{(q)}| \exp(2\mathrm{i}\varphi)$. The sticks can be compared to the orientation of the ellipses in Fig. 2.1 and the corresponding Cartesian coordinates of $\Gamma^{(q)}_{\mathrm{cart}}$ can be inferred.

Two examples of these patterns for $x_3 = 3$ and 10 arc minutes are shown in Figs. 4.11 and 4.12, respectively. These plots reveal a complicated pattern of each of the natural components of the 3PCF. The patterns for $x_3 = 3$ and 10 arc minutes look similar with, on average, a smaller shear amplitude for the second case. All natural components show regions where only the real or imaginary part is present, e.g. $\Gamma^{(0)}$ is mainly real and negative in an elliptical region between the two points $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$. These properties of the 3PCF can be exploited and integrated third-order quantities can be defined as the average over regions of constant 3PCF. This averaging results in a higher signal-to-noise ratio in comparison with the correlation function and might simplify the detection of third-order statistics of cosmic shear. In fact, the first significant measurement of third-order shear was achieved using such an integrated quantity (Bernardeau et al. 2002).

There are also bands or narrow stripes where the 3PCF or its real or imaginary part is vanishing. In some cases these bands are connected with the vertices $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$. In this framework, the symmetry properties that were found in the last sections can be easily recovered. The components $\Gamma^{(1)}$ and $\Gamma^{(2)}$ are real and positive on circles of radius x_3 centered on $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$, respectively, corresponding to isosceles triangles for which $\Gamma^{(1)}$ and $\Gamma^{(2)}$ were shown to be real in Sect. 4.1.4. Moreover, the parity transformation corresponding to a mirror image along a vertical line between $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ interchanges $\Gamma^{(1)}$ and $\Gamma^{(2)}$ — this can be seen clearly in the figures.

4.4 Summary of the study of third-order shear statistics

In this chapter, I study third-order statistics of cosmic shear using theoretical predictions of non-linear models of the large-scale structure, and compare the results with OCDM ray-

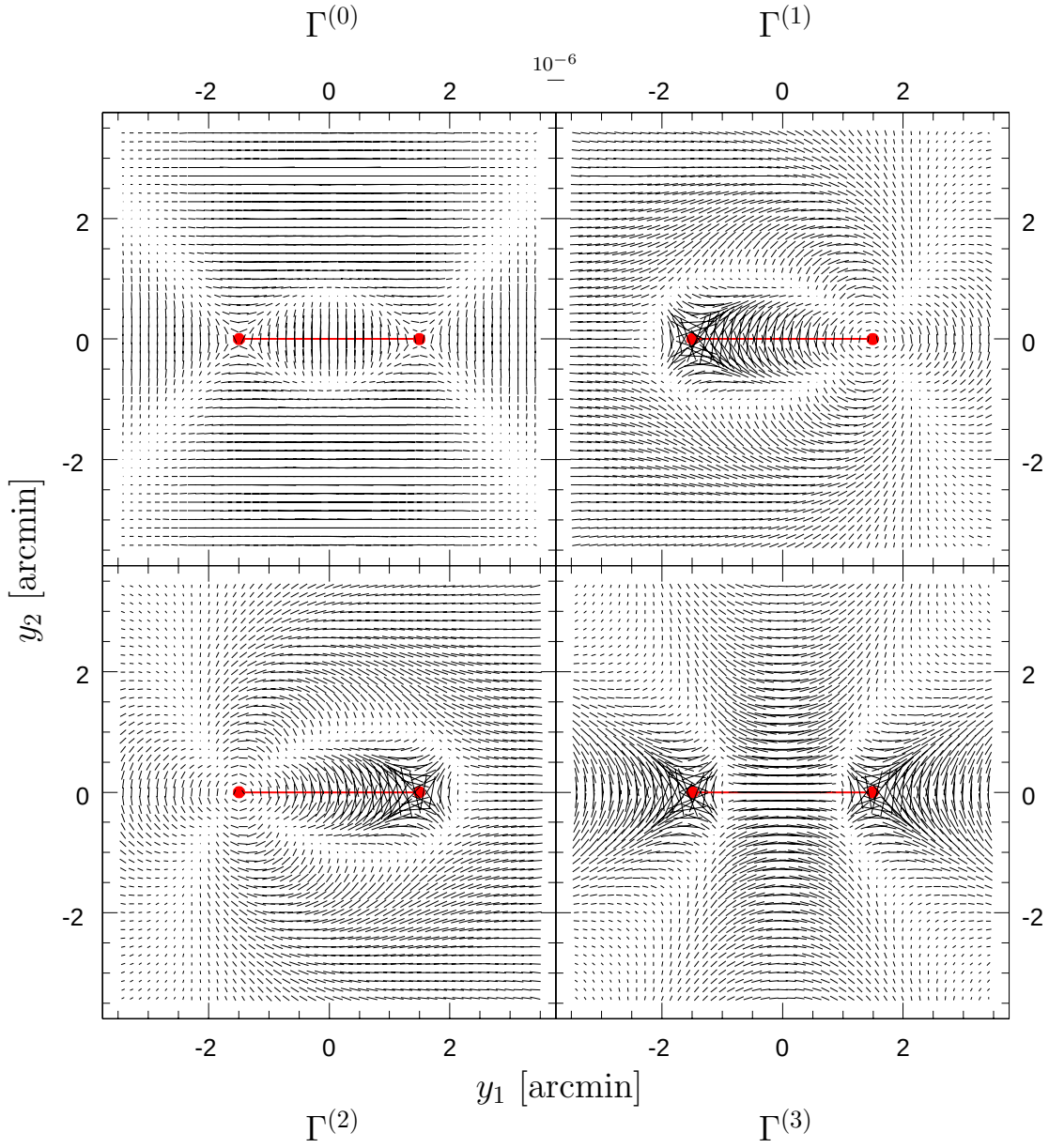

Figure 4.11: *Shear pattern for triangles with $x_3 = 3$ arc minutes. Each shear stick represents the amplitude and phase of $\Gamma^{(q)}_{\rm cart}$, $q = 0 \ldots 3$ for a triangle given by the stick position (y_1, y_2) and the two dots $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$. The size of a stick corresponding to 10^{-6} is indicated on top of the plot. The relation between the stick orientation and the polar angle of the complex shear can be inferred from Fig. 2.1.*

tracing simulations. Both the three-point correlation function and the generalized third-order aperture mass statistics from the theoretical models and the simulations agree reasonably well on angular scales larger than about one arc minute. The symmetry properties of the 3PCF as derived in Chapter 2 are verified numerically and it is shown that for a general triangle, all eight components of the 3PCF are non-zero and have comparably amplitude.

The dependence on cosmological parameters of the third-order aperture mass statistics is

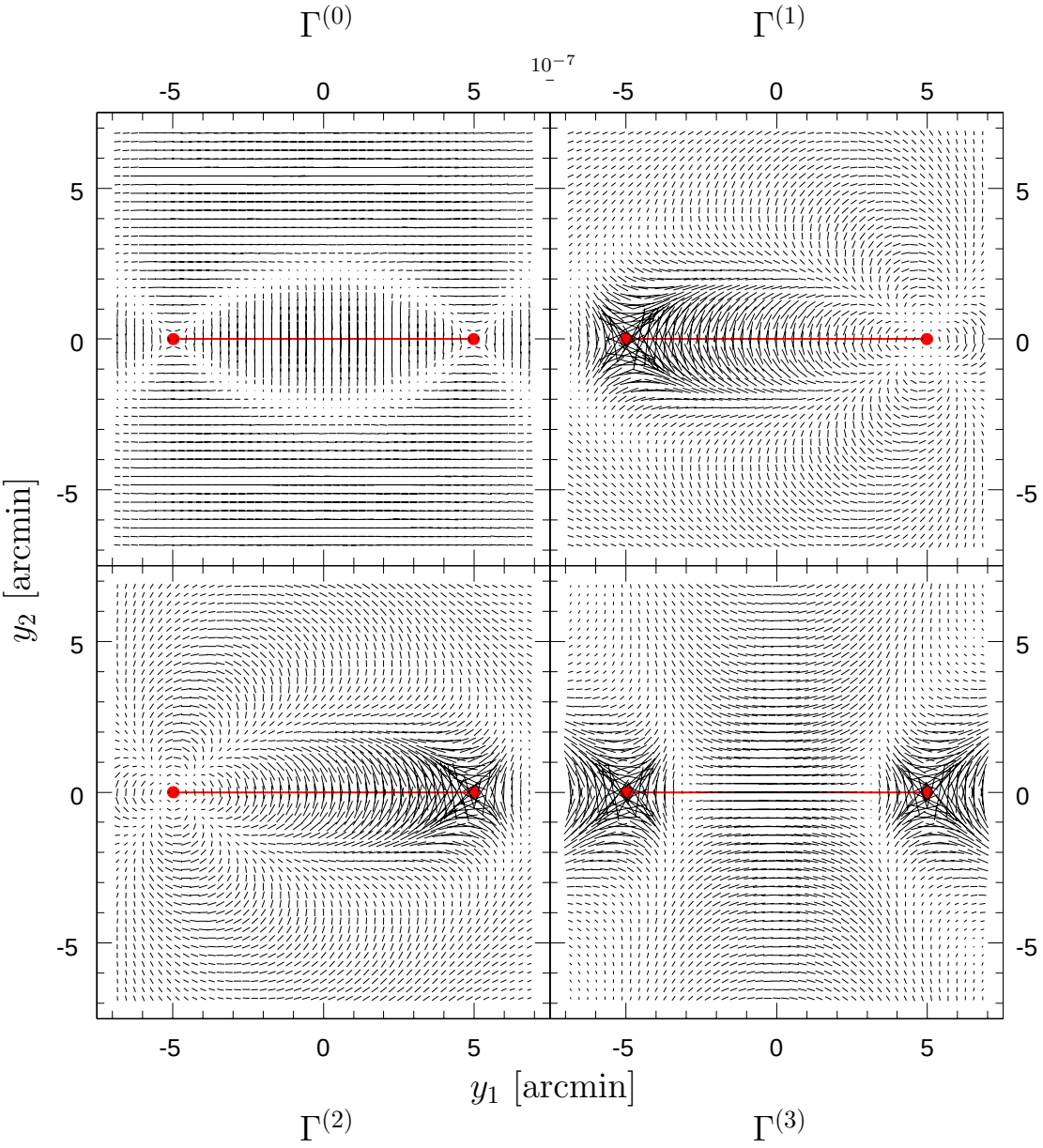

Figure 4.12: *Shear pattern for triangles with $x_3 = 10$ arc minutes, see previous figure for more details. The size of a stick corresponding to 10^{-7} is indicated on top of the plot.*

investigated and it is shown that it is sensitive to Γ and n_s in the same way as $\langle M_\mathrm{ap}^2 \rangle$, but differently sensitive to Ω_m and σ_8 when compared to $\langle M_\mathrm{ap}^2 \rangle$. This implies that whereas the former parameter pair remains highly degenerate under the combination of $\langle M_\mathrm{ap}^2 \rangle$ and $\langle M_\mathrm{ap}^3 \rangle$, the Ω_m-σ_8-degeneracy can be lifted. This prediction will be verified in Chapter 5, and can be seen when comparing the error ellipses for these parameters in Fig. 5.6.

Further, a two-dimensional visualization of the four complex natural components of the 3PCF is done which reveals a complicated pattern as a function of triangle parameters. These patterns can help to find high signal-to-noise integrals of the 3PCF which might be easier detectable than the 3PCF itself.

Chapter 5

Combined second- and third-order aperture mass statistics measurements

In this chapter predictions on cosmological parameter constraints from combined measurements of second- and third-order aperture mass statistics of cosmic shear are presented. The generalized third-order aperture mass statistics $\langle M_{\rm ap}^3 \rangle$ (Sect. 2.3.2) is shown to contain much more information about the bispectrum of the projected matter density than the skewness of the aperture mass $\langle M_{\rm ap,d}^3 \rangle$. From theoretical models (Sect. 1.2.7) as well as from ΛCDM ray-tracing simulations (Sect. 5.1) the moments $\langle M_{\rm ap}^2 \rangle$ and $\langle M_{\rm ap}^3 \rangle$ and their dependence on cosmological parameters are calculated. The covariance including shot noise and cosmic variance of $M_{\rm ap}^2$, $M_{\rm ap}^3$ and their cross-correlation are estimated using ray-tracing simulations (Sect. 5.2). A Fisher matrix analysis is performed and for various combinations of cosmological parameters 1σ-errors are predicted corresponding to a deep 29 square degree cosmic shear survey (Sect. 5.3). Although the parameter degeneracies cannot be lifted completely, the (linear) combination of second- and third-order aperture mass statistics reduces the errors significantly. The near-degeneracy between $\Omega_{\rm m}$ and σ_8, present for all second-order cosmic shear measures, is diminished substantially whereas less improvement is found for the near-degenerate pair consisting of the shape parameter Γ and the spectral index $n_{\rm s}$. Uncertainties in the source galaxy redshift z_0 increase the errors of all other parameters.

5.1 The aperture mass from ray-tracing simulations

In order to estimate the second- and third-order aperture mass statistics and their covariance (Sect. 2.4), 36 ΛCDM ray-tracing simulations are used, kindly provided by T. Hamana (for more details see Ménard et al. 2003). Each field consists of 1024^2 data points of the convergence κ and the shear γ, the pixel size is 0.2 arc minutes. I assume that every pixel corresponds to a galaxy, thus the source galaxy density is 25 per square arc minute. The Poisson noise is much smaller than the shape noise of the ellipticities, and apertures with radii smaller than one arc minute are discarded due to discreteness effects in the ray-tracing and in the underlying N-body simulations. All source galaxies are located at a redshift of about unity. The parameters of the ray-tracing simulations correspond to model 2, see Table B.1.

Because the field κ is given on a regular grid, the aperture mass (2.6) can be calculated very quickly using FFT with the ensemble average replaced by the average over all aperture centers $\boldsymbol{\vartheta}$. However, since for discrete Fourier transforms periodic boundary conditions are assumed which is not the case for the ray-tracing simulations, points near the borders have to be excluded from the averaging. This leads to an overestimation of the covariance of the $M_{\rm ap}$-statistics which increases with the aperture radius. To avoid this, one could calculate

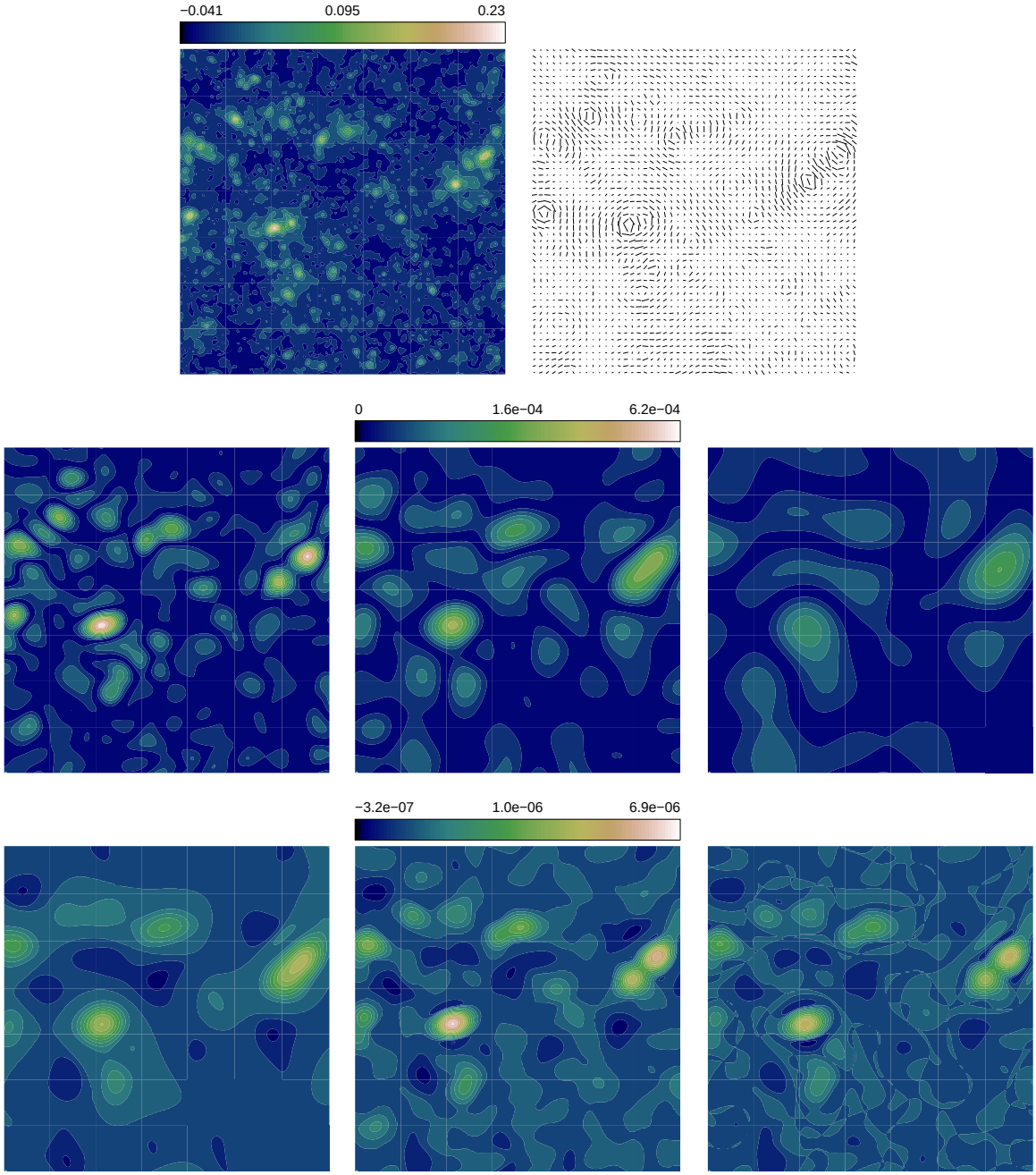

Figure 5.1: Upper row: *The convergence $\kappa(\boldsymbol{\vartheta})$ and the shear $\gamma(\boldsymbol{\vartheta})$ are plotted for an extract of one of the ΛCDM ray-tracing fields. The box size is 40 arc minutes.* Middle row, left to right: *The square of the aperture mass, $M_{\rm ap}^2(\theta, \boldsymbol{\vartheta})$, for $\theta = 1.4, 2.3$ and 3.9 arc minutes, for the same region.* Lower row, left to right: *The product of three aperture masses, $M_{\rm ap}(\theta_1, \boldsymbol{\vartheta}) M_{\rm ap}(\theta_2, \boldsymbol{\vartheta}) M_{\rm ap}(\theta_3, \boldsymbol{\vartheta})$, for $(\theta_1, \theta_2, \theta_3) = (2.3, 2.3, 3.9), (1.4, 1.4, 2.3)$ and $(1.4, 2.3, 3.9)$ arc minutes, for the same region. For the middle and lower row, the color coding is proportional to the square root of the levels in order to increase the contrast.*

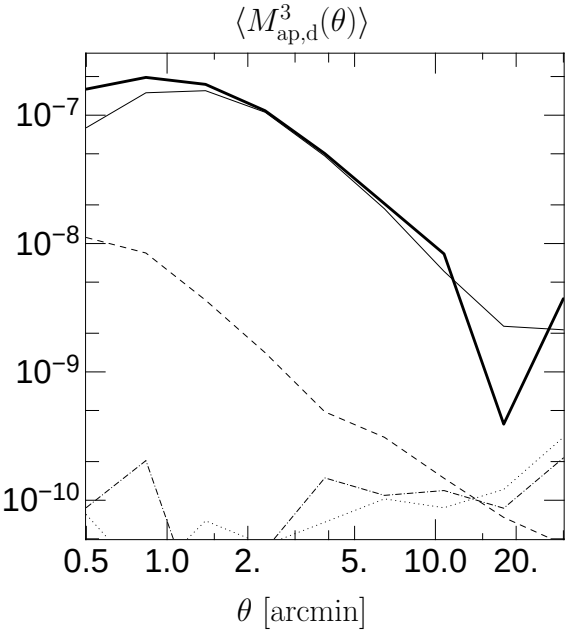

Figure 5.2: *$\langle M^3_{\rm ap,d}\rangle(\theta)$ from apertures (bold line), and from the integration over the 3PCF (thin line). Also plotted are the B-mode signals from the integration method $\langle M_{\rm ap}M_\perp^2\rangle$ (dashed), $\langle M_{\rm ap}^2M_\perp\rangle$ (dotted) and $\langle M_\perp^3\rangle$ (dash-dotted). The curves represent the mean from three of the ray-tracing fields.*

$\left\langle M^2_{\rm ap}\right\rangle$ and $\langle M^3_{\rm ap}\rangle$ from the shear correlation functions using the relations (2.17) and (2.43), which takes into account the complete area. This approach is not chosen here because of the time-consuming calculation of the 3PCF. The correction scheme I apply to the covariance matrices is described in Sect. 5.2.

In Fig. 5.1 an extract of one of the ray-tracing simulations is shown as an example. In the convergence map, over- and underdensities show up clearly which manifest themselves in the shear map with tangential shear patterns. In the maps of the square and cube of the aperture mass, the over- and underdensities in κ are smoothed out according to the filter scales.

Figures 4.7 and 4.8 show $\left\langle M^2_{\rm ap}\right\rangle$ and $\left\langle M^3_{\rm ap}\right\rangle$ from the ΛCDM simulations and the theoretical predictions of HEPT (Sect. 1.2.7.2). The non-linear fitting formulae reproduce reasonably well the results from the simulations for angular scales above ~ 1 arc minute. Since the Gaussian filter (2.10) is significantly larger than zero out to three times the aperture radius, the largest aperture which can be put onto the field is for $\theta_{\rm max} = a/6 = 34'$ where $a = 204.8'$ is the field size.

For comparison, I calculate $\left\langle M^3_{\rm ap}\right\rangle$ by integrating over the 3PCF, using eq. (2.43). Although the fast tree-code algorithm described in Jarvis et al. (2004) is used to calculate the 3PCF, it is still time-consuming. The computation time goes as $b^{-3.3}$ where b is the logarithmic bin width and a fine binning of the 3PCF is required in order not to introduce a large B-mode, see below. For the choice of $b = 0.075$ used here, the integration method takes about a factor of 500 longer than the aperture method using FFT.

The results are shown in Fig. 5.2 and represent the average over three of the ray-tracing fields. $\left\langle M^3_{\rm ap}\right\rangle$ as calculated via the FFT method cannot be determined for large radii because of the border effects, as mentioned above. Since $\left\langle M^3_{\rm ap}\right\rangle$ obtained via integrating over the 3PCF is based on the simulated shear field, I use the γ fields instead of the κ fields in order to calculate $\left\langle M^3_{\rm ap}\right\rangle$ via the FFT aperture method, using the second equality of the first line in (2.6). With $M_\perp(\theta) = \Im\left(Q'_\theta * \gamma\right)$, I also determine the statistics $\langle M^2_{\rm ap}M_\perp\rangle$, $\langle M_{\rm ap}M^2_\perp\rangle$ and $\langle M^3_\perp\rangle$ as indicators of a B-mode (Sect. 1.4.3). $\langle M_{\rm ap}M^2_\perp\rangle$ is expected to vanish if the ray-tracing simulations are B-mode-free. The two quantities with odd power in $M_\perp$ can only be

non-zero for a convergence field which is not parity-invariant (Schneider 2003). I find all three statistics to be three and more orders of magnitude below the pure E-mode, confirming that the ray-tracing simulations contain virtually no B-mode and are parity-symmetric. However, $\langle M_{\rm ap} M_\perp^2 \rangle$ inferred from the 3PCF is at a couple of percent of the E-mode. This is most probable due to the binning of the 3PCF — the B-mode gets smaller when the binning is refined. As can be seen in Fig. 5.2, there is good agreement between the two methods, except for very small angular scales (where the B-mode is of the order 10%) and large aperture radii (where a significant fraction of the field near the border cannot be taken into account with the aperture method).

5.2 Covariance matrices of the aperture mass statistics

I use the ray-tracing fields to estimate the covariance matrices of the second- and generalized third-order aperture mass statistics. The averaging in eq. (2.54) is performed over the different simulations. Because of the small number of realizations, I split up each of the 36 fields into 4 subfields and average over the resulting 144 realizations. This is necessary in order to avoid a singular covariance matrix, see Sect. 5.2.3. Adjacent subfields do not represent fully independent realizations of the convergence field but the correlations are negligible: when averaging over only a bootstrapped subset of subfields, no systematic deviation but only a noisier estimate of the covariance is obtained. Note that because of the splitting, the maximum usable aperture radius is now 17 arc minutes.

The fact that a region around the border cannot be used depending on the aperture radius θ results in an effective area $A_{\rm eff}(\theta)$ which is smaller than the original area $A = a^2$, namely $A_{\rm eff}(\theta) = (a - 6\theta)^2$. Since the covariance is inversely proportional to the observed area, a correction scheme is applied by multiplying each covariance matrix entry $C(\theta_1, \theta_2)$ with $\sqrt{A_{\rm eff}(\theta_1) A_{\rm eff}(\theta_2)}/A$ in the case of $\langle M_{\rm ap}^2 \rangle$ and $\langle M_{\rm ap,d}^3 \rangle$. For the generalized third-order aperture mass, where each matrix element corresponds to two triplets of aperture radii, the effective area corresponding to the maximum radius of each triplet is inserted for the correction factor. Altogether, this correction makes sure that the covariance matrix corresponds to the same survey area A for all aperture radii.

For the Fisher matrix analysis of cosmological parameters (Sect. 5.3), the covariance matrices obtained from the 2.9 square degree fields are rescaled to a corresponding survey area of 29 square degrees by division with 10, making use of their $1/A$-dependence. Note that this increase of survey area is not equivalent of extending a single patch on the sky, since this extended observed area will not sample independent but correlated parts of the large-scale structure and the decrease in cosmic variance will be less than the increase in area. This rescaling of the area corresponds to observing 10 independent lines of sight, each one 2.9 square in area.

5.2.1 Adding intrinsic ellipticities

In order to realistically model the noise coming from the intrinsic ellipticities of the source galaxies, one would have to add a random ellipticity to each shear value. It has been shown that this is equivalent to adding a noise term to the convergence κ (van Waerbeke 2000). For mass reconstructions, this noise has to be added to a smoothed κ map. Here however, no smoothing is required (for a justification see below), thus, to each pixel of κ, I add a random

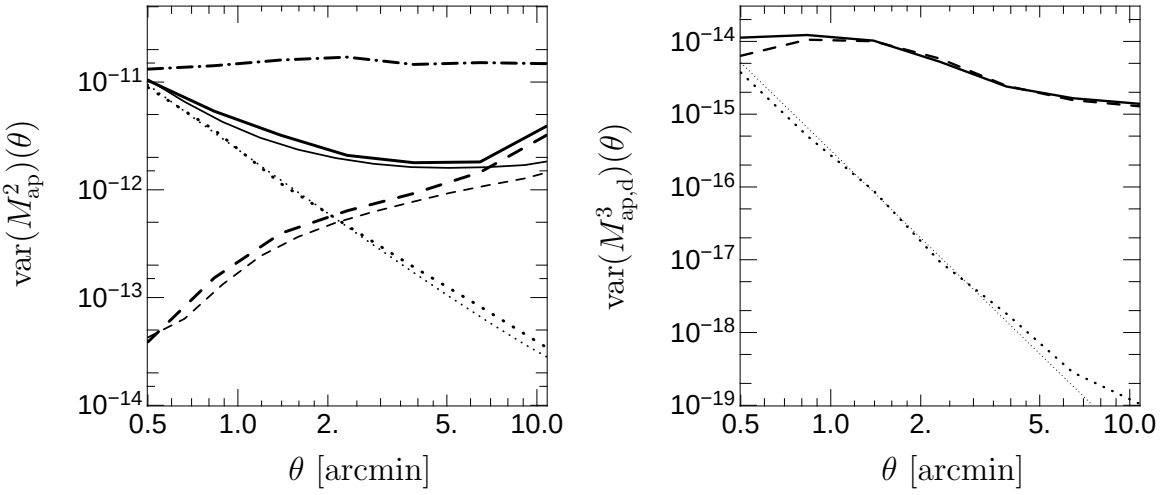

Figure 5.3: Left panel: *The variance of M_{ap}^2 from the Gaussianized ΛCDM simulations (bold lines), in comparison with the Monte-Carlo integration (thin lines), for a survey area of $A = (102.4')^2$. The dotted lines give the variance from shot noise only (due to the intrinsic ellipticity dispersion), the dashed curves correspond to cosmic variance only. The solid line includes both error sources; note that it is not the sum of the other two curves — there is a non-vanishing mixed term. The dash-dotted line indicates the cosmic variance term for the non-Gaussian case.* Right panel: *The diagonal of the covariance matrix of $M_{\mathrm{ap,d}}^3(\theta)$, for the same survey area. The dotted line (bold: simulations, thin: theoretical value, eq. 2.64) indicates shot noise only, the dashed curve is the cosmic variance contribution and the solid line includes both error sources.*

Gaussian variable with zero mean and dispersion $0.3/\sqrt{2}$. This is equivalent to adding to each shear value a complex random ellipticity with $\langle\varepsilon^{(\mathrm{s})}\rangle = 0$ and $\langle\varepsilon^{(\mathrm{s})}(\varepsilon^{(\mathrm{s})})^*\rangle^{1/2} = \sigma_\varepsilon = 0.3$.

In the case of $\langle M_{\mathrm{ap}}^2\rangle$, this method of adding ellipticity noise yields the predicted amplitude to the variance without need of smoothing, as can be seen in Fig. 5.3. The shot-noise contribution to the variance is in good agreement with the Monte-Carlo method described in Sect. 3.2. The shot-noise term of the variance of M_{ap}^3 agrees very well with the analytical expectation (2.64), see Fig. 5.3, except for large θ, where only few apertures can be placed onto the field which are not too close to the border. Apparently, adding intrinsic random ellipticities to each grid point without smoothing introduces no artefacts.

5.2.2 Gaussianized fields

In order to compare the covariance of M_{ap}^2 from the ray-tracing simulations with the Monte-Carlo integration of the analytical expressions (2.57) as a sanity check, I transform the ray-tracing simulations into Gaussian fields without changing the power spectrum. This is achieved by multiplying the Fourier transform $\hat{\kappa}$ of each convergence field by random phases (destroying the phase correlations). Then for each Fourier mode $\boldsymbol{k}$, a $\hat{\kappa}(\boldsymbol{k})$ value is picked randomly from one of the 36 fields repeatedly to create 36 new Gaussianized realizations.

Destroying the phase correlations for each individual field independently would not have led to the desired goal. Randomizing the phases cancels the connected 4-point term (kurtosis) of each individual realization, but not the kurtosis of the underlying ensemble. The estimator of the covariance (2.54) is independent of the kurtosis of each individual realization because

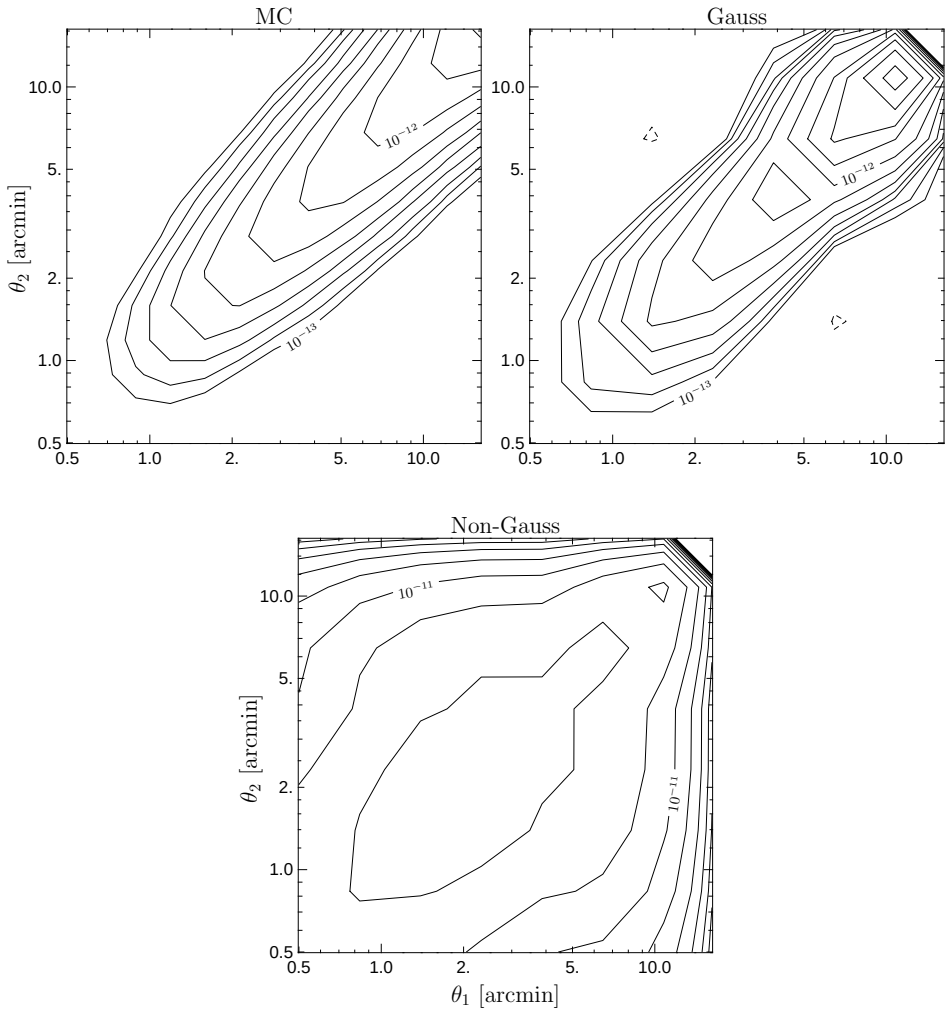

Figure 5.4: *Contours of the cosmic-variance term of the covariance of $M_{\rm ap}^2$, for the Monte-Carlo method (upper left panel), the Gaussianized ray-tracing fields (upper right) and the original fields (lower panel), for a survey area of $A = (102.4')^2$. The contours are logarithmically spaced.*

first $\langle M_{\rm ap}^2 \rangle$ is determined for each field and then the square of this quantity is averaged over all fields — thus for this averaging, only second-order quantities are taken into account. The process of remixing the $\hat{\kappa}$-fields in Fourier space annihilates the kurtosis of the underlying ensemble, and the resulting fields represent realizations of a Gaussian random field.

In Fig. 5.3, the variance (diagonal of the covariance) of $M_{\rm ap}^2$ is plotted. The results from the ray-tracing simulations are in fairly good agreement with the Monte-Carlo method, although the cosmic variance term from the ray-tracings is slightly higher than the one from the Monte-Carlo integration.

It is clear from this figure that non-Gaussianity increases the noise level on the diagonal by an enormous amount, about two orders of magnitude at $\sim 1'$. The ratio of the non-Gaussian to the Gaussian variance is $\propto \theta^{-2}$ for small θ and becomes less steep for larger θ.

On nearly all scales, cosmic variance dominates over the shot noise. From Fig. 5.4, one sees that due to mode-coupling, high cross-correlations between different angular scales are introduced, present on the off-diagonal of the covariance.

5.2.3 The case of $\langle M_{\rm ap}^3 \rangle$

The variance of the aperture mass skewness, $M_{\rm ap,d}^3(\theta) = M_{\rm ap}^3(\theta,\theta,\theta)$ is plotted in Fig. 5.3. As for the second-order case, cosmic variance dominates over shot noise on all but very small scales.

The covariance matrix of $M_{\rm ap}^3$ is not diagonal-dominant, as can be seen in Fig. 5.5, where the correlation coefficient of $\mathsf{C}(M_{\rm ap}^3)$ is plotted. The self-similar pattern and the many secondary diagonals originate from the reordering of $\langle M_{\rm ap}^3(\theta_1,\theta_2,\theta_3)\rangle$ into a single vector, which inevitably creates repeating entries of similar combinations of aperture radii. The correlation of $\langle M_{\rm ap}^2\rangle$ for two aperture radii $\theta_1 \geq \theta_2$ is a rapidly decreasing function of the ratio θ_1/θ_2. In the case of $\langle M_{\rm ap}^3\rangle$, however, there are many combinations of filter scales which show high correlation. This fact together with the small sample of realizations of κ-fields causes the covariance matrix to be very ill-conditioned. For the Fisher matrix analysis (Sect. 5.3.2) the covariance matrix has to be inverted. Stable results are obtained when the ratio of adjacent aperture radii is not too small, i.e. larger than about 1.5, and therefore, the number of data points for $M_{\rm ap}^3$ is not too large. This is in agreement with Pan & Szapudi (2005) who claim that the rank of a covariance matrix obtained from averaging over $N_{\rm sim}$ realizations cannot be more than $N_{\rm sim}$, and that in practice not more than $N_{\rm sim}/2$ modes should be used.

One way to determine whether the estimate of the covariance of $M_{\rm ap}^3$ is reasonable would involve 6-point statistics, which is not feasible analytically. Instead, I slightly modify the aperture radii used in the analysis and get a rough estimate of the accuracy of this method. I comment on the stability of the results in Sect. 5.3.4.

5.3 Constraints on cosmological parameters

From the simulated data, a data vector $\boldsymbol{x}$ is "observed", which consists of n values of $\left\langle M_{\rm ap}^2\right\rangle$ and/or $\left\langle M_{\rm ap}^3\right\rangle$ as a function of angular scales. Using a theoretical model, and approximating the observables as Gaussian variables, the likelihood function for an n-dimensional multivariate Gaussian distribution which is defined in eq. (3.2) is constructed.

5.3.1 The input data

The following five cases for the input data vector $\boldsymbol{x}$ and its covariance C are distinguished:

1. ('2') $x_l = \left\langle M_{\rm ap}^2(\theta_l)\right\rangle$, C = covariance of $M_{\rm ap}^2$.
2. ('3') $x_l = \left\langle M_{\rm ap}^3(\theta_i,\theta_j,\theta_k)\right\rangle$ for a combination of three filter radii which after relabeling corresponds to the index l as described in Sect. 2.5.2. x_l is organized such that $(\theta_i,\theta_j,\theta_k)$ is in lexical order and $\theta_i \leq \theta_j \leq \theta_k$. C = covariance of $M_{\rm ap}^3$.
3. ('3d') $x_l = \langle M_{\rm ap,d}^3(\theta_l)\rangle$, C = covariance of $M_{\rm ap,d}^3$.
4. ('2+3d') x_l = some element from the concatenated data vector containing $\left\langle M_{\rm ap}^2\right\rangle$ and $\langle M_{\rm ap,d}^3\rangle$. C is a block matrix containing the covariance matrices of $M_{\rm ap}^2$ and $M_{\rm ap,d}^3$ on the diagonal and the cross-correlation on the off-diagonal.
5. ('2+3') x_l = some element from the concatenated data vector containing $\left\langle M_{\rm ap}^2\right\rangle$ and $\left\langle M_{\rm ap}^3\right\rangle$. C is a block matrix containing the covariances of $M_{\rm ap}^2$ and $M_{\rm ap}^3$ on the diagonal and the cross-correlation on the off-diagonal.

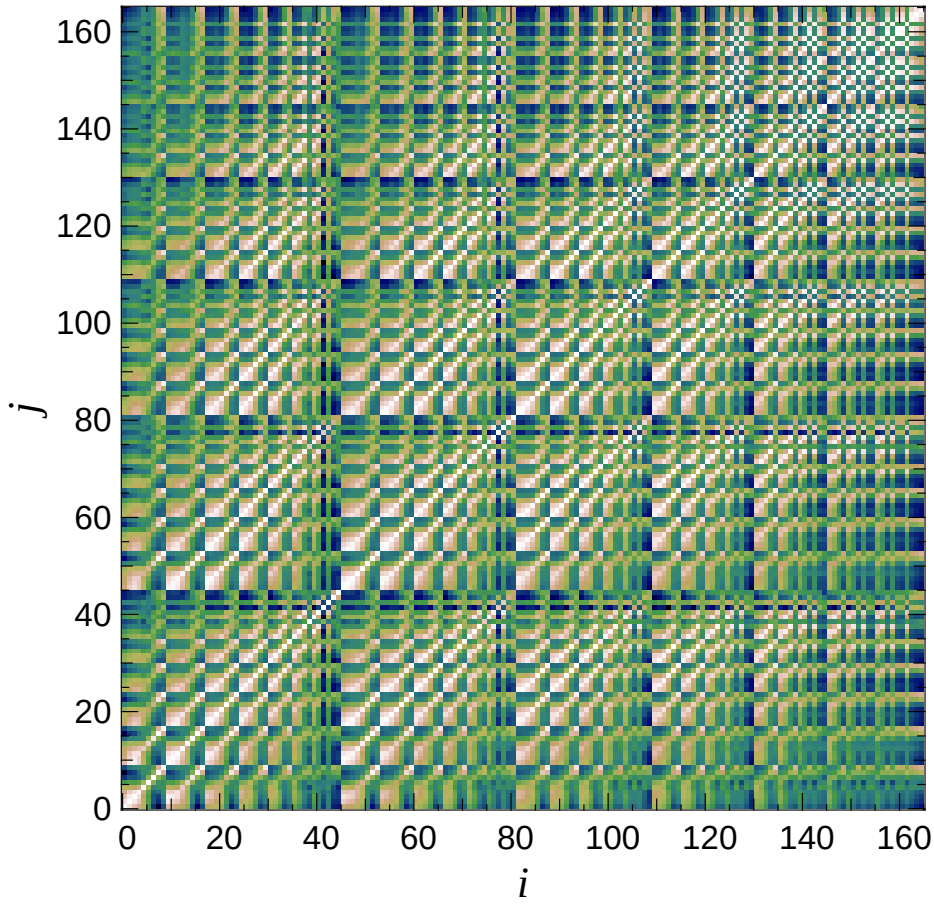

Figure 5.5: *The correlation coefficient* $r_{ij} = \mathsf{C}(M_{\rm ap}^3)_{ij}/[\mathrm{var}(M_{\rm ap}^3)_i\,\mathrm{var}(M_{\rm ap}^3)_j]^{1/2}$ *of the covariance of* $M_{\rm ap}^3$*, corresponding to a field size of* $A = (204.8')^2$*. 9 filter radii between 0.5 and 30 arc minutes are used (logarithmic spacing). See Sect. 2.5.2 for the correspondence between the matrix indices* i, j *and the filter scale triplets.*

The survey area is 29 square degrees corresponding to ten uncorrelated fields, each of size $A = (102.4')^2$. The covariance matrices are calculated by averaging over the realizations as described in Sect. 5.2. Six different filter radii are used, logarithmically spaced between 1 and 15 arc minutes and yielding six data points for both of $\langle M_{\rm ap}^2\rangle$ and $\langle M_{\rm ap,d}^3\rangle$ and 56 for $\langle M_{\rm ap}^3\rangle$.

5.3.2 Fisher matrix analysis

As in the case of the survey strategy analysis (Chapter 3), the Fisher matrix (3.7) corresponding to a local expansion of the likelihood around the maximum, instead of the full likelihood, is analyzed. The Fisher matrix elements for the five combinations of second- and third-order aperture mass statistics considered here (see Sect. 5.3.1) are given in Table 5.3.2.

For various combinations of cosmological parameters, the MVBs (3.5) are computed by inverting the Fisher matrix or a submatrix of it, corresponding to the parameter combination considered. First, the analysis is done for only two parameters, in order to graphically display the MVBs (Sect. 5.3.2.1). Then, simultaneous MVBs for three and more parameters are calculated (Sect. 5.3.2.2).

Table 5.1: *Fisher matrix for the five different input data as listed in Sect. 5.3.1, denoted by '2', '3', '3d', '2+3d' and '2+3', respectively. The survey area is 29 square degrees, the table entries are given in units of 10^4.*

		Ω_m	Ω_Λ	Γ	σ_8	n_s	z_0
2	Ω_m	1.766	−0.268	1.543	1.597	0.601	1.079
	Ω_Λ	−0.268	0.239	−0.784	−0.424	−0.336	−0.413
	Γ	1.543	−0.784	2.875	1.905	1.202	1.630
	σ_8	1.597	−0.424	1.905	1.618	0.766	1.200
	n_s	0.601	−0.336	1.202	0.766	0.507	0.674
	z_0	1.079	−0.413	1.630	1.200	0.674	0.975
3	Ω_m	1.698	−0.315	1.260	1.848	0.478	0.661
	Ω_Λ	−0.315	0.445	−1.304	−0.852	−0.552	−0.452
	Γ	1.260	−1.304	3.952	2.803	1.653	1.389
	σ_8	1.848	−0.852	2.803	2.705	1.137	1.146
	n_s	0.478	−0.552	1.653	1.137	0.701	0.579
	z_0	0.661	−0.452	1.389	1.146	0.579	0.544
3d	Ω_m	0.263	−0.079	0.291	0.342	0.109	0.119
	Ω_Λ	−0.079	0.146	−0.422	−0.255	−0.177	−0.141
	Γ	0.291	−0.422	1.236	0.794	0.513	0.418
	σ_8	0.342	−0.255	0.794	0.635	0.319	0.284
	n_s	0.109	−0.177	0.513	0.319	0.214	0.173
	z_0	0.119	−0.141	0.418	0.284	0.173	0.144
2+3d	Ω_m	3.256	−0.071	1.844	2.322	0.713	1.757
	Ω_Λ	−0.071	0.363	−0.999	−0.466	−0.434	−0.379
	Γ	1.844	−0.999	3.613	2.405	1.527	1.924
	σ_8	2.322	−0.466	2.405	2.180	0.974	1.592
	n_s	0.713	−0.434	1.527	0.974	0.653	0.799
	z_0	1.757	−0.379	1.924	1.592	0.799	1.337
2+3	Ω_m	7.046	0.348	2.457	4.440	0.939	3.219
	Ω_Λ	0.348	0.764	−1.775	−0.740	−0.773	−0.435
	Γ	2.457	−1.775	5.846	3.925	2.486	2.756
	σ_8	4.440	−0.740	3.925	4.162	1.597	2.635
	n_s	0.939	−0.773	2.486	1.597	1.076	1.159
	z_0	3.219	−0.435	2.756	2.635	1.159	2.169

5.3.2.1 Two parameters

In Fig. 5.6, the MVBs are shown as ellipses in two-dimensional subspaces of the parameter space. The hidden parameters are fixed. In all cases, the combination of $\langle M_\mathrm{ap}^2 \rangle$ and $\langle M_\mathrm{ap}^3 \rangle$ leads to a substantial reduction in the 1σ-error. As expected, the generalized third-order aperture mass statistics yields better constraints than the 'diagonal' version $\langle M_\mathrm{ap,d}^3 \rangle$. The direction of degeneracy is slightly different for some parameter pairs, most notably when the source redshift parameter z_0 is involved, making the combination of the statistics very effective in these cases. The Ω_m-σ_8-degeneracy is lifted partially and the combined Fisher

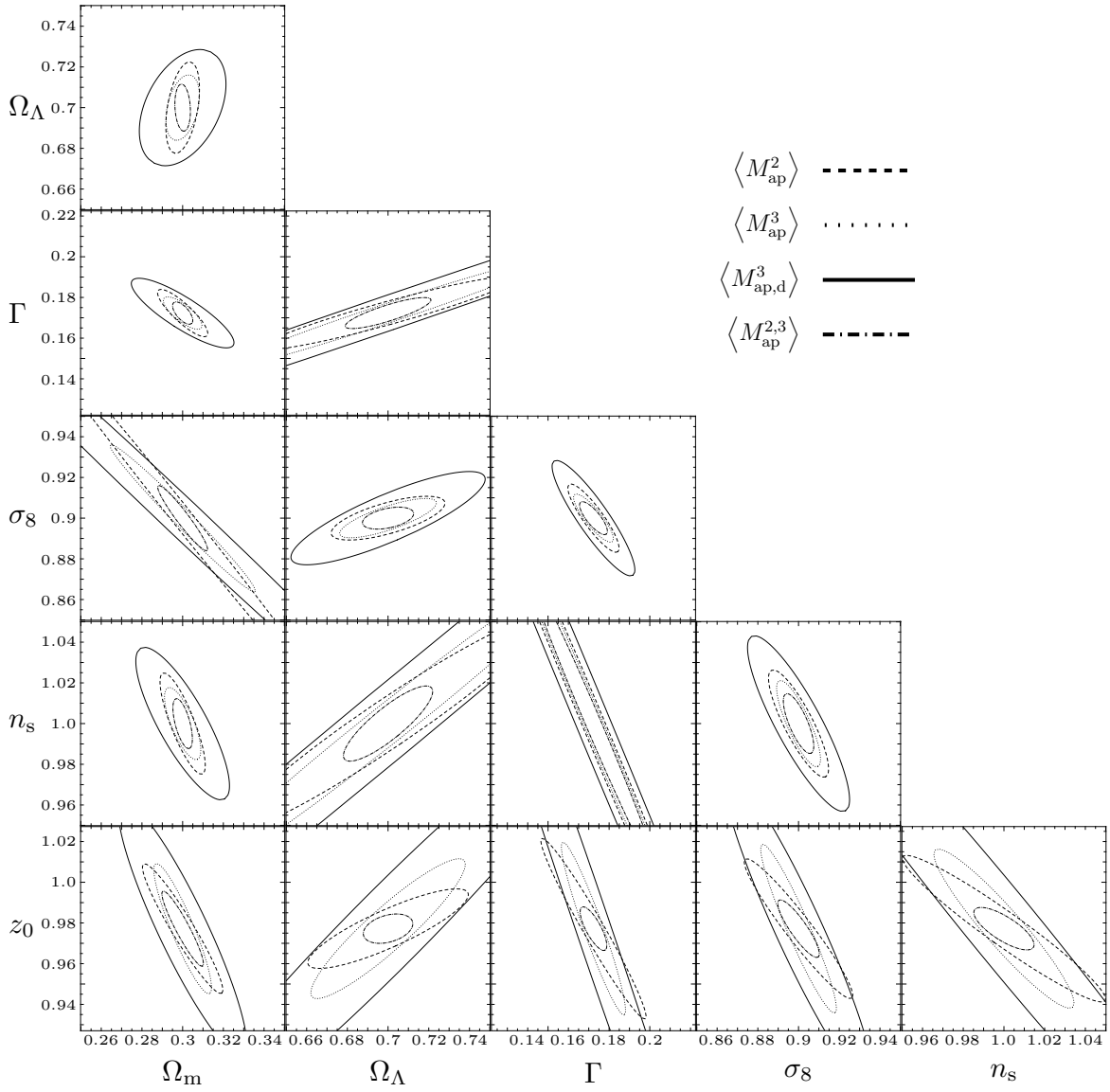

Figure 5.6: *1σ-error ellipses from the Fisher matrix. The hidden parameters are kept fixed. Dashed line: $\langle M_{\rm ap}^2 \rangle$, dotted line: $\langle M_{\rm ap}^3 \rangle$, solid line: $\langle M_{\rm ap,d}^3 \rangle$, dash-dotted line: combination of $\langle M_{\rm ap}^2 \rangle$ and $\langle M_{\rm ap}^3 \rangle$ as described in Sect. 5.3.1. If one of the parameters is $\Omega_{\rm m}$, a flat Universe is assumed (except for the $\Omega_{\rm m}$-Ω_Λ-plot). The assumed survey area is 29 square degrees.*

matrix analysis yields a large improvement on the error of the two parameters. Contrary to that, the pair $(\Gamma, n_{\rm s})$ is degenerate to a high level for both $\langle M_{\rm ap}^2 \rangle$ and $\langle M_{\rm ap}^3 \rangle$ as well as for their combination.

Note that the combined 1σ-errors are not completely determined by the product of the likelihoods of $\langle M_{\rm ap}^2 \rangle$ and $\langle M_{\rm ap}^3 \rangle$. The combined covariance is not the direct product of the covariance of $\langle M_{\rm ap}^2 \rangle$ and $\langle M_{\rm ap}^3 \rangle$ because of the contribution from the cross-correlation between both statistics.

It is not surprising that the directions of degeneracy between parameters are quite similar for $\langle M_{\rm ap}^2 \rangle$ and $\langle M_{\rm ap}^3 \rangle$, with larger differences existing when z_0 is one of the free parameters.

Both statistics depend on the convergence power spectrum, because in HEPT as well as in quasi-linear PT the bispectrum of the matter fluctuations is given in terms of the power spectrum (1.94). The differences between $\langle M_{\rm ap}^2 \rangle$ and $\langle M_{\rm ap}^3 \rangle$ mainly come from their different dependence on the projection prefactor and the lens efficiency G (eq. 1.89). The projection is most sensitive to the source redshift and of all parameters, changes in z_0 show up in a most distinct way for $\langle M_{\rm ap}^2 \rangle$ and $\langle M_{\rm ap}^3 \rangle$.

Since the degeneracy direction between $\langle M_{\rm ap}^2 \rangle$ and the skewness $\langle M_{\rm ap,d}^3 \rangle$ are very similar, not much improvement is obtained when these two statistics are combined, therefore the corresponding error ellipses are not drawn in Fig. 5.6.

5.3.2.2 Three and more parameters

I calculate the MVBs for three and more parameters simultaneously for various combinations of parameters and for each input data as described in Sect. 5.3.1. The results are given in Table 5.2. All hidden parameters are fixed to the fiducial values of model 2, see Table B.1. If not both $\Omega_{\rm m}$ and Ω_Λ vary, a flat Universe is assumed.

In most of the cases, the error bars from the generalized third-order aperture mass statistics $\langle M_{\rm ap}^3 \rangle$ are smaller than those from its second-order counterpart $\langle M_{\rm ap}^2 \rangle$. This trend gets stronger the more free cosmological parameters are involved, since the measurement of $\langle M_{\rm ap}^3 \rangle$ provides more data points and therefore more degrees of freedom[1]. The skewness of the aperture mass $\langle M_{\rm ap,d}^3 \rangle$ provides by far the worst constraints on the parameters.

In all of the cases, the combination of $\langle M_{\rm ap}^2 \rangle$ and $\langle M_{\rm ap}^3 \rangle$ yields an improvement on the parameter constraints. This improvement may be rather small, e.g. in the cases when both Γ and $n_{\rm s}$ are involved. Then the combined MVB is dominated mainly by the MVB of $\langle M_{\rm ap}^3 \rangle$, and the additional information from $\langle M_{\rm ap}^2 \rangle$ is unimportant. However, for the majority of parameter combinations the combined error is a factor of two and more smaller, indicating that the dependence of the two statistics on the cosmological parameters is different to some degree, and their combination lifts the near-degeneracy substantially. Amongst other, this is true for the pair $\Omega_{\rm m}$ and σ_8. Even if a rather good constraint on these two parameters from $\langle M_{\rm ap}^3 \rangle$ is combined with a large MVB coming from $\langle M_{\rm ap}^2 \rangle$, the combined error can be reduced by a factor of two and more, thus the most prominent parameter degeneracy for second-order cosmic shear between $\Omega_{\rm m}$ and σ_8 can efficiently be broken by adding third-order statistics.

When $\langle M_{\rm ap}^2 \rangle$ is combined with the generalized aperture-mass statistics (the case '2+3') and the skewness ('2+3d'), the former combination always yields better parameter constraints than the latter. For three free parameters, the first combination is typically a factor of two better, if more parameters are involved the improvement is even larger, up to a factor of ten when all six parameters are free. Thus, the preference of $\langle M_{\rm ap}^3 \rangle$ over the skewness of $M_{\rm ap}$ is justified also when it is combined with the second-order aperture mass statistics.

In general, constraints on the cosmological constant Ω_Λ are weaker than for the other parameters, and although the combination of second- and third-order aperture mass statistics gives some improvement on the error, Ω_Λ remains the least known parameter.

5.3.3 Correlation between parameters

Table 5.3 shows the correlation coefficient (3.20) between all cosmological parameters considered in this chapter. For the combination of $\langle M_{\rm ap}^2 \rangle$ and $\langle M_{\rm ap,d}^3 \rangle$ ('2+3d'), the correlation is

[1] This is true only to some extent since the data points are correlated.

Table 5.2: *MVBs for various combinations of three and more cosmological parameters, corresponding to a 29 square degree survey. The hidden parameters are kept fixed. '2', '3', '3d', '2+3d' and '2+3' stand for the five different input data as described in Sect. 5.3.1. If Ω_Λ is not a free parameter, a flat Universe is assumed. The table is continued on the next page.*

	Ω_m	Ω_Λ	Γ	σ_8	n_s	z_0
2	0.077			0.104		0.035
3	0.041			0.053		0.047
3d	0.189			0.219		0.116
2+3d	0.027			0.030		0.029
2+3	0.016			0.016		0.019
2	0.087		0.015	0.119		
3	0.063		0.017	0.078		
3d	0.267		0.043	0.300		
2+3d	0.027		0.012	0.044		
2+3	0.015		0.008	0.024		
2	0.083			0.110	0.029	
3	0.059			0.072	0.035	
3d	0.222			0.242	0.077	
2+3d	0.026			0.040	0.025	
2+3	0.014			0.022	0.017	
2	0.017		0.093		0.200	
3	0.010		0.051		0.113	
3d	0.343		0.158		0.347	
2+3d	0.010		0.063		0.138	
2+3	0.005		0.035		0.078	
2	0.087	0.106		0.113		
3	0.057	0.089		0.067		
3d	0.244	0.328		0.263		
2+3d	0.033	0.056		0.047		
2+3	0.019	0.037		0.028		
2	0.157		0.204	0.241	0.406	
3	0.063		0.056	0.080	0.115	
3d	0.976		0.786	1.204	1.389	
2+3d	0.030		0.089	0.056	0.175	
2+3	0.015		0.041	0.026	0.085	
2	0.095	0.744	0.353	0.285		
3	0.065	0.117	0.053	0.085		
3d	0.387	0.817	0.490	0.542		
2+3d	0.066	0.243	0.078	0.053		
2+3	0.028	0.088	0.026	0.030		

Table 5.2: *MVB table continued.*

	$\Omega_{\rm m}$	Ω_Λ	Γ	σ_8	$n_{\rm s}$	z_0
2	0.569		0.270	0.552		0.645
3	0.080		0.033	0.088		0.091
3d	1.113		0.498	1.084		1.332
2+3d	0.218		0.090	0.224		0.213
2+3	0.069		0.031	0.072		0.069
2	0.674	2.447	2.199	1.857	1.618	
3	0.065	0.127	0.065	0.085	0.125	
3d	7.306	1.349	7.613	9.686	8.344	
2+3d	0.075	0.263	0.101	0.070	0.186	
2+3	0.031	0.094	0.043	0.034	0.090	
2	3.554	5.535	2.677	3.719	2.169	3.553
3	0.110	0.328	0.084	0.092	0.140	0.220
3d	7.532	23.48	11.87	10.43	13.15	17.83
2+3d	0.928	1.101	0.620	0.996	0.701	0.812
2+3	0.079	0.124	0.050	0.080	0.090	0.072

very large for all parameter pairs, the difference to unity in some cases is only of the order of 10^{-3}. The degeneracy directions of $\langle M_{\rm ap}^2\rangle$ and $\langle M_{\rm ap,d}^3\rangle$ are very similar, thus the combination of the two causes the correlation between parameters to be very high.

5.3.4 Stability

In order to check the Fisher matrix analysis for consistency and stability towards small changes of the input data, the calculations are redone with slightly different aperture radii. For changes of a couple of percent in the aperture radii, the resulting Fisher matrix elements vary of the order of up to 10 percent. The MVBs (3.5) fluctuate by about the same amount if two or three parameters are considered to be determined from the data simultaneously. However, for four and five free parameters, the MVBs are less stable, since the Fisher matrix is numerically very ill-conditioned and the inversion is a non-linear operation. In general, the MVBs for $\langle M_{\rm ap}^3\rangle$ are less stable than the ones for $\langle M_{\rm ap}^2\rangle$.

The eigenvectors of $F_{\alpha\beta}^{-1}$ are less affected by a different sampling of the aperture radius. Angles between original and modified eigenvectors are typically only a few degrees. The variation of the correlation coefficient $r_{\alpha\beta}$ is less than ~ 0.1 if up to four parameters are considered. For a higher-dimensional Fisher matrix however, the variation can be higher, similar to the case of the MVB.

5.4 Summary and conclusions of the combined $\langle M_{\rm ap}^2\rangle$- and $\langle M_{\rm ap}^3\rangle$-analysis

The power spectrum of large-scale (dark-)matter fluctuations was until recently the most important quantity that has been measured — directly or indirectly — by cosmic shear.

Table 5.3: *The correlation coefficient $r_{\alpha\beta}$ (3.20) of the inverse Fisher matrix (3.20). '2', '3', '3d', '2+3d' and '2+3' stand for the five different input data as described in Sect. 5.3.1. Note that the correlation matrix r is symmetric and unity on the diagonal.*

		Ω_Λ	Γ	σ_8	n_s	z_0
2	Ω_m	−0.80	0.71	−0.94	−0.79	−0.98
	Ω_Λ		−0.15	0.56	0.28	0.90
	Γ			−0.90	−0.98	−0.57
	σ_8				0.95	0.87
	n_s					0.67
3	Ω_m	−0.81	−0.31	−0.84	0.35	−0.81
	Ω_Λ		0.66	0.41	−0.29	0.92
	Γ			−0.15	−0.69	0.63
	σ_8				−0.14	0.38
	n_s					−0.46
3d	Ω_m	−0.29	0.43	−0.81	−0.43	−0.24
	Ω_Λ		0.74	−0.33	−0.74	1.00
	Γ			−0.88	−1.00	0.77
	σ_8				0.87	−0.37
	n_s					−0.77
2+3d	Ω_m	−0.99	0.98	−1.00	−0.97	−1.00
	Ω_Λ		−0.94	0.98	0.96	0.97
	Γ			−0.99	−0.98	−0.99
	σ_8				0.97	1.00
	n_s					0.96
2+3	Ω_m	−0.87	0.46	−0.99	−0.12	−0.92
	Ω_Λ		−0.20	0.83	0.24	0.65
	Γ			−0.55	−0.66	−0.52
	σ_8				0.17	0.90
	n_s					−0.04

Interesting constraints on cosmological parameters like Ω_m and σ_8 have been obtained from second-order cosmic shear statistics. The bispectrum of the density fluctuations contains complementary information about structure evolution and cosmology. It is a measure of the non-Gaussianity of the large-scale structure. Current cosmic shear surveys are at the detection limit of measuring non-Gaussianity significantly, and future observations will most certainly determine the bispectrum with high accuracy.

Combined measurements of the power and the bispectrum yield additional constraints on cosmological parameters and partially lift degeneracies between parameters. The second- and generalized third-order aperture mass statistics $\langle M_\mathrm{ap}^2\rangle$ and $\langle M_\mathrm{ap}^3\rangle$ are local measures of the power and bispectrum, respectively. In this part of the thesis, cosmological parameter forecast from combined measurements of these two shear statistics is presented. Using ΛCDM ray-tracing simulations, the covariance of M_ap^2 and M_ap^3 and their cross-correlation are calculated. An extensive Fisher matrix analysis is performed and minimum variance bounds (MBVs) are

obtained for a variety of combinations of cosmological parameters.

The generalized third-order aperture mass statistics (Sect. 2.3.2) is the correlator of $M_{\rm ap}$ for three different aperture radii. In contrast to the skewness of $M_{\rm ap}$ which probes the bispectrum for equilateral triangles only, the generalized third-order aperture mass is in principle sensitive to the bispectrum on the complete ℓ-space. Therefore, it contains much more information about cosmology than the skewness alone.

The direction of degeneracy between cosmological parameters are similar for second- and third-order statistics. However, in most cases the combination of $\langle M_{\rm ap}^2\rangle$ and $\langle M_{\rm ap}^3\rangle$ gives substantial improvement on the predicted parameter constraints. The MVBs decrease by a factor of two or more for most of the parameter combinations. When the source redshift z_0 is not fixed but also to be determined from the data, the errors on the other parameters increase largely and the improvement by combining $\langle M_{\rm ap}^2\rangle$ and $\langle M_{\rm ap}^3\rangle$ is lowered.

I combine the second-order aperture mass statistics $\langle M_{\rm ap}^2\rangle$ both with the skewness and the generalized third-order aperture mass. The latter combination gives much better parameter constraints than the first one. For six parameters to be determined from the data simultaneously, the corresponding MVBs are better by a factor of about 10 for each parameter.

The $\Omega_{\rm m}$-σ_8-degeneracy is very prominent for both the second- and the third-order statistics of $M_{\rm ap}$ individually. However, by combining the two, the degeneracy is partially lifted — the 1σ-errors of both parameters drop by a factor of two or more, depending on which other parameters are also considered to be free. The Γ-$n_{\rm s}$-degeneracy, however, cannot be broken by combining $\langle M_{\rm ap}^2\rangle$ and $\langle M_{\rm ap}^3\rangle$, the determination of this pair of parameters is dominated by $\langle M_{\rm ap}^3\rangle$.

If the range of apertures is extended, would one expect the resulting improvement on the parameter estimation from the third-order aperture mass statistics to be higher than from second-order? For the former, the number of data points increases with the third power of the number of aperture radii, whereas for the latter, the increase is only linear. Thus, for an increase in the number of measured apertures, the constraints using $\langle M_{\rm ap}^3\rangle$ should improve more than those from $\langle M_{\rm ap}^2\rangle$. On the other hand, the data points are not at all uncorrelated, in fact as is shown in this work (Sect. 5.2.3) the correlation can be very high for various combinations of aperture radius triples. Moreover for large scales ($\theta \gtrsim 30'$, see Fig. 4.7), the linear regime of the large-scale structure is probed, non-Gaussian contributions are small, and the information content for the third-order shear statistics is not very important.

Summary

In this thesis I study the efficacy of cosmic shear measurements to constrain cosmological parameters. The first chapter shortly reviews the current standard model of cosmology and the parameters describing this model. The theory of structure formation in the Universe is introduced and the phenomenum of gravitational lensing, in particular weak lensing by large-scale matter inhomogeneities is elucidated. In Chapter 2, various second- and third-order statistics of cosmic shear are introduced together with estimators and their covariance. These quantities are used in the following chapters for the study of cosmological parameter determination from cosmic shear measurements. In Chapter 3, I examine the influence of cosmic shear survey characteristics on the resulting parameter errors. The properties of third-order cosmic shear statistics and their importance for cosmological measurements are presented in Chapter 4. Finally in Chapter 5, second- and third-order aperture mass statistics of cosmic shear are combined with the result that error bars and degeneracies for various cosmological parameters are largely reduced.

Weak lensing by the large-scale structure in the Universe has become an important cosmological tool and allows one to measure cosmological parameters independently and complementary to other cosmology experiments such as Cosmic Microwave Background anisotropy probes, Lyman-α or galaxy surveys. Future cosmic shear surveys will provide high-precision measurements of the large-scale distribution of matter in the Universe on linear and non-linear scales and yield tight constraints on various cosmological parameters.

Since weak lensing observations put a very high demand on atmospheric conditions, instrumental requirements and are very time-consuming, the current end even next-generation cosmic shear surveys will be able to cover only a couple of percent of the complete sky. One therefore has to carefully choose a strategy of distributing the lines of sight and find a trade-off between the following opposing scenarios. On the one hand, a large range of angular scales has to be covered by sampling a compact sky region with a fair number of individual lines of sight. On the other hand, one needs a considerable number of independent directions on the sky in order to minimize the sampling or cosmic variance. Apart from these considerations, bright stars and foreground galaxies inevitably render parts in the data as unusable causing an imaging survey to have a complicated geometry.

I examine the effect of the survey geometry on the accuracy of cosmological parameter determination in Chapter 3. I consider surveys consisting of a varying number of independent circular regions on the sky ("patches") which are sparsely sampled by a different number of lines of sight. These surveys are compared with ones consisting of single, uncorrelated fields of view. The latter have a very small cosmic variance but do not sample the shear correlation on scales larger than the individual images. From Monte-Carlo simulations of various survey geometries, I obtain the covariance of several second-order statistics encoding the measurement errors and correlation between different angular scales for a particular survey setting. Using a Fisher matrix analysis, I estimate the minimum variance bounds for various combinations of cosmological parameters such as the matter density $\Omega_{\rm m}$, the power spectrum

normalization σ_8, the shape parameter Γ, the spectral index of the initial fluctuations n_s, the cosmological constant Ω_Λ, and the characteristic source galaxy redshift z_0. Using the aperture mass dispersion $\langle M_\mathrm{ap}^2 \rangle$, the patch geometries yield smaller errors for the parameters indicating that information on medium and large angular scales are more important than a small cosmic variance. The result of this study is that by appropriately choosing the survey geometry a 25 percent improvement for the 1σ-error is possible.

In addition, a Karhunen-Loève (KL) eigenmode analysis is performed to study various estimators of second-order shear statistics and the influence of the survey geometry and various noise contributions to these estimators. The KL technique allows the range and combination of angular scales which contain most of the information about cosmological parameters to be quantified. Accordingly, although medium scales where the shot noise due to the intrinsic ellipticities and finite number of galaxies is small contribute most to the signal, all scales which are probed by the survey carry valuable information. The KL analysis shows that the second-order shear estimators can be linearly combined into two modes which essentially carry all cosmological information. Furthermore, the information contained in the covariance of the shear estimators is subdominant for survey areas larger than about half a square degree.

The accuracy of cosmological parameters, measured from second-order cosmic shear statistics, suffers from near-degeneracies between various parameters which leads to large error bars. It has been shown that the reduced skewness of the convergence κ, which is the non-linear combination of second- and third-order moments of κ, can break the most prominent near-degeneracy between Ω_m and σ_8 (Bernardeau et al. 1997; van Waerbeke et al. 1999). Until recently, the survey size and accuracy of the galaxy shape determination allowed for significant measurements only of second-order statistics. However, the first detections of a third-order shear signal have been reported (Bernardeau et al. 2002; Jarvis et al. 2004; Pen et al. 2003) and the next-generation surveys will most certainly measure non-Gaussian features in the shear field with high accuracy.

The power spectrum of the matter density contrast on large scales has been extensively modeled using numerical simulations and consequently, second-order statistics of cosmic shear which probe the projected power spectrum have been studied in great detail using theoretical predictions. On the other hand, the bispectrum of the cosmological dark matter distribution is less securely known and until now, theoretical predictions of third-order cosmic shear statistics have been scarce. In Chapter 4, the three-point correlation function and the third-order aperture mass statistics, which are both functions of the projected bispectrum, are calculated from theoretical non-linear models of structure formation. These predictions are tested and compared with ray-tracing simulations. The dependence of the third-order aperture mass statistics with respect to cosmological parameters is discussed and a two-dimensional visualization of the shear three-point correlation function is presented.

In Chapter 5, I show that the linear combination of second- and third-order aperture mass statistics, $\langle M_\mathrm{ap}^2 \rangle$ and $\langle M_\mathrm{ap}^3 \rangle$, respectively, highly reduces near-degeneracies between several cosmological parameters and improves the resulting error bars significantly. It is demonstrated that the generalized third-order aperture mass statistics contains much more information about cosmology than the skewness of M_ap. From ΛCDM ray-tracing simulations of the shear and the convergence, the covariance of M_ap^2, M_ap^3 and their cross-correlation are estimated which contain non-Gaussian contributions originating from the non-linear evolution of the large-scale structure. A Fisher matrix analysis quantifies error bars of and degeneracies between the cosmological parameters $\Omega_\mathrm{m}, \Gamma, \sigma_8, n_\mathrm{s}, \Omega_\Lambda$ and the source galaxy redshift z_0. For

most parameter pairs, their degeneracies for $\langle M_{\rm ap}^2 \rangle$ and $\langle M_{\rm ap}^3 \rangle$ are similar. However, the combination of the two statistics results in an improvement of the error bars by a factor of two or more. If all six parameters are determined simultaneously from the data, the corresponding error bars are reduced by a factor of 10 on combining $\langle M_{\rm ap}^2 \rangle$ and $\langle M_{\rm ap}^3 \rangle$. While the parameters Γ and $n_{\rm s}$ remain highly degenerate, the $\Omega_{\rm m}$-σ_8-degeneracy is lifted substantially by combining the second and generalized third moment of $M_{\rm ap}$ — the error bars of the two parameters drop by a factor of two to three if the source redshift z_0 is known.

Appendix A

Random fields

To deal with cosmological fields describable only in a statistical way, such as the large-scale matter distribution in the Universe or the cosmic shear field given in some sky region, the use of random fields is necessary. In this section, random fields and their properties, adapted to the needs of this work, are issued. For a strict mathematical treatment of random fields see e.g. Adler (1981).

A random field F is a measurable set of functions $F = \{g : \mathbb{R}^n \to \mathbb{R}\}$, where an element $g \in F$ is called a *realization* of the random field. It is further equipped by joint probability distribution functions

$$P\left[g(\boldsymbol{x}_1), g(\boldsymbol{x}_2), \ldots, g(\boldsymbol{x}_N)\right] \mathrm{d}g(\boldsymbol{x}_1)\mathrm{d}g(\boldsymbol{x}_2)\ldots\mathrm{d}g(\boldsymbol{x}_N) \tag{A.1}$$

for all $N \in \mathbb{N}$. $P[g(\boldsymbol{x})]\mathrm{d}g(\boldsymbol{x})$ is the probability for an element of F to take the value $g(\boldsymbol{x})$ at the point $\boldsymbol{x}$, $P[g(\boldsymbol{x}_1)g(\boldsymbol{x}_2)]\mathrm{d}g(\boldsymbol{x}_1)\mathrm{d}g(\boldsymbol{x}_2)$ gives the joint probability for having the values $g(\boldsymbol{x}_1)$ at $\boldsymbol{x}_1$ and $g(\boldsymbol{x}_2)$ at $\boldsymbol{x}_2$, etc.

The first moment or *mean* of the random field is defined as

$$\langle g(\boldsymbol{x})\rangle = \int g(\boldsymbol{x})P[g(\boldsymbol{x})]\,\mathrm{d}g(\boldsymbol{x}), \tag{A.2}$$

where the integration is carried out over the set F. This averaging process, indicated by the angle brackets, are called *ensemble average*. A random field g' with zero mean can be constructed from any random field g by setting $g'(\boldsymbol{x}) \equiv g(\boldsymbol{x}) - \langle g(\boldsymbol{x})\rangle$. From now on, only fields with zero mean will be considered.

The second moment or *two-point correlation function* (2PCF) of a random field is

$$\xi(\boldsymbol{x}_1, \boldsymbol{x}_2) = \langle g(\boldsymbol{x}_1)g(\boldsymbol{x}_2)\rangle = \int\int g(\boldsymbol{x}_1)g(\boldsymbol{x}_2)P[g(\boldsymbol{x}_1), g(\boldsymbol{x}_2)]\,\mathrm{d}g(\boldsymbol{x}_1)\mathrm{d}g(\boldsymbol{x}_2). \tag{A.3}$$

Analogously, the *three-point correlation function* (3PCF) $\zeta(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3) = \langle g(\boldsymbol{x}_1)g(\boldsymbol{x}_2)g(\boldsymbol{x}_3)\rangle$ and for every $N \in \mathbb{N}$ the *N-point correlation function* is defined. To fully characterize a random field, all N-point correlation functions are needed in general.

For a cosmological experiment there usually exists only one observable realization g, in particular because there is only one observable Universe. The determination of the properties of the underlying random field F from a single realization g seems to be a hopeless task. However, instead of ensemble averaging one can average over a sufficient large number of spatial points. The assumption that these two averaging processes are equivalent is the *ergodicity* hypothesis. For example, the large-scale structure is sampled for many different, statistically independent directions on the sky. The sampling variance of this spatial averaging process is called *cosmic variance*.

A.1 Homogeneous and isotropic random fields

A random field is called *homogeneous* if all joint probability distribution functions are invariant under simultaneous translation by the same vector. In this case, the 2PCF only depends on the difference of the two points, $\xi(\boldsymbol{x}_1, \boldsymbol{x}_2) = \xi(\boldsymbol{x}_1 - \boldsymbol{x}_2)$. Likewise, the 3PCF depends only on two vector differences an can be written as e.g. $\zeta(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3) = \zeta(\boldsymbol{x}_1 - \boldsymbol{x}_2, \boldsymbol{x}_1 - \boldsymbol{x}_3)$. A homogeneous random field is called *isotropic*, if all joint probability distribution functions are invariant under simultaneous rotation by the same rotation matrix. In particular, the 2PCF ξ is rotational invariant, $\xi(\boldsymbol{x}) = \xi(x)$ and the 3PCF only depends on three scalars which uniquely describe a triangle.

The study of the properties of homogeneous and isotropic random fields in Fourier space is very useful. For the function $g : \mathbb{R}^n \to \mathbb{R}$, its Fourier transform is the function $\hat{g} : \mathbb{R}^n \to \mathbb{C}$, related to g by the following two expressions:

$$\hat{g}(\boldsymbol{k}) = \int_{\mathbb{R}^n} \mathrm{d}^n x \, g(\boldsymbol{x}) \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}, \qquad g(\boldsymbol{x}) = \int_{\mathbb{R}^n} \frac{\mathrm{d}^n k}{(2\pi)^n} \hat{g}(\boldsymbol{k}) \mathrm{e}^{-\mathrm{i}\boldsymbol{x}\cdot\boldsymbol{k}}. \tag{A.4}$$

The 2PCF of the random field F in Fourier space is then given by

$$\left\langle g(\boldsymbol{k}) g^*(\boldsymbol{k}') \right\rangle = \int_{\mathbb{R}^n} \mathrm{d}^n x \, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\boldsymbol{x}} \int_{\mathbb{R}^n} \mathrm{d}^n x' \, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}'\boldsymbol{x}'} \left\langle g(\boldsymbol{x}) g(\boldsymbol{x}') \right\rangle = (2\pi)^n \delta_{\mathrm{D}}(\boldsymbol{k} - \boldsymbol{k}') P_g(k), \tag{A.5}$$

where δ_{D} is the Dirac delta "function" and P_g is the Fourier transform of the 2PCF ξ or the *power spectrum* of g,

$$P_g(k) = \int_{\mathbb{R}^n} \mathrm{d}^n y \, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\boldsymbol{y}} \xi(y). \tag{A.6}$$

Note that here and throughout this thesis, ergodicity is assumed and the random field F can be identified with one of its realizations g. Equation (A.5) states that pairwise different Fourier modes are uncorrelated which is a direct consequence of the translational invariance. Because of isotropy, the power spectrum only depends on the modulus of the wave vector.

Similarly to (A.5), the three-point correlation function can be written in Fourier space,

$$\begin{aligned} \langle g(\boldsymbol{k}_1) g(\boldsymbol{k}_2) g(\boldsymbol{k}_3) \rangle &= \int_{\mathbb{R}^n} \mathrm{d}^n x_1 \int_{\mathbb{R}^n} \mathrm{d}^n x_2 \int_{\mathbb{R}^n} \mathrm{d}^n x_3 \, \mathrm{e}^{\mathrm{i}(\boldsymbol{k}_1\boldsymbol{x}_1 + \boldsymbol{k}_2\boldsymbol{x}_2 + \boldsymbol{k}_3\boldsymbol{x}_3)} \langle g(\boldsymbol{x}_1) g(\boldsymbol{x}_2) g(\boldsymbol{x}_3) \rangle \\ &= (2\pi)^n \delta_{\mathrm{D}}(\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3) \Big[B_g(\boldsymbol{k}_1, \boldsymbol{k}_2) + B_g(\boldsymbol{k}_2, \boldsymbol{k}_3) + B_g(\boldsymbol{k}_3, \boldsymbol{k}_1) \Big], \end{aligned} \tag{A.7}$$

where B_g is the Fourier transform of 3PCF ζ or the *bispectrum* of g:

$$B_g(\boldsymbol{k}_1, \boldsymbol{k}_2) = \int_{\mathbb{R}} \mathrm{d}^n y_1 \int_{\mathbb{R}} \mathrm{d}^n y_2 \, \zeta(\boldsymbol{y}_1, \boldsymbol{y}_2) \, \mathrm{e}^{\mathrm{i}(\boldsymbol{y}_2\boldsymbol{k}_1 - \boldsymbol{y}_1\boldsymbol{k}_2)}. \tag{A.8}$$

As can be seen in eq. (A.7), homogeneity implies that only Fourier modes which form a closed triangle show a correlation.

A.2 Gaussian random fields

A random field is called *Gaussian* if its joint probability distribution functions are multivariate Gaussian distributions,

$$P\left[g_1, g_2, \ldots, g_N\right] \mathrm{d}g_1 \mathrm{d}g_2 \ldots \mathrm{d}g_N = \frac{\exp\left[-\frac{1}{2}\sum_{ij} g_i C_{ij}^{-1} g_j)\right]}{\left[(2\pi)^N \det \mathbf{C}\right]^{1/2}} \mathrm{d}g_1 \mathrm{d}g_2 \ldots \mathrm{d}g_N, \tag{A.9}$$

where $g_i = g(\boldsymbol{x}_i)$ and $C_{ij} \equiv \langle g_i g_j \rangle$ is the covariance matrix of g. For a Gaussian random field, all connected moments higher than second are zero and therefore, all odd moments vanish and all even moments can be written in terms of the second moment. Thus, a Gaussian random field is fully characterized by the two-point correlation function or equivalently, the power spectrum. Another important property is that for a Gaussian field the ergodicity hypotheses holds if and only if the power spectrum is a continuous function (Adler 1981).

Appendix B

Fiducial cosmological models

In this work several CDM cosmologies are used as fiducial models, the parameters of which are listed in Table B.1. The models 1 and 2 are ΛCDM cosmologies with a non-zero cosmological constant, model 3 represents an open Universe (OCDM) with $\Lambda = 0$. The parameters are the matter density parameter Ω_m, the density parameter of the cosmological constant Ω_Λ, the shape parameter Γ, the normalization σ_8 and the initial spectral index n_s of the power spectrum. See Sects. 1.1.5, 1.2.5 and 1.2.6.1 for more details on the parameters. Models 2 and 3 match the parameters of the N-body and inferred ray-tracing simulations.

For the models 1 and 3, the shape parameter has the form $\Gamma = \Omega_\mathrm{m} h$, see eq. (1.38). In model 2, Sugiyama's Γ (eq. 1.39) is assumed (Sugiyama 1995) — the numerical value of Γ corresponds to $h = 0.7$ and $\Omega_\mathrm{b} = 0.04$. Only for model 1, a continuous redshift distribution of the source galaxies (1.90) is chosen which is characterized by the two parameters z_0 and β. The ray-tracing simulations and therefore the models 2 and 3 have a single source redshift plane at z_0 around unity.

Table B.1: *Parameters of the cosmological models used in this work.*

	model 1	model 2	model 3
Cosmological parameters			
Ω_m	0.3	0.3	0.3
Ω_Λ	0.7	0.7	0.0
Γ	0.21	0.172	0.21
σ_8	1	0.9	1
n_s	1	1	1
Redshift distribution			
z_0	1	0.977	1
β	1.5	-	-
Ray-tracing parameter			
pixel number	-	1024	2048
pixel size	-	$0.2'$	$0.0949'$
number of realizations	-	36	7

Appendix C

Limber's equation for the bispectrum

The relation between a homogeneous and isotropic three-dimensional random field (e.g. the density contrast δ) and its projection onto two dimension (e.g. the convergence κ) in a cosmological context can be obtained using Limber's equation (Limber 1953). Applied in Fourier space, this equation yields the projection of the power spectrum of the random field (Kaiser 1992, 1998; Bartelmann & Schneider 2001). In this section the projection of the three-dimensional bispectrum is derived using Limber's equation. The result of this calculation is used for the expression of the convergence bispectrum (eq. 1.94).

In analogy to Bartelmann & Schneider (2001), three projections of the density contrast δ are defined,

$$g_i(\boldsymbol{\theta}) = \int_0^{w_\mathrm{lim}} \mathrm{d}w\, q_i(w)\, \delta\left[f_K(w)\boldsymbol{\theta}, w\right], \quad i = 1, 2, 3, \tag{C.1}$$

where the q_i are weight functions, $\boldsymbol{\theta}$ is a two-dimensional vector and w_lim is the limiting comoving distance of the projection. The three-point correlation function (3PCF) of the g_i is

$$\begin{aligned}
\zeta_g(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3) &= \langle g_1(\boldsymbol{\theta}_1) g_2(\boldsymbol{\theta}_2) g_3(\boldsymbol{\theta}_3)\rangle \\
&= \int \mathrm{d}w_1\, q_1(w) \int \mathrm{d}w_2\, q_2(w) \int \mathrm{d}w_3\, q_3(w)\, \langle \delta[f_K(w)\boldsymbol{\theta}_1, w]\delta[f_K(w)\boldsymbol{\theta}_2, w]\delta[f_K(w)\boldsymbol{\theta}_3, w]\rangle \\
&= \int \mathrm{d}w_1\, q_1(w) \int \mathrm{d}w_2\, q_2(w) \int \mathrm{d}w_3\, q_3(w) \int \frac{\mathrm{d}^3k_1}{(2\pi)^3} \int \frac{\mathrm{d}^3k_2}{(2\pi)^3} \int \frac{\mathrm{d}^3k_3}{(2\pi)^3} \\
&\quad \times \left\langle \hat{\delta}(\boldsymbol{k}_1, w_1)\hat{\delta}(\boldsymbol{k}_2, w_2)\hat{\delta}(\boldsymbol{k}_3, w_3)\right\rangle \mathrm{e}^{-\mathrm{i}f_K(w_1)\boldsymbol{\theta}_1\cdot\boldsymbol{k}_{1\perp}} \mathrm{e}^{-\mathrm{i}f_K(w_2)\boldsymbol{\theta}_2\cdot\boldsymbol{k}_{2\perp}} \\
&\quad \times \mathrm{e}^{-\mathrm{i}f_K(w_3)\boldsymbol{\theta}_3\cdot\boldsymbol{k}_{3\perp}} \mathrm{e}^{-\mathrm{i}w_1 k_{13}} \mathrm{e}^{-\mathrm{i}w_2 k_{23}} \mathrm{e}^{-\mathrm{i}w_3 k_{23}},
\end{aligned} \tag{C.2}$$

where the Fourier transform is applied in the last step and the three-dimensional Fourier mode vectors are decomposed into a two-dimensional perpendicular and a one-dimensional parallel component with respect to the projection direction, $\boldsymbol{k}_i = (\boldsymbol{k}_{i\perp}, k_{i3})$.

Following Bartelmann & Schneider (2001), I assume that no power on scales larger than some coherence scale exists and further, that the density contrast δ as well as the weight functions q_i do not vary appreciably over this coherence scale. Consequently, one sets $f_K(w_1) = f_K(w_2) = f_K(w_3) =: f_K(w)$ and $q_1(w_1) = q_1(w), q_2(w_2) = q_2(w), q_3(w_3) = q_3(w)$. Next, the bispectrum B_δ of the density contrast (1.42) is inserted and the k_3-integral is carried out making use of the delta function. One gets

$$\begin{aligned}
\zeta_g(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3) = \int \mathrm{d}w\, q_1(w) q_2(w) q_3(w) \int \frac{\mathrm{d}^3k_1}{(2\pi)^3} \int \frac{\mathrm{d}^3k_2}{(2\pi)^3} \Big[& B_\delta(\boldsymbol{k}_1, \boldsymbol{k}_2, w) \\
& + B_\delta(\boldsymbol{k}_2, -\boldsymbol{k}_1 - \boldsymbol{k}_2, w) + B_\delta(-\boldsymbol{k}_1 - \boldsymbol{k}_2, \boldsymbol{k}_1, w)\Big] \\
& \mathrm{e}^{-\mathrm{i}f_K(w)\boldsymbol{\theta}_1\cdot\boldsymbol{k}_{1\perp}} \mathrm{e}^{-\mathrm{i}f_K(w)\boldsymbol{\theta}_2\cdot\boldsymbol{k}_{2\perp}}
\end{aligned}$$

$$\times\, \mathrm{e}^{-\mathrm{i}f_K(w)\boldsymbol{\theta}_3\cdot(-\boldsymbol{k}_{1\perp}-\boldsymbol{k}_{2\perp})}\mathrm{e}^{-\mathrm{i}wk_{13}}\int \mathrm{d}w_2\,\mathrm{e}^{-\mathrm{i}w_2k_{23}}\int \mathrm{d}w_3\,\mathrm{e}^{-\mathrm{i}w_3(-k_{13}-k_{23})}\,. \tag{C.3}$$

The trivial w_2- and w_3-integrations yield $(2\pi)^2\delta_D(k_{13})\delta_D(k_{23})$, canceling out all Fourier modes which are not perpendicular to the projection direction (Blandford et al. 1991),

$$\begin{aligned}\zeta_g(\boldsymbol{\theta}_1,\boldsymbol{\theta}_2,\boldsymbol{\theta}_3) = &\int \mathrm{d}w\, q_1(w)q_2(w)q_3(w)\int\frac{\mathrm{d}^2k_{1\perp}}{(2\pi)^2}\int\frac{\mathrm{d}^2k_{2\perp}}{(2\pi)^2}\\ &\times\Big[B_\delta(\boldsymbol{k}_{1\perp},\boldsymbol{k}_{2\perp},w)+2\text{ terms}\Big]\mathrm{e}^{-\mathrm{i}f_K(w)(\boldsymbol{\theta}_1-\boldsymbol{\theta}_3)\cdot\boldsymbol{k}_{1\perp}}\mathrm{e}^{-\mathrm{i}f_K(w)(\boldsymbol{\theta}_2-\boldsymbol{\theta}_3)\cdot\boldsymbol{k}_{2\perp}}.\end{aligned} \tag{C.4}$$

Because of homogeneity, the 3PCF only depends on the two vector differences $\boldsymbol{\vartheta}_1=\boldsymbol{\theta}_1-\boldsymbol{\theta}_3$ and $\boldsymbol{\vartheta}_2=\boldsymbol{\theta}_2-\boldsymbol{\theta}_3$. Inserting one of the three terms of the last equation into the expression for the bispectrum of g (eq. A.8), one gets

$$\begin{aligned}B_g(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2) = &\int \mathrm{d}w\, q_1(w)q_2(w)q_3(w)\int\frac{\mathrm{d}^2k_1}{(2\pi)^2}\int\frac{\mathrm{d}^2k_2}{(2\pi)^2}\\ &\times B_\delta(\boldsymbol{k}_1,\boldsymbol{k}_2)\int\mathrm{d}^2\vartheta_1\,\mathrm{e}^{\mathrm{i}(f_K(w)\boldsymbol{k}_2-\boldsymbol{\ell}_2)\cdot\boldsymbol{\vartheta}_1}\int\mathrm{d}^2\vartheta_2\,\mathrm{e}^{-\mathrm{i}(f_K(w)\boldsymbol{k}_1-\boldsymbol{\ell}_1)\cdot\boldsymbol{\vartheta}_2},\end{aligned} \tag{C.5}$$

where the subscripts indicating the perpendicular components of the wave vectors have been dropped. The ϑ_1-integral yields $(2\pi)^2\delta_\mathrm{D}(f_K(w)\boldsymbol{k}_2-\boldsymbol{\ell}_2)=(2\pi)^2/f_K^2(w)\cdot\delta_\mathrm{D}(\boldsymbol{k}_2-\boldsymbol{\ell}_2/f_K(w))$; another delta-function emerges from the ϑ_2-integration. Finally,

$$B_g(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2)=\int\frac{\mathrm{d}w}{f_K^4(w)}q_1(w)q_2(w)q_3(w)B_\delta\left(\frac{\boldsymbol{\ell}_1}{f_K(w)},\frac{\boldsymbol{\ell}_2}{f_K(w)}\right). \tag{C.6}$$

In the case of the bispectrum of the convergence κ, the expression

$$q_i(w)=\frac{3}{2}\left(\frac{H_0}{c}\right)^2\frac{\Omega_\mathrm{m}}{a(w)}\int_w^{w_\mathrm{lim}}\mathrm{d}w'\,p(w')\frac{f_K(w-w')}{f_K(w')}f_K(w)=G(w)f_K(w) \tag{C.7}$$

for $i=1,2,3$ has to be taken. With that, one gets

$$B_\kappa(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2)=\int_0^{w_\mathrm{lim}}\frac{\mathrm{d}w}{f_K(w)}G^3(w)B_\delta\left(\frac{\boldsymbol{\ell}_1}{f_K(w)},\frac{\boldsymbol{\ell}_2}{f_K(w)}\right). \tag{C.8}$$

Bibliography

Adler, R. J. 1981, The Geometry of Random Fields (Wiley) 125, 127

Bacon, D. J., Réfrégier, A. R., & Ellis, R. S. 2000, MNRAS, 318, 625 30

Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986, ApJ, 304, 15 16

Bartelmann, M. & Schneider, P. 2001, Phys. Rep., 340, 297 21, 30, 131

Bennett, C. L., Halpern, M., Hinshaw, G., et al. 2003, ApJS, 148, 1 9

Bernardeau, F., Colombi, S., Gaztañaga, E., & Scoccimarro, R. 2002, Phys. Rep., 367, 1 13, 15

Bernardeau, F., Mellier, Y., & van Waerbeke, L. 2002, A&A, 389, L28 102, 122

Bernardeau, F., van Waerbeke, L., & Mellier, Y. 1997, A&A, 322, 1 30, 68, 100, 122

Bernardeau, F., van Waerbeke, L., & Mellier, Y. 2003, A&A, 397, 405 102

Blandford, R. D., Saust, A. B., Brainerd, T. G., & Villumsen, J. V. 1991, MNRAS, 251, 600 132

Bond, J. R. 1995, Physical Review Letters, 74, 4369 70

Bradač, M., Lombardi, M., & Schneider, P. 2004, A&A, 424, 13 29

Brainerd, T. G., Blandford, R. D., & Smail, I. 1996, ApJ, 466, 623 32

Brown, M. L., Taylor, A. N., Hambly, N. C., & Dye, S. 2002, MNRAS, 333, 501 36

Catelan, P., Kamionkowski, M., & Blandford, R. D. 2001, MNRAS, 320, L7 36

Contaldi, C. R., Hoekstra, H., & Lewis, A. 2003, Physical Review Letters, 90, 221303/1 30

Crittenden, R. G., Natarajan, P., Pen, U.-L., & Theuns, T. 2001, ApJ, 559, 552 36

Crittenden, R. G., Natarajan, P., Pen, U.-L., & Theuns, T. 2002, ApJ, 568, 20 39, 41, 44

Croft, R. A. C. & Metzler, C. A. 2000, ApJ, 545, 561 36

Dye, S., Taylor, A. N., Greve, T. R., et al. 2002, A&A, 386, 12 29

Einstein, A. 1916, Annalen der Physik, 49, 769 3

Eisenstein, D. J. & Hu, W. 1998, ApJ, 496, 605 17

Firmani, C., D'Onghia, E., Chincarini, G., Hernández, X., & Avila-Reese, V. 2001, MNRAS, 321, 713 10

Friedmann, A. 1922, Z. Phys., 10, 377 5

Friedmann, A. 1924, Z. Phys., 21, 326 5

Hamana, T., Miyazaki, S., Shimasaku, K., et al. 2003, ApJ, 597, 98 41

Hamilton, A., Kumar, P., Lu, E., & Matthews, A. 1991, ApJ, 374, L1 19

Hamilton, A. J. S., Tegmark, M., & Padmanabhan, N. 2000, MNRAS, 317, L23 70

Heavens, A. 2003, MNRAS, 343, 1327 70, 72

Heavens, A., Réfrégier, A., & Heymans, C. 2000, MNRAS, 319, 649 36

Hirata, C. M., Mandelbaum, R., Seljak, U., et al. 2004, MNRAS, 353, 529 36

Hirata, C. M. & Seljak, U. 2004, Phys. Rev. D, 70, 063526 56

Hoekstra, H., Yee, H. K. C., & Gladders, M. D. 2002, ApJ, 577, 595 30, 41

Hogg, D. W., Eisenstein, D. J., Blanton, M. R., et al. 2005, ApJ, 624, 54 4

Hu, W. & Tegmark, M. 1999, ApJ, 514, L65 30

Huterer, D. & White, M. 2005, Phys. Rev. D, 72, 043002 70

Jain, B., Seljak, U., & White, S. 2000, A&A, 530, 547 32

Jarvis, M., Bernstein, G., Fischer, P., & Smith, D. 2003, AJ, 125, 1014 30, 41

Jarvis, M., Bernstein, G., & Jain, B. 2004, MNRAS, 352, 338 49, 52, 91, 107, 122

Jenkins, A., Frenk, C. S., Pearce, F. R., et al. 1998, ApJ, 499, 20 22

Jing, Y. P. 2002, MNRAS, 335, L89 36

Kaiser, N. 1992, ApJ, 388, 272 33, 131

Kaiser, N. 1995, ApJ, 439, L1 35

Kaiser, N. 1998, ApJ, 498, 26 131

Kaiser, N., Squires, G., Fahlman, G., & Woods, D. 1994, in Clusters of galaxies, Proceedings of the XIVth Moriond Astrophysics Meeting, Méribel, France, 269, also astro-ph/9407004 38

Kaiser, N., Wilson, G., & Luppino, G. 2000, astro-ph/0003338 30

Karhunen, K. 1946, Ann. Acad. Sci. Fennicae, 37 70

Kendall, M. G. & Stuart, A. 1969, The Advanced Theory of Statistics, Vol. II (London: Griffin) 65

Kilbinger, M. 2002, Master's thesis, Universität Bonn 64, 88

Kilbinger, M. & Munshi, D. 2005, MNRAS in press, also astro-ph/0509548 2

Kilbinger, M. & Schneider, P. 2004, A&A, 413, 465 2, 58, 63, 64, 67, 88

Kilbinger, M. & Schneider, P. 2005, A&A, 442, 69 2

King, L. & Schneider, P. 2002, A&A, 396, 411 36

King, L. & Schneider, P. 2003, A&A, 398, 23 36

Lemaître, G. 1927, Ann. Soc. Sci. Brux, A47, 49 5

Lemaître, G. 1931, MNRAS, 91, 483 5

Limber, D. N. 1953, ApJ, 117, 134 33, 131

Linde, A. 2005, New Astronomy Review, 49, 35 10

Loève, M. 1948, Processes Stochastiques et Mouvement Brownien (Hermann, Paris/France) 70

Ménard, B., Hamana, T., Bartelmann, M., & Yoshida, N. 2003, A&A, 403, 817 105

Netterfield, C. B. et al. 2002, ApJ, 571, 604 9

Pan, J. & Szapudi, I. 2005, submitted to MNRAS, also astro-ph/0505422 111

Peacock, J. A. & Dodds, S. J. 1994, MNRAS, 267, 1020 17, 19

Peacock, J. A. & Dodds, S. J. 1996, MNRAS, 280, L19 20, 99

Peebles, P. J. E. 1980, The Large-Scale Structure of the Universe (Princeton University Press) 11

Peebles, P. J. E. 1993, Principles of Physical Cosmology (Princeton University Press) 31

Pen, U.-L., Zhang, T., van Waerbeke, L., et al. 2003, ApJ, 592, 664 49, 122

Penzias, A. A. & Wilson, R. W. 1965, ApJ, 142, 419 8

Perlmutter, S. et al. 1999, ApJ, 517, 565 10

Press, W. H., Teukolsky, S. A., Flannery, B. P., & Vetterling, W. T. 1992, Numerical Recipes in C (Cambridge University Press) 63, 65, 72, 74

Ratnatunga, K. U., Ostrander, E. J., Griffiths, R. E., & Im, M. 1995, ApJ, 453, L5 25

Riess, A. G. et al. 1998, Astron. J., 116, 1009 10

Robertson, H. P. 1935, ApJ, 82, 284 4

Schirmer, M., Erben, T., Schneider, P., Wolf, C., & Meisenheimer, K. 2004, A&A, 420, 75 39

Schneider, P. 1996, MNRAS, 283, 837 38

Schneider, P. 2003, A&A, 408, 829 36, 49, 54, 108

Schneider, P. 2005, in Kochanek, C.S. and Schneider, P. and Wambsganss, J.: Gravitational Lensing: Strong, Weak & Micro. Lecture Notes of the 33rd Saas-Fee Advanced Course, G. Meylan, P. Jetzer & P. North (eds.) (Springer-Verlag: Berlin), 273 21

Schneider, P., Ehlers, J., & Falco, E. E. 1992, Gravitational Lenses (Springer) 21

Schneider, P., Kilbinger, M., & Lombardi, M. 2005, A&A, 431, 9 2, 44, 48, 49, 52, 53

Schneider, P. & Lombardi, M. 2003, A&A, 397, 809 44, 46

Schneider, P., van Waerbeke, L., Jain, B., & Kruse, G. 1998, MNRAS, 296, 873 36, 39, 41, 49, 100

Schneider, P., van Waerbeke, L., Kilbinger, M., & Mellier, Y. 2002, A&A, 396, 1 2, 57, 58, 60, 88

Schneider, P., van Waerbeke, L., & Mellier, Y. 2002, A&A, 389, 729 35, 36, 41, 43, 44

Scoccimarro, R. & Couchman, H. M. P. 2001, MNRAS, 325, 1312 20, 35, 99

Scoccimarro, R. & Frieman, J. A. 1999, ApJ, 520, 35 20

Scoccimarro, R., Zaldarriaga, M., & Hui, L. 1999, ApJ, 527, 1 58

Seljak, U., Makarov, A., McDonald, P., et al. 2005, Phys. Rev. D, 71, 103515 18

Simon, P. 2004, in Proceedings of "Baryons in Dark Matter Halos". Novigrad, Croatia, 5-9 Oct 2004. Editors: R. Dettmar, U. Klein, P. Salucci. Published by SISSA, 97– 18

Smoot, G. F., Bennett, C. L., Kogut, A., et al. 1991, ApJ, 371, L1 4

Spergel, D. N., Verde, L., Peiris, H. V., et al. 2003, ApJS, 148, 175 6, 7, 10

Stoehr, F., White, S. D. M., Springel, V., Tormen, G., & Yoshida, N. 2003, MNRAS, 345, 1313 10

Sugiyama, N. 1995, ApJS, 100, 281 17, 129

Szalay, A. S., Jain, B., Matsubara, T., et al. 2003, ApJ, 591, 1 70

Takada, M. & Jain, B. 2003, ApJ, 583, L49 44, 47

Takada, M. & Jain, B. 2004, MNRAS, 348, 897 52

Tegmark, M., Taylor, A., & Heavens, A. 1997, ApJ, 480, 22 65, 70

van Waerbeke, L. 2000, MNRAS, 313, 524 108

van Waerbeke, L., Bernardeau, F., & Mellier, Y. 1999, A&A, 342, 15 30, 122

van Waerbeke, L., Mellier, Y., Erben, T., et al. 2000, A&A, 358, 30 30

van Waerbeke, L., Mellier, Y., & Hoekstra, H. 2005, A&A, 429, 75 18, 30, 36, 41

van Waerbeke, L., Mellier, Y., Pelló, R., et al. 2002, A&A, 393, 369 30, 71

Walker, A. G. 1936, Proc. Lond. Math. Soc, 2, 90 4

Walsh, D., Carswell, R. F., & Weymann, R. J. 1979, Nature, 279, 381 21

Watkins, R., Feldman, H. A., Chambers, S. W., Gorman, P., & Melott, A. L. 2002, ApJ, 564, 534 70

White, M. & Hu, W. 2000, ApJ, 537, 1 32

Wittman, D. M., Tyson, J. A., Kirkman, D., Dell'Antonio, I., & Bernstein, G. 2000, Nature, 405, 143 30

Zaldarriaga, M. & Scoccimarro, R. 2003, ApJ, 584, 559 44

Zhang, L. L. & Pen, U.-L. 2005, New Astronomy, 10, 569 92

Acknowledgements

This thesis would not have been possible without the help and inspiration from many people. First of all, I want to thank Peter Schneider for giving me the opportunity to work with him and to write a thesis under his supervision. I profited profusely from our countless discussions. I am grateful to him for his support during the times I got stuck in my work and also for giving me the freedom to follow my line of thought independently.

I gained a lot from discussions with Lindsay King, Masahiro Takada, Patrick Simon, Peter Watts, Marco Lombardi, Thomas Erben, Oliver Czoske, Ludo van Waerbeke, Mike Jarvis, Jan Hartlap, Marco Hetterscheidt, Johannes Schmid-Burgk, Heino Falcke, Ismael Tereno, Francis Bernardeau, Karim Benabed and Brice Ménard. I am grateful to Takashi Hamana and Bhuvnesh Jain for kindly providing their ΛCDM ray-tracing simulations and Mike Jarvis for his tree-code algorithm for the 3PCF. A big thanks goes to all members of the lens group and the Sternwarte Bonn for creating a pleasant and warm atmosphere that made work enjoyable.

A substantial part of this thesis emerged out of a project with Dipak Munshi. It was a pleasure to start this fruitful collaboration that enabled us to exchange ideas.

During my PhD, I had the possibility to visit other institutes and give presentations on my work and I am particularly grateful for the invitations from Yannick Mellier (IAP Paris), Dipak Munshi (IoA Cambridge) and K. S. Dwarkanath (RRI Bangalore).

This thesis benefitted plenty from the diligent and tedious work of reading the manuscript and making corrections and suggestions by Peter Schneider, Rupal Mittal, Jasmin Pielorz and Peter Watts.

All computer-related issues and problems were quickly dealt with by Günther Lay, Oliver Cordes and Ole Marggraf. Kathy Schrüfer was irreplaceable in guiding me through the jungle of University administration. Thanks to all of you!

I would like to express my appreciation to the co-referee of my thesis, Johannes Schmid-Burgk. I thank Michael Kobel and Martin Rumpf for willing to participate in my exam as fachnahes Mitglied and fachangrenzendes Mitglied, respectively.

Diese Arbeit wurde unterstützt durch das Bundesministerium für Bildung und Forschung (BMBF) durch die Deutsche Forschungsgemeinschaft (DFG) unter dem Projekt SCHN 342/3-1, durch die Bonn International Graduate School (BIGS) und die International Max Planck Research School (IMPRS).