Theoretical Aspects of Cosmic Shear and its Ability to constrain Cosmological Parameters

DISSERTATION

zur
Erlangung des Doktorgrades (Dr. rer. nat)
der
Mathematisch–Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich–Wilhelms–Universität Bonn

vorgelegt von

TIM FREDERIK EIFLER
aus
Meisenheim

Bonn, 2008

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

1. Referent: Prof. Dr. Peter Schneider (Universität Bonn)

2. Referent: Prof. Dr. Yannick Mellier (Institut d'Astrophysique de Paris)

Tag der Promotion: 06.02.2009

Erscheinungsjahr: 2009

Diese Dissertation ist auf dem Hochschulschriftenserver der ULB Bonn unter http://hss.ulb.uni-bonn.de/diss_online elektronisch publiziert.

Contents

Chapter 1

Introduction

In the last 20 years cosmology has progressed from a data-starving to a prospering science branch. Today we know that the Universe was in a very dense state at early times and since then has expanded to its current size. The filamentary structure of galaxies and clusters we observe today originates from small perturbations in the primordial density field, which grow due to self-gravitation. The increasing number of cosmological experiments and the huge amount of high-quality data offer great possibilities to study the structure and evolution of our Universe. Cosmologists developed the so-called ΛCDM or "concordance" model, based on the assumptions of large-scale homogeneity and isotropy, which provides a robust explanation for cosmological observations and has predicted several of them. Within this picture the energy density of the Universe consists of 4 components, i.e. baryons ($\Omega_b \approx 5\%$), photons ($\Omega_r \approx 0.005\%$), cold dark matter ($\Omega_{cdm} \approx 20\%$), and dark energy ($\Omega_\Lambda \approx 75\%$), where the physical nature of the latter two is still an open question.

Various cosmological experiments address the issues of dark matter and dark energy; one of the most promising is weak gravitational lensing by the large-scale structure, also called cosmic shear. Light bundles emitted from distant galaxies travel through the Universe and are continuously deflected by the gravitational field of the inhomogeneous matter distribution. As a consequence the shapes of galaxy images are distorted and the statistical properties of these distortions reflect those of the matter density field.

In 2000, four different groups independently measured the first cosmic shear signal (Bacon et al. 2000; Kaiser et al. 2000; van Waerbeke et al. 2000; Wittman et al. 2000). Since then it has proven to be a very powerful method to constrain cosmological parameters, in particular the normalization of the power spectrum σ_8 in combination with Ω_m (e.g. Fu et al. 2008). The remarkable results cosmic shear has achieved so far will improve in the future with large upcoming surveys like Pan-STARRS, KIDS, DES, JDEM, Euclid or LSST. Soon the first two surveys, starting in 2009, will provide data which enables us to estimate the shear signal with less than 1% statistical error. These small statistical errors make cosmic shear an ideal tool for future dark energy studies (see Peacock et al. 2006; Albrecht et al. 2006).

Still, there are unsolved systematics and uncertainties which limit parameter estimation with cosmic shear. On the observational side an insufficient PSF-correction is the most important source of contamination. The (still ongoing) Shear TEsting Program (STEP) has significantly improved

on this issue (for latest results see Heymans et al. 2006; Massey et al. 2007a). Solutions to this problem are further discussed in Bridle et al. (2008). Furthermore, astrophysical contaminations must be removed, such as intrinsic alignment of source galaxies, which can be excluded in the signal if redshift information is available (King & Schneider 2003). Shape-shear correlation was predicted by Hirata & Seljak (2004) and detected by Mandelbaum et al. (2006). Recently, Joachimi & Schneider (2008) showed how to exclude the shape-shear contribution, again if accurate photometric redshifts are available. Besides excluding astrophysical contaminations, redshift information significantly increases constraints on cosmological parameters (Massey et al. 2007b); accurate redshifts will be of major importance for future cosmic shear surveys. For precision cosmology with cosmic shear not only the quality of the data and the removal of astrophysical contaminations is important. In addition, we have to improve on theoretical issues in lensing (e.g. reduced shear approximation, inaccuracy of Limber's equation or the flat-sky approximation) and cosmology (models for the non-linear power spectrum, non-Gaussian covariances, models for higher-order moments of the density/shear field). These improvements are important to obtain accurate predictions to which we compare the high-precision data.

Theoretial predictions and observational results meet each other in the likelihood analysis. If there remain significant problems on either side, the likelihood analysis will give biased results. In addition, the likelihood analysis itself is of great importance; even with precise data and accurate theoretical predictions the constraints on cosmological parameters can be biased by an improper likelihood analysis. Developing statistical tools which especially fit the needs of the considered cosmological probe(s) is therefore of great importance for the inference of cosmological interpretations.

The subject of my PhD thesis is to improve on theoretical concepts related to the estimation of cosmological parameters with cosmic shear. This involves the optimization of cosmic shear data vectors and accurate descriptions of the corresponding covariances as well as the statistical methods used in the likelihood analysis. Furthermore, I applied my theoretical work on the most recently developed cosmic shear measure, namely the ring statistics, to cosmic shear data. In my thesis I optimize the ring statistics' signal strength and employ it to constrain cosmological parameters using data from the Canada-France-Hawaii Telescope Legacy Survey (CFHTLS).

The thesis is structured as follows: In chapter 2 I start with a general introduction to cosmology and structure formation. Chapter 3 reviews the basic concepts of gravitational lensing with the focus on cosmic shear, its achievements, problems and future prospects. My research results are presented in the following chapters starting with an introduction to cosmic shear parameter estimation (chapter 4), in particular, I examine binning effects on cosmic shear covariances and how these effects influence the parameter constraints. The issue of covariances is further examined in the chapters 7 and 8. In chapter 7 I investigate the cosmology-dependence of cosmic shear covariances and develop improved methods for a likelihood analysis, which take the cosmology-dependence into account. A second problem in the derivation of covariances is the inclusion of the shear field's non-Gaussianity. In chapter 8 I examine the impact of non-Gaussianity on covariances and parameter constraints, in particular regarding dark energy parameters. In this chapter I also verify the calibration factor formalism introduced by Semboloni et al. (2007) to account for non-Gaussianity. Results of my research on cosmic shear data vectors are presented in chapter 5, where I compare and optimize cosmic shear data vectors with respect to their in-

formation content and robustness against contaminations, and in chapter 6. The latter chapter contains my personal highlight of this thesis, namely the constraints on σ_8 using CFHTLS data. I summarize all results in chapter 9 and give a brief outlook on future projects in chapter 10.

Chapter 2

Cosmology

Cosmology aims to describe the Universe as a whole, its structure and evolution in time. Twenty years ago this goal seemed hardly achievable, when cosmology was a vague and data-starving science. Nowadays, the situation has completely changed. There exist many different cosmological experiments providing huge amounts of data and this trend will most likely intensify in the next years. Cosmologists developed the so-called "standard" or "concordance model" of cosmology, based on the assumptions of large-scale homogeneity and isotropy, which provides a robust explanation for cosmological observations and has predicted several of them.

In this chapter we summarize the main aspects of cosmology, with the focus on details which are most relevant for this thesis. Starting from General Relativity and the dynamics in our Universe, we explain today's picture of structure formation and conclude with a description of the main experiments leading to the concordance model of cosmology. For more details on these topics we refer the reader to e.g. Peacock (1999) or Schneider et al. (2006).

2.1 The homogeneous and isotropic Universe

2.1.1 Einstein's Field Equation

By developing General Relativity (1907 to 1915), Albert Einstein set up a theory of gravitation. Einstein's field equation relates a matter distribution given by the energy-momentum tensor $T_{\mu\nu}$ to the curvature of four-dimensional spacetime described by the Einstein-Tensor $G_{\mu\nu}$

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

with G as the gravitational constant, c the speed of light and Λ the so-called cosmological constant. Today one usually shifts this term to the other side of the equation, interpreting Λ as a *vacuum energy density* (see Sect. 2.1.6). The tensor $G_{\mu\nu}$ is defined as

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \, , \tag{2.2}$$

with $R_{\mu\nu}$ as the Ricci tensor and R the curvature scalar. Both can be obtained as contractions of the Riemann tensor which provides a covariant description of the curvature of spacetime and can be expressed in terms of the metric $g_{\mu\nu}$.

2.1.2   Robertson-Walker Metric

On small scales the Universe is obviously not homogeneous. Galaxies and clusters of galaxies are large overdensities of matter compared to the intercluster medium. Large galaxy redshift surveys like the SDSS survey (latest data release from Adelman-McCarthy et al. 2008) indicate that the matter distribution is nearly homogeneous if we average over scales larger than 300 Mpc. The strongest evidence that the Universe is isotropic comes from the Cosmic Microwave Background (Smoot et al. 1991), which show only deviations of order 10^{-5}K from a mean temperature of 2.73K. Both observations strongly support the assumption that the Universe can be considered homogeneous and isotropic. This assumption, named cosmological principle, implies that any observer without peculiar velocity (so-called *comoving observer*) makes the same observations independent from his position. A metric describing such a universe is given by the Robertson-Walker metric

$$
\begin{aligned}
\mathrm{d}s^2 &= g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu \\
&= c^2\mathrm{d}t^2 - a^2(t)\left[\mathrm{d}w^2 + f_K^2(w)\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2\right)\right],
\end{aligned}
\tag{2.3}
$$

where w is the comoving radial distance and t the time measured by a comoving observer. ϕ and θ are angular coordinates. The expansion factor $a(t)$ describes the past and future of our Universe, whether it expands forever, contracts at some point in the future, or is static. The scale factor at present time $a(t_0)$ is normalized to be 1. The function $f_K(w)$ describing the comoving angular diameter distance (see Sect. 2.1.4) depends on the curvature K of the metric

$$
f_K(w) = \begin{cases} \frac{1}{\sqrt{K}}\sin(\sqrt{K}w) & K > 0 \\ w & K = 0 \\ \frac{1}{\sqrt{-K}}\sinh(\sqrt{-K}w) & K < 0, \end{cases}
\tag{2.4}
$$

where $K > 0$ stands for a closed, $K < 0$ for an open, and $K = 0$ for a spatially flat universe.

2.1.3   Friedmann equations

Inserting the Robertson-Walker metric into Einstein's field equation one derives the Friedmann equations which describe the dynamics in our Universe. The energy-momentum tensor takes the form of a perfect fluid $T_{\mu\nu} = \mathrm{diag}\left(\rho c^2, p, p, p\right)$, where ρ (the density) and p (the pressure) depend only on time

$$
\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{Kc^2}{a^2} + \frac{\Lambda}{3}
\tag{2.5}
$$

and

$$
\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right) + \frac{\Lambda}{3} .
\tag{2.6}
$$

The l.h.s. of equation (2.5) is the square of the so-called *Hubble parameter* $H(t)$ and its value today is called *Hubble constant* H_0 (see also Sect. 2.3). The Hubble constant H_0 is often parameterized as $H_0 = h\ 100(\text{km s}^{-1}\text{Mpc}^{-1})$, where h expresses the uncertainty in the measurement of H_0. Other measurements, depending on H_0 (e.g. distances), are similarly parameterized in terms of h. As mentioned in Sect. 2.1.1 the cosmological constant Λ, which was originally introduced by Einstein to allow for static universes, is nowadays interpreted as vacuum energy density. In the above two equations ρ only accounts for the energy densities of "normal" (pressureless) matter and radiation. From now on we adapt the concept of a constant vacuum energy density, therefore drop the Λ-term and extend the ρ to $\rho = \rho_m + \rho_r + \rho_\Lambda$. This concept is mathematically equivalent to a cosmological constant, but has a reasonable physical interpretation. Note, that there are alternative theories stating that the vacuum energy density must not necessarily be constant; we give more details on this topic in Sect. 2.1.6.

In order to determine the scale factor via the Friedmann equations we need information on $\rho(t)$. Differentiating (2.5) and inserting it into (2.6) we derive the so-called *adiabatic equation*

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\rho a^3 c^2\right) + p(t)\frac{\mathrm{d}a^3(t)}{\mathrm{d}t} = 0 \ . \tag{2.7}$$

It relates the change of energy density in a comoving volume to the pressure multiplied by the spatial change in proper volume[1] and can be interpreted as a conservation law of energy in the Universe. The three matter components, pressureless matter, radiation, and vacuum energy density, evolve differently in time and are dominant through different epochs. Pressureless matter, characterized as particles with velocities much smaller than c, has an *equation of state (EOS)* which reads $p = 0$. From (2.7) we derive

$$\rho_m \propto a^{-3}. \tag{2.8}$$

For particles with velocities close or equal to c the pressure term reads $p = \rho c^2/3$. Inserting this into (2.7) leads to

$$\rho_r \propto a^{-4} \ . \tag{2.9}$$

The corresponding equation for the constant vacuum energy density reads

$$\rho_\Lambda = \text{const} \ . \tag{2.10}$$

The density for which our Universe is flat is named *critical density*. Assuming that $\rho = \rho_m + \rho_r + \rho_\Lambda$, hence dropping the Λ-term in (2.5), it can be calculated by inserting $K = 0$. We obtain

$$\rho_{cr} = \frac{3H_0^2}{8\pi G} \ . \tag{2.11}$$

Density values are scaled with the critical density, for the three different "sorts of matter" we define

$$\Omega_m := \frac{\rho_m}{\rho_{cr}} \ , \qquad \Omega_r := \frac{\rho_r}{\rho_{cr}} \ , \qquad \Omega_\Lambda := \frac{\rho_\Lambda}{\rho_{cr}} = \frac{\Lambda}{3H_0^2} \ . \tag{2.12}$$

---

[1]Proper coordinates $r(t)$ are related to comoving coordinates x by the scale factor $a(t)$, $r(t) = a(t)x$.

The main contribution to Ω_r is given by photons of the Cosmic Microwave Background (CMB) (see Sect. 2.3) and the energy density of these photons follow a Planck distribution. By measuring the temperature (T_{CMB} = 2.73 K), Ω_r can be obtained from the Stefan-Boltzmann law; one finds that Ω_r is fairly small compared to Ω_m and Ω_Λ and can be neglected today. However, at earlier times, radiation must have been dominant because its density scales with $a(t)^{-4}$ whereas Ω_m is only proportional to $a(t)^{-3}$. The time when Ω_r equals Ω_m is denoted a_{eq}. Using the parameters defined above and the definition of the Hubble constant we can rewrite (2.5)

$$H^2(t) = H_0^2 \left[\frac{\Omega_r}{a(t)^4} + \frac{\Omega_m}{a(t)^3} - \frac{Kc^2}{a(t)^2 H_0^2} + \Omega_\Lambda \right] . \tag{2.13}$$

To describe the curvature of our current epoch we insert $a(t) = 1$ and derive

$$K = \left(\frac{H_0}{c} \right)^2 (\Omega_{tot} - 1) , \tag{2.14}$$

with $\Omega_{tot} = \Omega_r + \Omega_m + \Omega_\Lambda$. From this expression of the curvature in terms of density parameters we see that

- the spatial hypersurfaces of constant t are flat for $\Omega_{tot} = 1$ ($K = 0$)

- hyperbolic for $\Omega_{tot} < 1$ ($K < 0$)

- and $K > 0$ for $\Omega_{tot} > 1$.

Latest observations indicate that the total curvature of the Universe is small; $|\Omega_{tot} - 1| \lesssim 0.017$ (Komatsu et al. 2008), therefore it is generally assumed that the Universe has zero curvature. Nevertheless, if this assumption is relaxed constraints on other (curvature dependent) parameters weaken significantly.

The Hubble parameter and therefore the evolution of our Universe is completely determined by the density parameters, in terms of which we rewrite (2.13)

$$\left(\frac{H(t)}{H_0} \right)^2 = \left[\frac{\Omega_r}{a(t)^4} + \frac{\Omega_m}{a(t)^3} - \frac{1 - \Omega_{tot}}{a(t)^2} + \Omega_\Lambda \right] . \tag{2.15}$$

At very early times we expect the Universe to be hot and dense. Going back in time, density and temperature further increase until we finally reach $a(t) = 0$. Since this so-called Big Bang until today, $a(t)$ was a monotonically increasing function. In order to determine its evolution in the future we neglect Ω_r in (2.15) and consider only Ω_m, and Ω_Λ. Felten & Isaacman (1986) have shown that the Universe will recollapse in case $\Omega_\Lambda < 0$. If the vacuum energy is positive, expansion will continue to infinity if $\Omega_m < 1$. In case $\Omega_m > 1$ the collapse is still avoided if Ω_m does not exceed a certain threshold, which is determined by the value of Ω_Λ. These facts are nicely illustrated in Fig. 2.1, which shows the different states of the Universe (open, closed, flat, expanding, recollapsing) in dependence of the cosmological parameters Ω_m and Ω_Λ. In addition, we see the most likely parameter regions, constrained by different cosmological experiments (see Sect 2.3 for further details). The area where the results of the three experiments overlap suggests that the Universe is (almost) flat, with positive Ω_Λ, and Ω_m much smaller than 1.
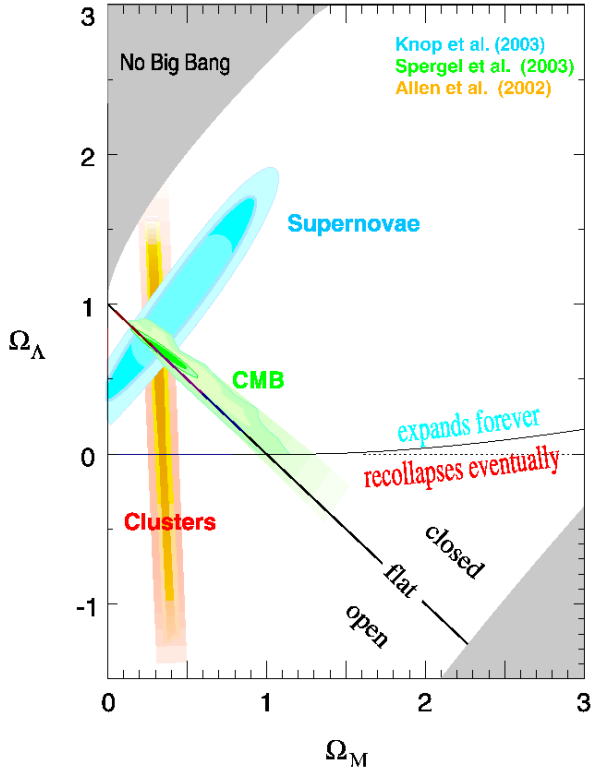
Figure 2.1: *From this figure one can deduce the expansion history and the curvature in our Universe for given Ω_m and Ω_Λ. The shaded area corresponds to parameter combinations which exclude a Big Bang, instead these regions represent bouncing or loitering universes. The diagonal solid line distinguishes parameter sets which lead to an open, closed, or flat Universe, respectively. The almost horizontal line divides expanding and potentially recollapsing cosmologies. In addition, we see the constraints on the parameter space obtained from 3 different cosmological probes which favor a close to flat Universe, with positive Ω_Λ, and Ω_m much smaller than 1. (Figure taken from Knop et al. 2003)*

2.1.4 Redshift and distances

Due to the expansion of the Universe, photons are redshifted on their way from the source to a comoving observer. The redshift is defined as

$$z := \frac{\lambda_0 - \lambda_e}{\lambda_e} \, , \qquad (2.16)$$

where λ_e is the emitted and λ_0 the observed wavelength. For low redshifts $z \ll 1$, we can relate the redshift of a galaxy to its recession velocity via the standard Doppler formula $z \simeq v/c$. Nevertheless, this description fails when going out to larger redshifts. The reason for this is that cosmological redshift is a relativistic effect due time dilation which can be seen from the following derivation. Photons propagate along null geodesics ($ds = 0$), hence we see from the comoving metric (2.3) that $c \, dt = -a \, dw$. Assume that two signals are emitted from a source a time t_e and $t_e + \Delta t$ which we observe at time t_0 and $t_0 + \Delta t_0$. The comoving separation between source and observer is constant; we can set up the following relation

$$\int_{t_0}^{t_e} \frac{c \, dt'}{a(t')} = w = \int_{t_0 + \Delta t_0}^{t_e + \Delta t_e} \frac{c \, dt'}{a(t')} = \int_{t_0}^{t_e} \frac{c \, dt'}{a(t')} + \frac{c \, \Delta t_0}{1} - \frac{c \, \Delta t_e}{a(t_e)} \, , \qquad (2.17)$$

where the time intervals Δt are small enough to neglect variations in $a(t)$. Comparing the left- and the right-hand side of the equation we see that Δt_e is time dilated, more precisely $\Delta t_e = a(t_e)\Delta t_0$.

The frequencies ν_e and ν_0 of the signal are the inverse of the corresponding time intervals, hence

$$\frac{1}{a(t_e)} = \frac{\nu_e}{\nu_0} = \frac{\lambda_0}{\lambda_e} = 1 + z \, . \tag{2.18}$$

In Euclidian space the definition of distance is unique, meaning that different methods to measure a distance give the same result. In contrast, there is no unique distance measure in a curved and expanding Universe, where the definition of distance must always be seen in context with the method it was obtained.

2.1.4.1 Comoving distance

The comoving distance is characterized as the coordinate distance on a spatial hypersurface between a comoving source (z_2) and a comoving observer (z_1). Similar to the above derivation of cosmological redshift we write

$$dw = -\frac{c}{a} \, dt = -\frac{c}{a^2 \, H} \, da \, , \tag{2.19}$$

where we use $dt = da/\dot{a}$ and $H(t) = (\dot{a}/a)$. Inserting (2.15) we can express the comoving distance in terms of density parameters as

$$w(z_1, z_2) = \frac{c}{H_0} \int_{a(z_2)}^{a(z_1)} \left[a \, \Omega_m + a^2 \left(1 - \Omega_m - \Omega_\Lambda \right) + a^4 \Omega_\Lambda \right]^{-1/2} da \, . \tag{2.20}$$

2.1.4.2 Angular diameter distance

Of special importance for gravitational lensing is the angular diameter distance which can be derived by easy geometrical consideration. It is given by the diameter of the source dL at redshift z_2 and the observed angular diameter $d\theta$ at redshift z_1. Using $dL = a(t) f_K(w) d\theta$, we define

$$D_{ang}(z_1, z_2) := \frac{dL}{d\theta} = a(z_2) f_K(w(z_1, z_2)) \, , \tag{2.21}$$

as the angular diameter distance.

2.1.4.3 Luminosity distance

The luminosity distance is defined as

$$D_{lum}(z_1, z_2) := \sqrt{\frac{L}{4\pi S}} \, . \tag{2.22}$$

S is the observed flux (at z_1) and L the luminosity of the source (at z_2). A general relation between luminosity distance and angular diameter distance was found by Etherington (1933) and reads

$$D_{lum}(z_1, z_2) = (1 + z)^2 D_{ang}(z_1, z_2) \, . \tag{2.23}$$

2.1.4.4 Horizon

Causal contact between particles in the past is only possible if the comoving distance between them is smaller than the comoving distance light has travelled since the Big Bang and the considered cosmic time. This maximum distance is called *comoving horizon* at time t. Using (2.19), it can be expressed as

$$r_H(a) = \int_0^t \frac{c \, dt'}{a(t')} = \int_0^{a(t)} \frac{c \, da'}{a'^2 H(a')} \, . \tag{2.24}$$

We mention two special cases, namely the comoving horizon during matter and radiation-dominated phase of the Universe

$$r_H(a) = 2\frac{c}{H_0 \sqrt{\Omega_m}} \sqrt{a} \qquad \text{for} \qquad a \gg a_{eq} \, , \tag{2.25}$$

$$r_H(a) = \frac{c}{H_0 \sqrt{\Omega_r}} a \qquad \text{for} \qquad a \ll a_{eq} \, , \tag{2.26}$$

where a_{eq} defines the scale factor at matter-radiation equality

$$a_{eq} = \Omega_r/\Omega_m = (32000 \, \Omega_m h^2)^{-1} \, . \tag{2.27}$$

In the above calculation we assume that the dominant contribution to Ω_r comes from the CMB photons and neutrinos (which at this epoch were still relativistic and therefor contribute to the radiation density), hence $\Omega_r = 1.68\Omega_{CMB} = 3.2 \times 10^{-5} h^{-2}$.

2.1.5 Big Bang Nucleosynthesis and Cosmic Microwave Background

At very early times the Universe is in an extremely hot and dense state where all constituents are in thermal equilibrium. Then the radiation field is described by a Planck distribution, and therefore solely depends on the temperature of the Universe. Note that temperature and energy uniquely define a cosmic time and are used as synonyms in this section.

At high temperatures, the formation of atomic nuclei is prohibited, as they are immediately destroyed by high energy photons. When the Universe expands and cools, light elements can form. This process starts with deuterium, which has a binding energy of 2.24 MeV. Still, deuterium does not form at this energy, due to the much higher number density of photons compared to baryons. More precisely, for temperatures > 0.07 MeV there are sufficiently many photons in the high energy tail of the Planck distribution to destroy deuterium via photodissociation. In principle, helium could have formed earlier; it is able to sustain high energy photons at temperatures of $T \approx 0.28$ MeV. However, helium is only formed through an intermediate step of deuterium (*deuterium bottleneck*). As soon as the deuterium number density is high enough this results in an immediate burst of helium production, binding nearly all neutrons ($T \approx 0.1$ MeV). Due to this almost instant process, the baryon fraction of helium depends solely on the ratio of protons to neutrons at $T \approx 0.1$ MeV

$$X_{^4He} = \frac{4(n_n/2)}{n_n + n_p} = \frac{2(n_n/n_p)}{1 + n_n/n_p} \, . \tag{2.28}$$

Protons and neutrons have formed earlier during the phase of baryongenesis and are held in equilibrium only through weak interactions, e.g.

$$p + e^- \longleftrightarrow n + \nu \qquad \text{and} \qquad p + \bar\nu \longleftrightarrow n + e^+ . \tag{2.29}$$
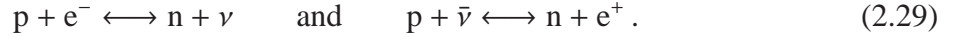
From the cross section the above particle interactions, we can estimate the temperature at which the corresponding reaction rate becomes too low to hold the particle in thermal equilibrium. Via the above interactions, protons and neutrons are in equilibrium until $T \approx 0.7$ MeV. The ratio of neutrons to protons at this time is 1/6. For ≈ 3 minutes, between $T \approx 0.7$ MeV and $T \approx 0.1$ MeV (when helium forms), the neutrons decay according to $n \longrightarrow p + e + \bar\nu$ (the decaying time of the neutron is well known to be $\tau_n = 886s$). This process changes the ratio of neutrons to protons to $n_n/n_p = 1/7$. We can now calculate the baryon fraction of helium to be $X_{^4\mathrm{He}} = 1/4$, which is in perfect agreement with today's observations (see Sect 2.3). In addition to ^{4}He, small amounts of deuterium, ^{3}He, ^{7}Li, and ^{7}Be are created during Big Bang nucleosynthesis, but no heavier elements.

When the Universe cools even more ($T \approx 1$ eV), the nuclei (mainly protons) and free electrons form neutral atoms, mainly hydrogen. This process of *recombination* takes place at $z \simeq 1100$, well after matter-radiation equality. The binding energy of hydrogen is 13.6 eV, still it cannot form at the corresponding time due to the same argument we mentioned in the context of deuterium formation. From $z \simeq 1100$ on, the photons and baryons decouple and stream freely throughout the Universe.The energy density of the photons is described through a Planck distribution, today we observe this Planck spectrum, highly redshifted, as the *Cosmic Microwave Background (CMB)*. Since its first detection (Penzias & Wilson 1965), many experiments have analyzed its properties. As an example we mention *WMAP (Wilkinson Microwave Anisotropy Probe)*, which is a satellite mission launched in 2001 to measure the anisotropies of the CMB. After subtracting the temperature differences due to our peculiar velocity and microwave emission from the galactic plane, we see small fluctuations of order $\Delta T/T = 10^{-5}$ around an average temperature of $T \approx 2.73$ K (see Fig. 2.2). These small temperature perturbations mirror the fluctuations of density, potential and peculiar velocity of matter at the time of recombination. WMAP confirmed the foregoing results of the COBE mission (for which John C. Mather and George F. Smoot were awarded a Nobel prize) with a much higher resolution and had a large impact on the current picture of our Universe (see Sect. 2.3.2).

2.1.6 Dark Energy

There are two main observations which indicate that there is a third source contributing to the energy density in our Universe, besides pressureless matter and radiation. First, we now from CMB measurements that the curvature of the Universe is very close to zero, hence the overall energy density of the Universe is close to the critical density. The contribution of matter and radiation to ρ can be inferred by observations and is approximately only one quarter of ρ_{crit}; the rest must be contributed from a third component, namely dark energy. Second, we know from supernovae experiments (Riess et al. 1998; Perlmutter et al. 1999), that the expansion of the Universe accelerates ($\ddot{a} > 0$). Using (2.6), $\ddot{a} > 0$ can only occur if dark energy has negative
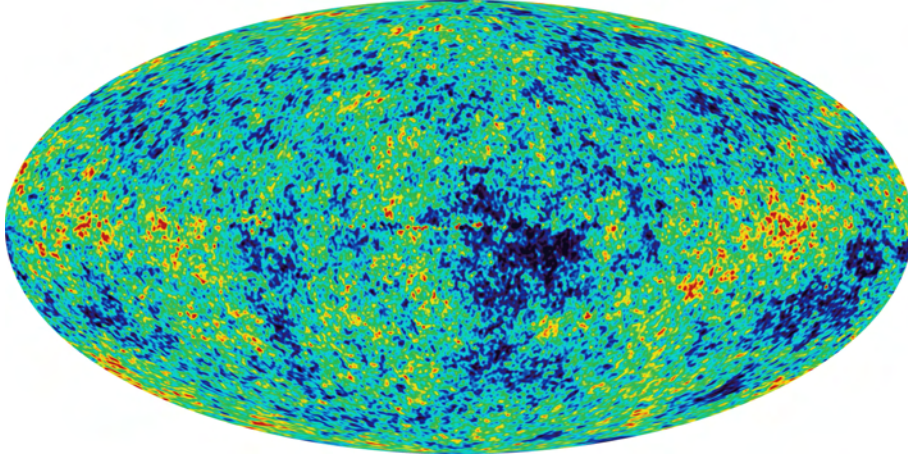
Figure 2.2: *The temperature fluctuations of the CMB measured by WMAP over a range of 400 μK around the average of 2.73 K. Blue imprints refer to cold spots, red to warmer regions. Foreground contaminations (e.g. galaxy dipole) are removed in this picture. (Figure 11 from Bennett et al. 2003)*

pressure, i.e. ($\rho + \frac{3p}{c^2} < 0$). The relation of energy density to pressure is quantified in the *equation of state parameter*

$$w := \frac{p_\Lambda}{\rho_\Lambda c^2} \,, \tag{2.30}$$

with which we can rewrite the adiabatic equation (2.7) for dark energy as

$$\frac{d\rho_\Lambda}{dt} + 3\frac{\dot{a}}{a}\rho_\Lambda (1 + w) = 0 \,. \tag{2.31}$$

For time dependent w, (2.31) can be integrated to obtain

$$\rho_\Lambda \propto \exp\left(-3 \int_1^a \frac{da'}{a'} \left[1 + w(a')\right]\right) \,, \tag{2.32}$$

which specifies for constant w to

$$\rho_\Lambda(a) = \rho_{\Lambda 0}\, a^{-3(1+w)} \,, \tag{2.33}$$

where $\rho_{\Lambda 0}$ is the energy density measured today. Note that the above statements and derivations apply similarly to pressureless matter and radiation, obtaining $w = 0$ for the first case and $w = 1/3$ for the latter. The favored interpretation of dark energy today is constant vacuum energy density.

Constant vacuum energy density, corresponds to $w = -1$. This model has nice properties, e.g. from the argument that the energy-stress tensor of the vacuum is Lorentz invariant, we can directly deduce $\rho = p/c^2$, which is exactly the behavior of a cosmological constant and

explains the accelerated expansion. Furthermore, quantum mechanics predicts that in a vacuum, particles and anti-particles are created and annihilate around a "zero point" energy and that the lifetime of these particles, although short, gives the vacuum a non-zero ground-state energy. On first sight, this all seems to be a convincing theory, nevertheless there is a large drawback, when attempting to estimate the value of vacuum energy density via quantum field theory. Carroll et al. (1992) show that the theoretical estimate for the vacuum energy density is $\rho_\Lambda \approx 10^{92} \text{ergs/cm}^3$. Normalized with ρ_{crit} this corresponds to $\Omega_\Lambda \approx 10^{120}$, deviating by 120 magnitudes from the observed $\Omega_\Lambda \approx 0.75$. In the aforementioned paper, ρ_Λ is calculated as an integral over the energy density of quantum fields (leptons, quarks, gauge fields), which is cut off at an energy scale where quantum field theory breaks down and a new theory of physics (quantum gravity) would be required. This is the so-called Planck energy (10^{19} GeV). The cut-off can be shifted towards lower energy scales which reduces the predicted value of ρ_Λ. Nevertheless one must cut off the integral at 10^{-2} eV to obtain $\rho_\Lambda \approx 0.7$. This is out of question, as physics at this energy level is perfectly described by quantum mechanics.

Assuming that supersymmetry holds solves part of the problem. Then, fermions and bosons have SUSY partners with equal energy density but different signs. However, to avoid a perfectly zero vacuum energy density, SUSY must be broken at a certain mass scale M, such that the remnant particles with $m < M$ form today's vacuum energy density. Still, no supersymmetric particle has been observed, indicating that they are quite massive, and that SUSY is broken at high energies ($M > 1$ TeV). Assuming that particles with higher rest mass cancel, one arrives at a vacuum energy density of $\Omega_\Lambda \approx 10^{60}$, which still is an enormous discrepancy to the observed value. We see that vacuum energy density is far from being understood, we will briefly mention alternative models.

Quintessence or scalar field   models subsume dynamical dark energy models, i.e. the equation of state parameter w can vary in time. The dynamic is described by a scalar field ϕ with Langrangian density

$$\mathcal{L} = 1/2 \partial^\mu \phi \partial_\mu \phi - V(\phi) \,. \tag{2.34}$$

Therefrom we can calculate the pressure and the density of the scalar field to be

$$\rho = \dot{\phi}^2/2 + V(\phi) \qquad p = \dot{\phi}^2/2 - V(\phi) \tag{2.35}$$

where ϕ is supposed to be spatially homogeneous, but time dependent, $\dot{\phi}^2/2$ denotes the kinetic energy of the field and $V(\phi)$ is the potential energy. Applying the definition of w (2.30) we derive for scalar fields

$$w = \frac{\dot{\phi}^2/2 - V(\phi)}{\dot{\phi}^2/2 + V(\phi)} = \frac{-1 + \dot{\phi}^2/2V(\phi)}{1 + \dot{\phi}^2/2V(\phi)} \,. \tag{2.36}$$

From the above equation we see that $w \approx -1$ if the scalar field acts as a slowly changing vacuum energy density ($\dot{\phi}^2/2V(\phi) \ll 1$) and $w \approx 1$ for a rapidly changing energy density. In both cases w changes in time; Linder (2003) introduced the following parameterization

$$w(a) = w_0 + w_a(1 - a) \,. \tag{2.37}$$

The above parameterization can be used to distinguish between the two classes of scalar field models, so-called "thawing" and "freezing" models. To explain these models we introduce the equation of motion for a scalar field, which is obtained from Einstein's field equation. It reads

$$\ddot{\phi} + 3H\dot{\phi} + V'(\phi) = 0\,, \tag{2.38}$$

where the prime denotes the derivative with respect to ϕ. Freezing models are characterized by a potential $V(\phi)$ which drops more rapidly than the friction term $3H\dot{\phi}$ as time progresses. Hence, the friction term in the above equation starts to dominate at late epochs. Thawing models have opposite behavior; here the friction term dominates at early times and the field changes rapidly for late times. The different behavior of both models is reflected in $w(a)$. With the future precision of cosmological data, we can use the above parameterization (2.37) to distinguish between thawing or freezing models. There exist also more complicated scalar field models, obtained by changing the kinetic term in the Lagrangian (e.g. k-essence), nevertheless there is no compelling physical explanation for either of them.

Modified gravity is another suggested explanation. The basic idea is that the geometrical part of Einstein' equation (2.1) must be modified (i.e. the Einstein tensor) rather than adding something to the energy-momentum tensor. For more details on this topic the reader is referred to Bekenstein (2004). Note that this theory faces severe problems when explaining the third peak of the CMB power spectrum (Spergel et al. 2007) or the bullet cluster (Clowe et al. 2006).

Other models basically question the cosmological principle, in particular the assumption of the Universe being homogeneous. Although this assumption is in good agreement with observations (Sect 2.1.2), it is not a direct consequence of them. Clarkson et al. (2008) suggest a purely geometrical method, which is based on the relation of Hubble parameter to a modified angular diameter distance, to test for the cosmological principle. But even if we can prove that on large scales the universe is homogeneous, on small scales it is truly inhomogeneous. When describing the dynamics in the Universe, should not we first estimate the dynamics of local inhomogeneities and then average over these, instead of first averaging over the inhomogeneities and then calculate the dynamics of this homogeneous average? This question is intensively discussed in the review of Buchert (2008); it is true that averaging Einstein's equation on small scales leads to a repulsive, so-called *backreaction* term. Such a term mimics the behavior of dark energy, however the strength of this effect is very much under debate. In very recent work, Wiltshire (2007a,b) presents a solution for the dynamics in an inhomogeneous universe, which replace the Friedmann equations. In this work inhomogeneities are incorporated by considering two scales, i.e. voids, which expand rapidly, and bubble walls containing clusters and galaxies, which surround theses voids. Wiltshire (2007b) performs several observational tests with this new theory claiming that it is viable, so far.

Final comment We do not know what drives the expansion of the Universe, whether it is any of the dark energy models or some General Relativity effect, which has been neglected so far. Investigating dark energy is one major task of future work and surveys. For more details on

upcoming dark energy surveys and possibilities how to constrain dark energy models best, the reader is referred to Albrecht et al. (2006) and Peacock et al. (2006). In this thesis we consider dark energy parameter estimation with cosmic shear, which is done in chapter 8.

2.2   Structure formation

On large scales the Universe is homogeneous and isotropic. On smaller scales this is obviously and fortunately not the case. The structure we observe today, like clusters and galaxies, has most likely evolved from small primordial overdensities, which further increase through gravitational processes. Due to gravitational instability, an initial overdense region expands slower compared to the expansion rate of the surrounding Universe, causing a further increase of its density contrast. If the initial overdensity was large enough, at some point it decouples from the expansion of the Universe, collapses and forms a cluster. This scenario is supported by the fact that we observe imprints of these primordial perturbations in the CMB. The density contrast at comoving spatial coordinates ($\mathbf{x}$) and cosmic time t is defined as

$$\delta(\mathbf{x}, t) := \frac{\rho(\mathbf{x}, t) - \bar{\rho}(t)}{\bar{\rho}(t)} \, , \tag{2.39}$$

where $\rho(\mathbf{x}, t)$ denotes the density at $(\mathbf{x}, t)$ and $\bar{\rho}$ is the mean density of the Universe.

2.2.1   Linear structure formation

We start with a qualitative description of the different physical processes, which affect a density perturbation. This can be done best in Fourier space, we therefore introduce the Fourier transform of $\delta(\mathbf{x})$ as

$$\hat{\delta}(\mathbf{k}) = \int_{\mathbb{R}^3} \mathrm{d}^3x \, \delta(\mathbf{x}) \, e^{i\mathbf{x}\cdot\mathbf{k}} \qquad \longleftrightarrow \qquad \delta(\mathbf{x}) = \int_{\mathbb{R}^3} \frac{\mathrm{d}^3k}{(2\pi)^3} \, \hat{\delta}(\mathbf{k}) \, e^{-i\mathbf{x}\cdot\mathbf{k}} \tag{2.40}$$

with $\mathbf{k}$ as the comoving wave vector. These Fourier modes of the density field evolve independently as long as the perturbation is in the linear regime, i.e. $|\delta(\mathbf{x}, t)| \ll 1$, furthermore the evolution of a mode depends only on the magnitude of $\mathbf{k}$. This magnitude is related to a characteristic length scale in real space, i.e. $\lambda = 2\pi/k$. If this length scale is larger than the comoving horizon (Sect. 2.1.4.4), the perturbation is not affected by any physical processes. Once this scale becomes smaller than the comoving horizon ("the mode enters the horizon"), physical processes start to influence the perturbation. Knowing that the physical conditions during the radiation-dominated phase of the universe are substantially different compared to the matter-dominated phase, we deduce that the evolution of a mode which enters the horizon before matter-radiation equality (see Sect. 2.1.4.4), differs from a mode with $a_{\mathrm{enter}} \gg a_{\mathrm{eq}}$, where we define a_{enter} as the expansion factor at the time when the mode enters the horizon.

Qualitatively the evolution of a density fluctuation in time is described as

$$\hat{\delta}(\mathbf{k}, a) = \hat{\delta}_i(\mathbf{k}) \, T(k) \, \frac{D_+(a)}{D_+(a_i)} \, , \tag{2.41}$$

where $\hat{\delta}_i(\mathbf{k})$ denotes the density perturbation at an initial epoch $a_i \ll a_{\mathrm{eq}}$, $T(k)$ the so-called *transfer function* and D_+ is the *growth factor* (the subscript is explained in Sect 2.2.1.1). The transfer function accounts for the evolution of $\hat{\delta}_i(\mathbf{k})$ during the radiation-dominated phase and through matter-radiation transition (see Sect. 2.2.1.2). The growth factor describes the late-time evolution in the matter-dominated and dark energy-dominated universe.

2.2.1.1 Growth factor

The evolution of sub-horizon modes can be described by Newtonian gravity. In contrast, we must refer to General Relativity to describe the super-horizon fluctuations. The dominant particle species is collisionless dark matter which implies that multi-streams can occur; hence, there is no well defined velocity field. The proper way of describing such a system employs the collisionless Boltzmann equations; a thorough treatment can be found in Dodelson (2003). However, we approximate matter as a pressureless fluid, which is a valid assumption on large scales or at early times when multi-streams are negligible. The complete set of evolution equations reads

$$\frac{\partial \rho(\mathbf{r}, t)}{\partial t} + \nabla_r \cdot [\rho \mathbf{u}(\mathbf{r}, t)] = 0 \qquad \text{Continuity equation,} \qquad (2.42)$$

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla_r) \mathbf{u} = -\nabla_r \phi \qquad \text{Euler equation,} \qquad (2.43)$$

$$\nabla_r^2 \phi = 4\pi G \rho - \Lambda \qquad \text{Poisson equation,} \qquad (2.44)$$

with $\mathbf{u}$ as the velocity of the fluid, ϕ the gravitational potential and $\mathbf{r}$ the proper coordinates. The Poisson equation was modified by the Λ term to allow for a cosmological constant. The homogeneous solution of the above set of equations reproduces the Friedmann equation (2.5). We transform equations (2.42) - (2.44) to comoving coordinates using $\mathbf{x} = \frac{\mathbf{r}}{a(t)}$, defining the comoving density $\rho(\mathbf{r}, t) = \hat{\rho}(\frac{\mathbf{r}}{a(t)}, t) = \hat{\rho}(\mathbf{x}, t)$ and decomposing the velocity into homogeneous expansion and peculiar velocity $\mathbf{u}(\mathbf{r}, t) = \dot{a}\mathbf{x} + \mathbf{v}(\mathbf{x}, t)$. In addition, we insert the definition of the density contrast (2.39) for the comoving density $\hat{\rho}$ into the Continuity equation which yields the following evolution equations in comoving coordinates

$$\frac{\partial \delta}{\partial t} + \frac{1}{a} \nabla_x \cdot [(1 + \delta)\mathbf{v}] = 0 \qquad \text{Continuity equation,} \qquad (2.45)$$

$$\frac{\partial \mathbf{v}}{\partial t} + \frac{\dot{a}}{a} \mathbf{v} + \frac{1}{a} (\mathbf{v} \cdot \nabla_x) \mathbf{v} = -\frac{1}{a} \nabla_x \Phi \qquad \text{Euler equation,} \qquad (2.46)$$

$$\nabla_x^2 \Phi = \frac{3H_0^2 \Omega_{\mathrm{m}}}{2a} \delta \qquad \text{Poisson equation,} \qquad (2.47)$$

with $\Phi(\mathbf{x}, t) := \phi(a\mathbf{x}, t) + \frac{\ddot{a}a}{2}|\mathbf{x}|^2$ as the comoving gravitational potential. A unperturbed expanding Universe corresponds to the solution $\delta \equiv 0$, $\mathbf{v} \equiv 0$, $\Phi \equiv 0$. Recall that we are in the linear regime, hence $|\delta| \ll 1$ and similarly the peculiar velocity $\mathbf{v}$ is small compared to the homogeneous expansion of the Universe. We therefore linearize the above equations in δ and $\mathbf{v}$. Taking the time derivative of the linearized continuity equation and the divergence of the linearized Euler equation, we combine both with the Poisson equation. Thereby we eliminate the peculiar velocity

and Φ deriving

$$\frac{\partial^2 \delta}{\partial t^2} + \frac{2\dot{a}}{a}\frac{\partial \delta}{\partial t} - \frac{3H_0^2 \Omega_{\mathrm{m}}}{2a^3}\delta = 0 \; . \tag{2.48}$$

This homogeneous differential equation describes the evolution of density perturbations in the linear regime. As the partial derivatives in (2.48) are only taken with respect to time and the coefficients are independent of time we can factorize the equation and obtain the solution

$$\delta(\mathbf{x}, t) = D_+(t)\delta_{0+}(\mathbf{x}) + D_-(t)\delta_{0-}(\mathbf{x}) \; , \tag{2.49}$$

with $\delta_{0\pm}$ denoting the density perturbation at a specific time (e.g. today). The functions $D_\pm$ are linearly independent solutions of

$$\frac{\partial^2 D}{\partial t^2} + \frac{2\dot{a}}{a}\frac{\partial D}{\partial t} - \frac{3H_0^2 \Omega_{\mathrm{m}}}{2a^3}D = 0 \; . \tag{2.50}$$

The function D_- decreases in time (*decaying mode*), even if it was present at early times it does not contribute to today's perturbations. We only consider D_+ (*growing mode*) as relevant for structure formation. Equation (2.49) reduces to

$$\delta(\mathbf{x}, t) = D_+(t) \, \delta_{0+}(\mathbf{x}) \, , \tag{2.51}$$

which indicates that the initial shape of the density fluctuations in comoving coordinates does not change with time. A general solution for (2.51) is given by

$$D_+(a) \propto \frac{H(t)}{H_0} \int_0^a \frac{\mathrm{d}a'}{\left[\Omega_{\mathrm{m}}/a' + \Omega_\Lambda a'^2 - (\Omega_{\mathrm{m}} + \Omega_\Lambda - 1)\right]^{\frac{3}{2}}} \, , \tag{2.52}$$

with the additional constraint to normalize the function to $D_+(t_0) = D_+(a = 1) = 1$. With this normalization $\delta_0(\mathbf{x})$ in (2.51) is today's density contrast according to linear perturbation theory. The normalized growing mode is called growth factor (see Fig. 2.3). Note that for an EdS-universe ($\Omega_{\mathrm{m}} = 1.0$, $\Omega_\Lambda = 0$), the growth factor scales as the scale factor.

2.2.1.2   Transfer function

The growth factor only describes the evolution of a perturbations inside the horizon and in a matter-dominated universe. In addition we need a description for superhorizon perturbations, both in the matter and radiation-dominated phase, and for modes which enter the horizon before a_{eq}. For such superhorizon perturbations we must refer to perturbation theory of General Relativity and can no longer use Newtonian gravity. For a detailed derivation the reader is again referred to Dodelson (2003). Furthermore, if a mode enters the horizon before a_{eq}, radiation pressure prevents the perturbation to grow. This means $\hat{\delta}$ of this mode is constant until the Universe becomes matter-dominated. In summary, one has to distinguish three phases for a mode of a given length scale and corresponding a_{enter}

$$\hat{\delta} \propto a^2 \qquad \text{if} \qquad a \ll a_{\mathrm{enter}} < a_{\mathrm{eq}} \, , \tag{2.53}$$

$$\hat{\delta} \propto \text{const} \qquad \text{if} \qquad a_{\mathrm{enter}} < a < a_{\mathrm{eq}} \, , \tag{2.54}$$

$$\hat{\delta} \propto a \qquad \text{if} \qquad a > a_{\mathrm{eq}} \, . \tag{2.55}$$
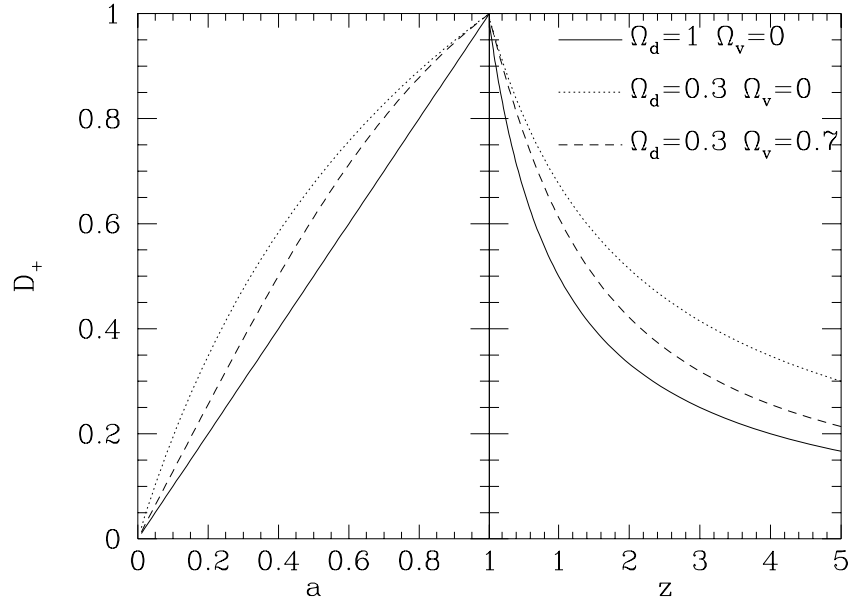
Figure 2.3: *Illustration of the growth factor D_+ depending on scale factor (left panel) and red-shift (right panel) for different density parameters. In case of an Einstein-de Sitter Universe (Ω_m =1, Ω_Λ=0) D_+ evolves similarly to the scale factor. (Figure from Schneider et al. 2006)*

We define the transfer function as

$$\frac{\hat{\delta}(k, a = 1)}{\hat{\delta}(k_{\text{small}}, a = 1)} := T(k) \frac{\hat{\delta}(k, a_{\text{early}})}{\hat{\delta}(k_{\text{small}}, a_{\text{early}})} , \tag{2.56}$$

where the l.h.s. expresses today's ratio of two density perturbations, and the r.h.s. shows the same ratio at early times. The transfer function relates both epochs and accounts for a possible suppression of growth. Note that in the above definition k is variable, whereas k_{small} is fixed to the scale of a perturbation which enters the horizon at late times. The scale factor a_{early} is fixed such that all scales of interest are outside the horizon at this time. Two limits of the transfer functions are easy to derive. First, for small k it approaches unity. Second, large k perturbations are suppressed, which is illustrated in Fig. 2.4. We see that the perturbation is reduced by a factor of $T(k) = (a_{\text{enter}}(k)/a_{\text{eq}})^2$. For $a \ll a_{\text{eq}}$ we can relate k to a_{enter} as follows

$$k \approx 1/r_{\text{H}}(a_{\text{enter}}(k)) = \frac{H_0 \sqrt{\Omega_{\text{r}}}}{c \, a_{\text{enter}}} \tag{2.57}$$

where in the last step we use (2.25). We see that $k \propto 1/a_{\text{enter}}$, hence $T \propto k^{-2}$ for large k.

Last, the transfer function depends on the type of dark matter which is dominant in the Universe. If the Universe was dominated by particles with relativistic velocities, so-called *Hot Dark Matter* (for example massive neutrinos), small potential well would not attract enough particles to form

Figure 2.4: *A schematic sketch of the transfer function. Modes which enter the horizon before a_{eq} are suppressed by the factor $(a_{\mathrm{enter}}/a_{eq})^2$. (Figure taken from Schneider 2006)*
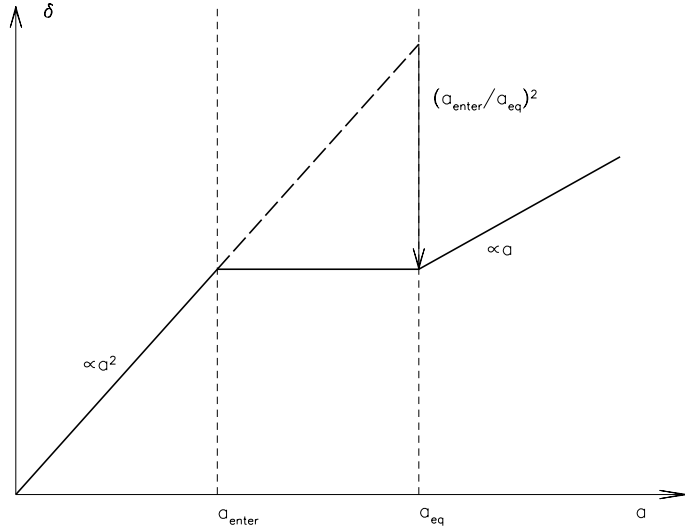
a perturbation due to the high velocity of the particles. This implies that large perturbations such as clusters and superclusters of galaxies should have formed first. However, when going to higher redshift, we do not observe the clusters and superclusters, but many isolated galaxies. For this reason we conclude that Hot Dark Matter only contributes a small fraction to Ω_m and prefer the model of *Cold Dark Matter* (CDM) where particles move with velocities which are highly non-relativistic.

In this thesis we use the transfer function fitting formula found by Efstathiou et al. (1992) which reads

$$T(k) = \left[1 + \left(6.4q + (3.0q)^{3/2} + (1.7)^2\right)^{\nu}\right]^{-1/\nu} \text{ with } \nu = 1.13 \; . \tag{2.58}$$

The definitions of q reads

$$q := \frac{k\mathrm{Mpc}}{h\Gamma} \; , \tag{2.59}$$

and for the shape parameter Γ we use the expression given in Sugiyama (1995)

$$\Gamma = \Omega_m h \, \exp\left[-\Omega_b(1 + \sqrt{2h}/\Omega_m)\right] \; . \tag{2.60}$$

2.2.2 Statistical properties of the density field

Structure formation is not able to describe the specific density field in our Universe, we can only achieve a description of its statistical properties as a function of time. We consider the density field of our Universe as one realization of a random field, whose statistical properties we want to determine. In principle, this determination requires an averaging process over many independent realizations, however we have only one observable Universe, hence the *ensemble average* must be replaced by a *volume average*. We average over different regions of the Universe with a separation that is large enough to consider these regions as independent. Then, the statistical

properties of this random field are described through its *moments* (mean, two-point correlation function, three-point correlation function, etc.).

For the special case of a Gaussian density field the statistical properties are fully described by the mean and two-point correlation function (or its Fourier-space equivalent, the power spectrum). The WMAP 5-years analysis of CMB fluctuations by Komatsu et al. (2008) show no measureable deviations from Gaussianity (note that this is questioned by the result of Yadav & Wandelt 2008). If structure growth was linear, today's density field is a Gaussian random field. However, structure growth is driven by gravitational instability which is a non-linear process and the approximation of linear perturbation theory only holds in case the density contrast (2.39) is small. On small scales, when δ grows larger than 1, non-linear structure growth becomes non-negligible and causes a coupling of Fourier modes, which implies a non-Gaussian density field today even if the initial field was Gaussian. Here, structure growth becomes non-linear. We have to consider non-linear models and higher-order moments of the density field.

2.2.2.1 Power spectrum

The mean of the density field vanishes($\langle \delta(\mathbf{x}) \rangle = 0$); we define its two-point correlation function as

$$\langle \delta(\mathbf{x})\delta^*(\mathbf{y}) \rangle := C_{\delta\delta}(|\mathbf{x} - \mathbf{y}|). \tag{2.61}$$

The homogeneity and isotropy of the density field imply that the two-point correlation function only depends on the separation $|\mathbf{x} - \mathbf{y}|$. Defining the Fourier transform of $\delta(\mathbf{x})$ as in (2.40) and using expression (2.61), we can calculate

$$\langle \hat{\delta}(\mathbf{k})\hat{\delta}(\mathbf{k}') \rangle = (2\pi)^3 \delta_{\mathrm{D}}(\mathbf{k} - \mathbf{k}')P_\delta(|\mathbf{k}|) \, , \tag{2.62}$$

where we defined the power spectrum P_δ as the Fourier transform of the correlation function

$$P_\delta(|\mathbf{k}|) = \int_{\mathbb{R}^3} \mathrm{d}^3 y \, e^{i\mathbf{y}\cdot\mathbf{k}} C_{\delta\delta}(|\mathbf{y}|) \, . \tag{2.63}$$

The relation (2.41) together with (2.61) and (2.63) enables us to relate a power spectrum at time t to today's power spectrum $P_0(k)$

$$P_\delta(k, a) = T^2(k)\frac{D_+^2(a)}{D_+^2(a_i)}P_i \, , \tag{2.64}$$

with P_i being the initial power spectrum. Due to the fact that in the very early universe no characteristic length scale is preferred and that a power law is the only scale invariant function, one chooses $P_i = Ak^{n_\mathrm{s}}$ (see Sect. 2.2.2.2), where A is a normalization constant and n_s is the spectral index, which is assumed to be ≤ 1. The normalization must be determined through observations. The dispersion of the density field smoothed on scales R is defined as

$$\sigma^2(R) = \left\langle \delta_R^2(\mathbf{x}) \right\rangle = \int \frac{\mathrm{d}^3 k}{(2\pi)^3} \left| \tilde{W}_R(k) \right|^2 P_\delta(k) \, , \tag{2.65}$$

with

$$\tilde{W}_R(k) = 3\frac{\sin(kR) - kR\cos(kR)}{(kR)^3} \, . \tag{2.66}$$

The above function is the Fourier transform of a spherical top-hat filter function with radius R. If the dark matter distribution follows the galaxy distribution one can measure $\sigma(R)$ simply through counting galaxies inside spheres of comoving radius R and calculating the dispersion as

$$\sigma^2(R) = \frac{\left\langle (N(R) - \langle N(R)\rangle)^2 \right\rangle}{\langle N\rangle^2} \, . \tag{2.67}$$

Commonly the power spectrum is normalized by σ_8 which is defined as

$$\sigma_8 = \sqrt{\sigma^2(R = 8h^{-1}\mathrm{Mpc})} \, . \tag{2.68}$$

This parameter describes the dispersion of density fluctuations in a sphere with radius $8\,h^{-1}$ Mpc.

2.2.2.2   Inflation and initial power spectrum

In the framework of the standard model two main problems occur, the *flatness problem* and the *horizon problem*. The first questions the extreme fine tuning of the curvature to $K = 0$ in the early Universe, which is required to explain CMB constraints on $\Omega_{\mathrm{tot}} \in [0.97; 1.04]$. The second problem addresses the uniform temperature of the CMB. Although different patches of the sky have never been in causal contact (the horizon size was smaller than $2°$ at the time of recombination), there exist only small fluctuations. An explanation for both questions is provided by inflation (Guth 1981), which predicts that the Universe had a very rapid phase of expansion, at a very early time. This implies a much larger horizon size before inflation which then decreases due to the rapid expansion. If all scales of the observed CMB were in causal contact before inflation (which solves the horizon problem) the Universe must have expanded exponentially by ≈ 64 e-folds during this epoch. Note, that this also solves the fine tuning of the flatness at the beginning of the radiation-dominated phase, because any curvature is smoothed out during the rapid expansion. Very similar to the quintessence dark energy models (see Sect. 2.1.6), one assumes that inflation is driven by a scalar field ϕ with negative pressure and a potential. This potential must be sufficiently flat for the energy density of ϕ to be approximately constant, as the field *rolls down* the potential to its energy ground state (e.g. Linde 1982). Note, that inflation predicts a scale-free power spectrum with a spectral index slightly smaller than 1. This small deviation from $n_s = 1$ quantifies the flatness of the potential $V(\phi)$, with the limit of $n_s = 1$ for $V(\phi)$ being flat. The WMAP-5years results indicate a slight deviation from a Harrison Zel'dovich power spectrum, i.e. $n_s \approx 0.96^{+0.014}_{-0.013}$.

2.2.2.3   The non-linear power spectrum - HKLM method

As mentioned above, the assumption of a Gaussian density field breaks down on small scales. In this section, we describe a method to obtain a non-Gaussian power spectrum using a fit-formula which is calibrated from numerical simulations, the so-called HLKM-method (Hamilton et al.

1991).

Considering a virialized cluster of galaxies, the internal density structure of such an object stays constant in time (*hypothesis of stable clustering*). It is not affected by the expansion of the universe, therefore in comoving coordinates we can consider a virialized object to collapse. Hamilton et al. (1991) introduced two different scalings, the linear one (r_L) which refers to the evolution of the cluster before its decoupling from the expansion and the nonlinear one (r_{NL}) referring to the time after the collapse. The collapse of a cluster is described as a transition from linear to non-linear scale. The matter content of the cluster is constant; this matter conservation yields a relation between the two scales

$$r_L^3 = (1 + \delta_{NL})r_{NL}^3 \; . \tag{2.69}$$

An important step in the ansatz of Hamilton et al. (1991) is to relate the density δ_{NL} to the volume averaged correlation function $\bar{\xi}(r) = r^{-3} \int_0^r dx^3 \xi(x)$. In order to collapse, the cluster has to exceed a certain density threshold. Correspondingly this threshold exists in terms of $\bar{\xi}(r)$; more precisely, if the averaged correlation function exceeds a critical value the cluster collapses. We can rewrite (2.69) as

$$r_L = [1 + \bar{\xi}_{NL}(r_{NL})]^{\frac{1}{3}} r_{NL} \; . \tag{2.70}$$

Hamilton et al. (1991) conjectured the existence of a general relation between linear and non-linear correlation functions.

$$\bar{\xi}_{NL}(r_{NL}) = f_{NL}[\bar{\xi}_L(r_L)] \tag{2.71}$$

This assumption was verified by numerical simulations (for EdS and ΛCDM) and turns out to be very effective. From the correlation function one can calculate the corresponding (dimensionless) power spectrum (Peacock & Dodds 1994) and rewrite (2.70) as

$$k_L = [1 + \Delta_{NL}^2(k_{NL})]^{-\frac{1}{3}} k_{NL}^3 \; , \tag{2.72}$$

with $\Delta_{NL}^2 = k^3/(2\pi^2)P_\delta$ denoting the dimensionless power spectrum and k_{NL} is the wavenumber. When transforming $\bar{\xi}$ into Fourier space, one interpretes the volume averaged correlation function as a measure for the power at an effective scale k_{eff}. The corresponding relation to (2.71) in Fourier space reads

$$\Delta_{NL}^2(k_{NL}) = f_{NL}[\Delta_L^2(k_L)] \; . \tag{2.73}$$

The function f_{NL} is fitted from numerical simulations by Peacock & Dodds (1996).

Furthermore, Smith et al. (2003) develop a fit function for the non-linear power spectrum which is based on the halo model (see Cooray & Sheth 2002, for a review). Here, one assumes that all mass of the universe is contained in separated spherical or, as an improvement, triaxial objects, the so-called haloes. These originate from early overdensities in the Universe, which decouple from the overall expansion, finally collapse and form a halo. By assuming that the typical distance of two haloes is large compared to their extent, this model allows for the distinction of two extreme cases, the highly non-linear and the quasi-linear regime. Both are associated with a power spectrum, i.e. P_H for the first and P_Q for the second. The power spectrum P_H, which accounts for large k (non-linear scales), is determined solely through the assumed mass profiles
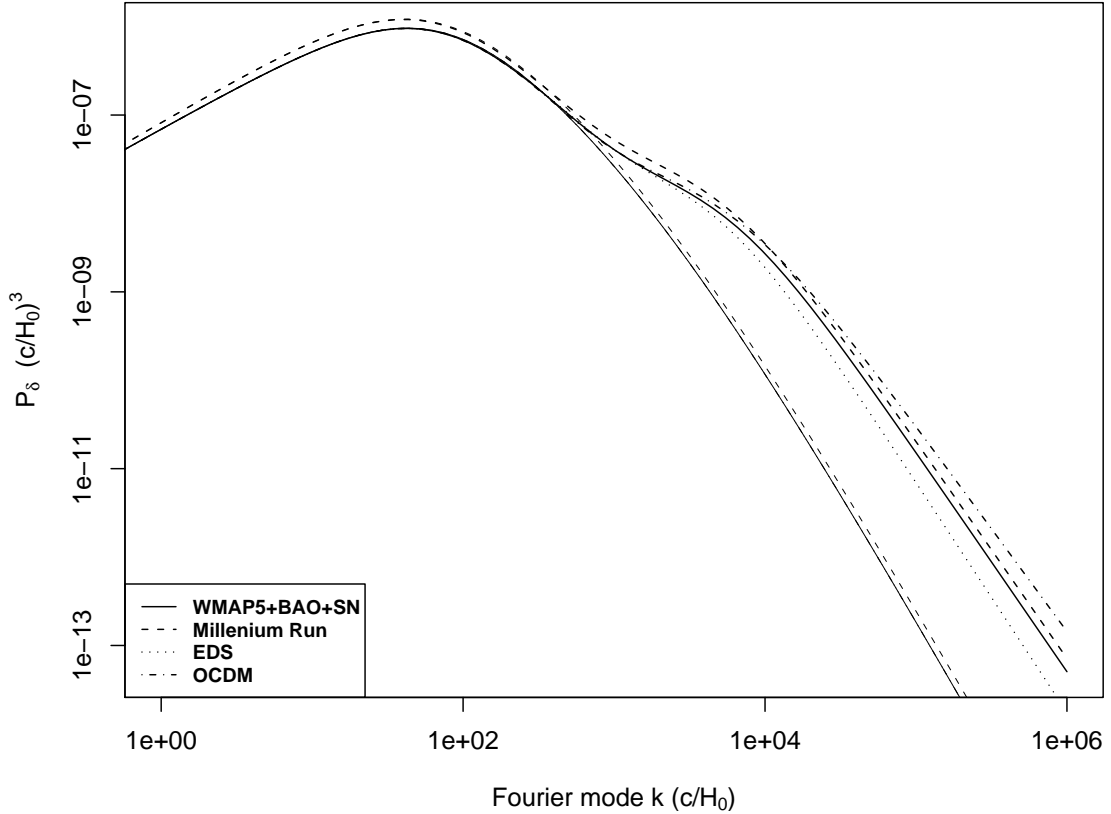
Figure 2.5: *The power spectrum of density fluctuations P_δ. Thin lines correspond to the linear, thick lines represent the non-linear P_δ (Smith et al. 2003). Solid and dotted lines correspond to a flat ΛCDM model, more precisely, to the cosmology favored by the WMAP 5-years analysis (table 2.1 last column) for the solid and to the cosmology of the Millennium simulation (Springel et al. 2005) for the dotted line. The dashed line corresponds to an EdS-universe, the dotted-dashed line to an open cold-dark-matter (OCDM) universe ($\Omega_m = 0.3, \Omega_\Lambda = 0$). For the EdS and the OCDM universe, all other parameters are similar to the WMAP5 results.*

of the haloes. In contrast, linear scales (small k), described by P_Q, are given by the spatial distribution of the haloes. Smith et al. (2003) combine the two power spectra into the final non-linear power spectrum

$$P_{NL}(k) = P_H(k) + P_Q(k) \tag{2.74}$$

Although based on the halo model the two individual terms are fitted to numerical simulations. Examples of the power spectrum for various cosmological are shown in Fig. 2.5. Note that the spatial mass distribution of the haloes needed to derive P_Q depends on the number density of haloes which again depends on cosmology; Press & Schechter (1974) give an analytic expression which quantifies the change in number density of haloes dn with respect to their mass M and the considered cosmic time t. This expression is based on the *spherical collapse model* (for a detailed

description see Schneider 2008) and reads

$$\frac{\mathrm{d}n}{\mathrm{d}M}(M,t) = -\frac{\sqrt{2}\Omega_{\mathrm{m}}\rho_{\mathrm{crit}}}{\sqrt{\pi}M}\frac{\delta_{\mathrm{c}0}(t)}{\sigma^2(R)}\frac{\mathrm{d}\sigma(R)}{\mathrm{d}M}\mathrm{e}^{-\frac{\delta_{\mathrm{c}0}^2(t)}{2\sigma^2(M))}} \; . \tag{2.75}$$

Here, $\sigma(R)$ is defined in (2.65), and $\delta_{\mathrm{c}0}(t)$ denotes the critical density contrast today (extrapolated using linear structure growth) which is needed for a perturbation to collapse before cosmic time t. The quantity $\delta_{\mathrm{c}0}(t)$ depends on cosmology, for an EdS-universe $\delta_{\mathrm{c}0}(t) = 1.68/a(t)$ holds. The original Press-Schechter theory underwent several improvements, e.g. by replacing the spherical with the ellipsoidal collapse, Sheth & Tormen (1999) provide a mass function with much better agreement to numerical simulations. However, their mass function becomes somewhat inaccurate in case of ΛCDM models. This deficit is removed in the fit-formula of Jenkins et al. (2001a) which shows excellent agreement to numerical simulations of a broad range of cosmologies and redshifts.

There are other attempts to derive a non-linear power spectrum, e.g. from perturbation theory (Bernardeau et al. 2002). Note, that this approach fails in the highly non-linear regime, also the fit-formula of Smith et al. (2003) is not sufficiently accurate for precision cosmology. We return to this issue in chapter 8.

2.2.2.4 Higher order moments - the bispectrum

Although the non-linear power spectrum partly describes non-Gaussian features in the density field, a precise treatment requires knowledge of higher-order moments. The lowest-order diagnostic after the power spectrum is the Fourier space equivalent to the three-point correlation function, the so-called bispectrum. Its definition reads

$$\left\langle \hat{\delta}(\boldsymbol{k}_1,t)\hat{\delta}(\boldsymbol{k}_3,t)\hat{\delta}(\boldsymbol{k}_3,t)\right\rangle = (2\pi)^3\delta_{\mathrm{D}}(\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3)\left(B_\delta(\boldsymbol{k}_1,\boldsymbol{k}_2,t) + B_\delta(\boldsymbol{k}_2,\boldsymbol{k}_3,t) + B_\delta(\boldsymbol{k}_3,\boldsymbol{k}_1,t)\right) \; . \tag{2.76}$$

Assuming an initial Gaussian density field the bispectrum solely originates from non-linear gravitational clustering; it can be approximated through second-order perturbation theory (for a detailed derivation see Bernardeau et al. 2002), explicitly

$$B_\delta(\boldsymbol{k}_1,\boldsymbol{k}_2,t) = F_2(k_1,k_2,\cos\varphi)\,P_\delta(k_1,t)\,P_\delta(k_2,t)\,, \tag{2.77}$$

where P_δ denotes the linear power spectrum and $\cos\varphi = (\boldsymbol{k}_1\boldsymbol{k}_2)/(k_1k_2)$. In contrast to the power spectrum, which only depends on the modulus of a wave vector, the bispectrum depends on three Fourier vectors. For the case of a EdS-universe the kernel $F_2(k_1,k_2,\cos\varphi)$ can be calculated analytically

$$F_2(k_1,k_2,\cos\varphi) = \frac{10}{7} + \cos\varphi\left(\frac{k_1}{k_2} + \frac{k_2}{k_1}\right) + \frac{4}{7}\cos^2\varphi \; . \tag{2.78}$$

The cosmology dependence of $F_2(k_1,k_2,\cos\varphi)$ is extremely weak and the bispectrum is proportional to P_δ^2, hence the reduced bispectrum

$$Q(k_1,k_2,\cos\varphi,t) = \frac{B_\delta(\boldsymbol{k}_1,\boldsymbol{k}_2,t) + B_\delta(\boldsymbol{k}_2,\boldsymbol{k}_3,t) + B_\delta(\boldsymbol{k}_3,\boldsymbol{k}_1,t)}{P_\delta(k_1,t)\,P_\delta(k_2,t) + P_\delta(k_2,t)\,P_\delta(k_3,t) + P_\delta(k_3,t)\,P_\delta(k_1,t)} \tag{2.79}$$

is almost independent of time and cosmology. It solely reflects the dependence on the configuration of the three Fourier vectors k_1, k_2, k_3. Note that the bispectrum is invariant under rotations or translations of this Fourier space triangle as a result of statistical homogeneity and isotropy of the density field.

Similar to the HKLM method there exist methods to calculate a non-linear model for the bispectrum of density fluctuations. This so-called hyper-extended perturbation theory (HEPT) was first developed for an EdS Universe (Scoccimarro & Frieman 1999) and later extended to various CDM models (Scoccimarro & Couchman 2001). In this theory, the coefficients of the kernel (2.78) are modified by amended fitting functions. These fitting functions are calibrated from numerical simulations, accounting for the non-linear evolution of the bispectrum with an average accuracy of 15%. Note that recently a new ansatz was proposed by Pan et al. (2007), who calculate the non-linear bispectrum as a function of the non-linear power spectrum, similar to the corresponding linear relation in (2.77).

2.3   Measurements of cosmological parameters

In the last 20 years the results on cosmological parameters progressed from vague estimates with large error bars to an impressive accuracy today, which will be improved even more with future high-precision data. In this section we outline the most important results and give references for further reading.

2.3.1   The Hubble constant

The first measurement of a cosmological parameter was performed by Edwin Hubble (Hubble 1929), who recognized that the recession velocity of nearby objects is proportional to their distance and quantified this relation via the Hubble constant H_0 (see Sect. 2.1.3). Hubble underestimated the influence of peculiar velocities of the observed objects and overestimated the Hubble constant, giving a value of $500 \ \mathrm{km \ s^{-1} \ Mpc^{-1}}$. More recently, the *HST Key Project* measured the period and brightness of cepheids in other galaxies and determined their luminosity distance using the fact that we can calculate the maximum luminosity from the period. Plotting the recession velocity of the observed galaxies against their distance we can determine the Hubble constant (Freedman et al. 2001). In addition to cepheids, Freedman et al. (2001) use various distance measures to determine the Hubble constant, e.g. Type Ia supernovae, the Tully-Fisher relation, surface brightness fluctuations, Type II supernovae, and the fundamental plane. They combine and weight the individual results and find good agreement with

$$H_0 = (72 \pm 8)\frac{\mathrm{km}}{\mathrm{s \ Mpc}} \,. \tag{2.80}$$

This value is confirmed by other experiments. For example, Riess et al. (2005) claim that earlier disagreement of SN Ia estimates of H_0 was only due to bad data, with their new data sample and

analysis they give

$$H_0 = 73 \pm 4(\text{statistical}) \pm 5(\text{systematic})\frac{\text{km}}{\text{s Mpc}} . \tag{2.81}$$

This agrees well with the result from WMAP. In their analysis H_0 is not measured directly, however under the assumption of zero curvature and the existence of a cosmological constant term, one can put the tightest constraints on H_0 (see Komatsu et al. 2008), i.e

$$H_0 = (70.1 \pm 1.3)\frac{\text{km}}{\text{s Mpc}} . \tag{2.82}$$

Last but not least, the Hubble constant can be obtained from the relative time delays between multiple images of a lensed quasar. From 10 of such multiple-image systems Kochanek & Schechter (2004) derive a Hubble constant of

$$H_0 = (71 \pm 3)\frac{\text{km}}{\text{s Mpc}} . \tag{2.83}$$

For more information on the use of time delays to estimate H_0 the reader is referred to the projects *COSMOGRAIL*[2] and *HOLIGRAIL* [3].

2.3.2 Density parameters from CMB

Due to small inhomogeneities at the time of recombination (see Sect. 2.1.5) one expects small anisotropies in the CMB. These temperature fluctuations are due to many effects, we distinguish *primary* and *secondary* anisotropies. The primary anisotropies result from physical processes before recombination, secondary anisotropies occur later, while photons propagate through the Universe (e.g. the Sunyaev-Zel'dovich-effect). The temperature fluctuations are expanded in spherical harmonics C_ℓ, where ℓ is the frequency mode which is inverse proportional to the angular scale. On scales larger than the horizon, the temperature fluctuations can be approximated analytically (Sachs & Wolfe 1967). The authors calculate that $\ell(\ell + 1)C_\ell$ is constant for small ℓ in an EdS-universe (*Sachs-Wolfe effect*). This is the reason why CMB scientists plot $\ell(\ell + 1)C_\ell$ against ℓ, instead of C_ℓ only (see Fig. 2.6). The most prominent features in the CMB power spectrum of temperature fluctuations are the *acoustic peaks*. Before recombination, the baryons and photons can be described as a perfect fluid. On scales smaller than the horizon, the baryon fluid is attracted towards the potential wells, formed by the dark matter perturbations. The radiation pressure counteracts this attraction, as a result the baryon-photon fluid starts to oscillate. When the photons decouple this oscillation is frozen into the CMB with a characteristic wavelength, the so-called *sound horizon* at the time of recombination $\lambda_{\max} \simeq t_{\text{rec}}c_{\text{s}}$. Due to the domination of the photons in the baryon-photon fluid, the sound speed is given by $c_{\text{s}} \simeq c/\sqrt{3}$, hence we can relate the angular scales of the sound horizon to the actual horizon via $\theta_{\text{s}} \simeq \theta_{\text{H}}/\sqrt{3}$. The size of the horizon is directly related to cosmological parameters; for a flat universe the sound horizon corresponds to an angular scale of $\theta_{\text{s}} \sim 1°$, which corresponds to $\ell \sim 200$ in spherical harmonics.

---

[2]http://www.cosmograil.org
[3]http://www.astro.uni-bonn.de/~holigrail
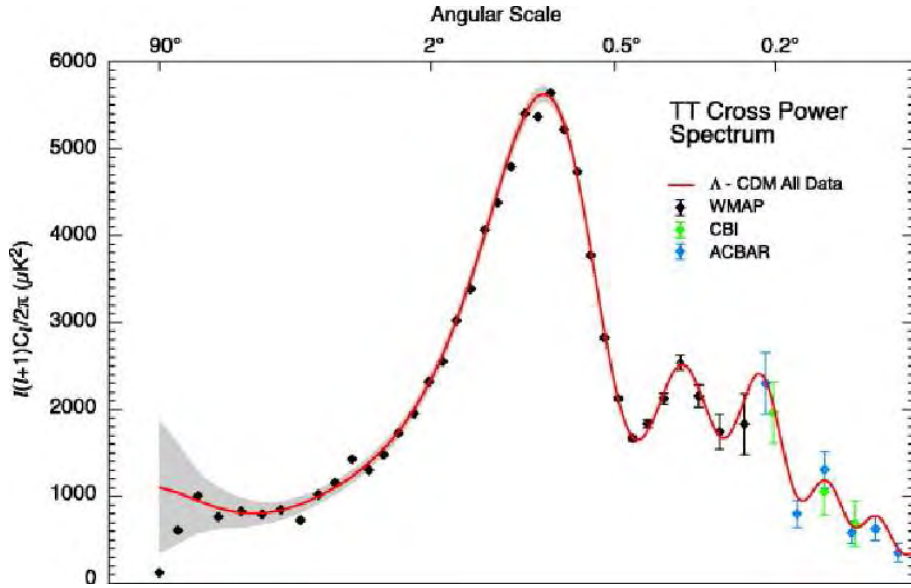
Figure 2.6: *The power spectrum of temperature fluctuation as obtained from WMAP. One nicely sees the acoustic peaks at $\ell \sim 200$ and integer multiples thereof. The shaded area to the left illustrates the uncertainty due to cosmic variance. (Figure from Hinshaw et al. 2003)*

Here, and on integer multiples of $\ell \sim 200$, we expect maxima in the CMB power spectrum. For a positive curvature, θ_s decreases, shifting the peaks in the power spectrum to higher ℓ, for negative curvature vice versa. This behavior and the dependence of the CMB temperature fluctuations on other parameters (Ω_Λ, Ω_b, Ω_m) are illustrated in Fig. 2.7. A change in Ω_Λ hardly has any impact which indicates that dark energy was not dominant at the time of recombination. Variations in Ω_b and Ω_m basically changes the amplitude of the acoustic peaks but hardly affect their positions. Note that the strongest constraints coming from the CMB are on the curvature; the assumption of a flat Universe is mainly justified by the CMB.

2.3.3   Galaxy surveys and Baryonic Acoustic Oscillations

In Sect. 2.2.1 we describe how the power spectrum of density fluctuations P_δ depends on cosmological parameters; models for the power spectrum can be obtained e.g. via the methods described in Sect. 2.2.2.3. By comparing these models to the measured power spectrum one can estimate cosmological parameters. Galaxy redshift surveys map the galaxy distribution and assume that this is, up to a constant (*bias*) factor, a good tracer for the dark matter distribution. The bias factor relates the power spectrum of the number density of galaxies to P_δ, i.e. $P_\delta = b^2 P_{gal}$. Strictly speaking, this bias factor cannot be derived from theory, nevertheless it seems reasonable in the linear regime but definitely breaks down on non-linear scales. With the definition of σ_8 (2.68), we see that $\sigma_8 = \sigma_{8,gal}/b$. The bias factor can be determined through numerical simulations by measuring the simulated dark matter distribution to the simulated galaxy distri-
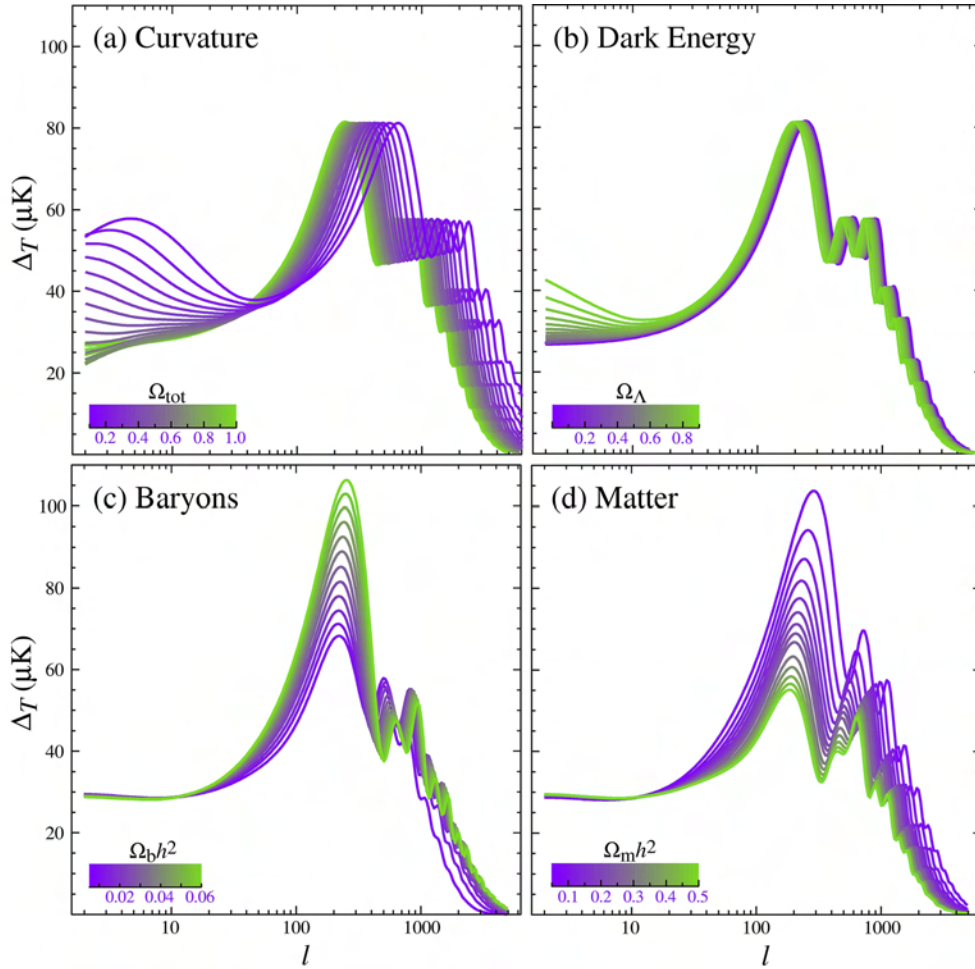
Figure 2.7: *Variation of the acoustic peaks in the anisotropy power spectrum of the CMB with respect to several parameters. (Figure from Hu & Dodelson 2002)*

bution. Hence, by measuring $\sigma_{8,gal}$ from data, one can directly constrain σ_8. Simon et al. (2007) use weak lensing data to constrain the bias factor arriving at $b \sim 0.8 \pm 0.11$. In addition to σ_8, one can determine the shape parameter Γ (see equation 2.60) from a measured P_δ, and thereby constrain Ω_m and Ω_b. The matter power spectrum is one of the most important quantities in cosmology, the level to which we can constrain P_δ quantifies our ability to constrain cosmological parameters. Note that there are several experiments which constrain P_δ (see Fig. 2.8), all with different sensitivities at different scales.

Very recently a new probe for cosmology has emerged from galaxy redshift surveys, the so-called *Baryonic Acoustic Oscillations (BAOs)*. Recall that the characteristic acoustic peaks in the CMB power spectrum occur because the baryons are driven out of the dark matter potential wells by the photon pressure. After recombination the separation of baryon overdensities remain separated from the dark matter overdensities by a characteristic comoving scale. As the Universe evolves both overdensities attract further matter, resulting in overdense regions. Hence, these

Figure 2.8: *The power spectrum of density fluctuations. This figure especially illustrates the variety of cosmological experiments which can constrain the power spectrum. Note that every experiment has a preferred scale at which its sensitivity is high. Only by combining several experiments we are able to probe all scales in the power spectrum. Figure from Max Tegmark's homepage.*
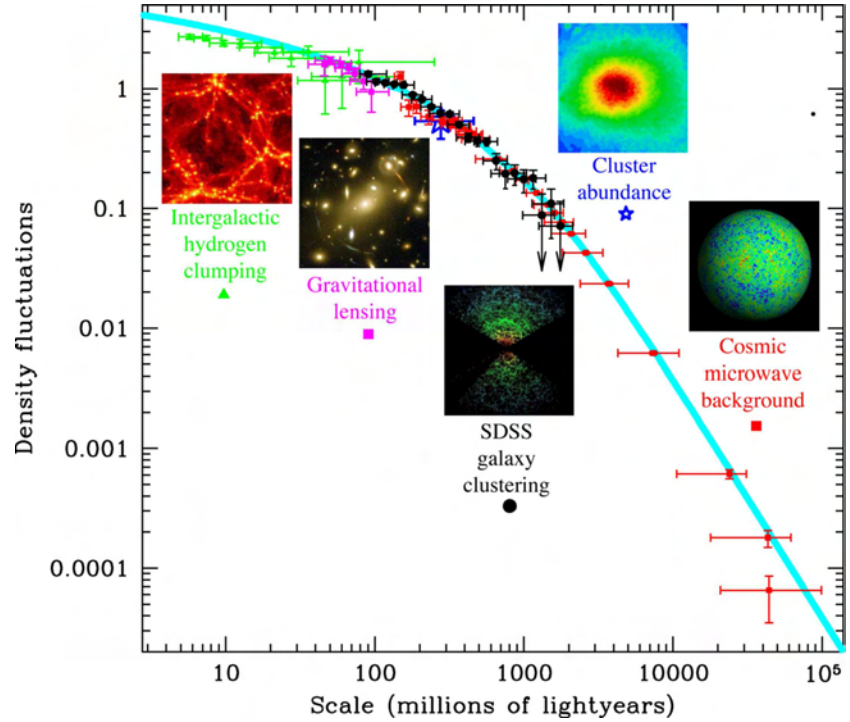
characteristic scales are also imprinted in today's matter (or galaxy) density power spectrum, resulting in the so-called baryonic acoustic peaks. Measuring the characteristic scales in the CMB temperature fluctuations power spectrum and in today's matter power spectrum provides a relation between angular diameter distance and redshift (a so-called "standard ruler"), which allows us to constrain the geometry and therefore the density parameters in the Universe. The first of these acoustic peaks has been detected in the matter power spectrum by Eisenstein et al. (2005); Cole et al. (2005); Huetsi (2005). Recently Gaztanaga et al. (2008a) claim to observe this peak also in the three-point correlation function of the galaxies. Gaztanaga et al. (2008b) measure the Hubble constant using BAOs with a result of $H_0 = 71.7 \pm 1.6 \frac{\mathrm{km}}{\mathrm{s\,Mpc}}$ and also combine BAOs with other cosmological probes to constrain the dark energy equation of state parameter $w = 0.96 \pm 0.05$. Note that this method strongly depends on the bias factor which again depends on the wave number (Sánchez & Cole 2008) and varies with redshift. A proper treatment of this issue is needed for precision cosmology constraints with BAO.

2.3.4 Clusters

Based on the model of spherical collapse, the Press-Schechter mass function predicts the number density of dark matter haloes depending on their mass and redshift. As mentioned earlier this model has been improved by numerical simulations (Jenkins et al. 2001b; Evrard et al. 2002); it depends strongly on cosmology, e.g. in a flat EdS-universe, the growth function has a smaller amplitude (see Fig. 2.3) compared to a ΛCDM model. Going back in redshift one therefore expects much less structure for an EdS-universe compared to a ΛCDM model. By comparing

the observations of cluster number density to the predictions of various models we can constrain cosmological parameters (Henry 2004; Kravtsov et al. 2006; Mantz et al. 2008). Allen et al. (2008) use the gas-to-mass fraction (f_{gas}) inside clusters to probe the accelerated expansion of the Universe. This method assumes that f_{gas} is constant in redshift, which has been checked for by numerical simulations. In addition, one can use f_{gas} to estimate Ω_m (Allen et al. 2008, and references therein). Based on the fact that the intra cluster gas contains the dominant fraction of baryonic matter, and assuming that this fraction is representative for the Universe, one can calculate $\Omega_m \simeq \Omega_b/f_{gas}$. The second assumption is justified by the facts that clusters are the largest bound structures in the Universe and that f_{gas} is almost constant for different clusters. Cluster samples can also be used to constrain the matter power spectrum, similar to galaxy surveys. This method is especially useful to constrain the large scales of the power spectrum (see Fig. 2.8).

2.3.5 Supernova Type Ia

We already mentioned in Sect. 2.1.6 that SN Ia experiments provided the first evidence for an accelerated expansion of the Universe and hence for dark energy (Riess et al. 1998; Perlmutter et al. 1999). A SN Ia is an explosion of a white dwarf which accretes mass and exceeds a critical mass limit, which is the same for all SN Ia. Therefore, one might assume that the luminosity of all SN Ia explosions is the same, meaning SN Ia are so-called "standard candles". In a strict sense this is not true, since the maximum luminosity of a SN Ia varies, however, there is a relation between shape of the lightcurve and the luminosity, which enables us to standardize the SN Ia. Hence, we can measure the luminosity distance of these supernovae, which evolves differently in redshift for different cosmological models (see Fig. 2.9). As a reference model we consider the evolution of the luminosity distance in an empty universe. For cosmological models without cosmological constant (e.g. an EdS universe) the luminosity distance will be lower. The fact that for $z \lesssim 1$ we observe a higher apparent magnitude compared to an empty universe can only be explained through an accelerated expansion and therefore implies a dark energy component. Since the first results in 1998, the search for SN Ia has been extended to a much larger sample (Astier et al. 2006) and to higher redshift, i.e. $z = 1.7$ (Kuznetsova et al. 2008). When going to these high redshifts, the apparent magnitude of SN Ia is lower compared to what one expects in an empty universe. This can be explained by the fact that the Universe at these redshifts was matter-dominated in this epoch. The main problem of SN Ia are evolutionary effects, such as a redshift-dependent critical mass limit for the explosion, which could results in a lower luminosity for high-redshift SNe Ia.

2.3.6 The concordance model

Table 2.1 shows the most recent highlight in cosmological parameter estimation, i.e. the 5 years data analysis from WMAP (Komatsu et al. 2008, and references therein). The table is divided into two parts, the upper describing parameter which are directly observable through WMAP, the lower contains parameters which are derived including prior information. The third column summarizes the constraints from a joint parameter estimate of WMAP, a combined SN type Ia sample
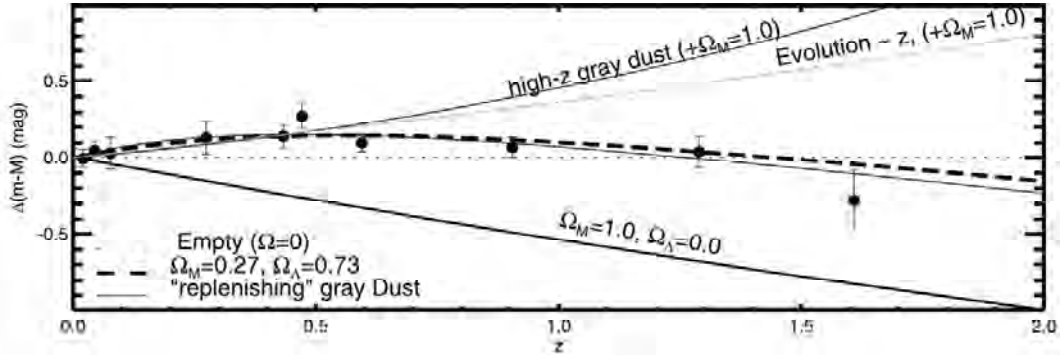
Figure 2.9: *This figure shows the difference between the distance modulus of SN Ia measurements and that expected in an empty universe as a function of redshift. This function can be predicted depending on the cosmological models, e.g. EdS (solid line) and a ΛCDM model (dashed line). The data points clearly favor the latter model. (From Riess et al. 2004)*

Table 2.1: *Cosmological parameters for a ΛCDM model from the WMAP 5 years analysis (Komatsu et al. 2008)*

Parameter	WMAP 5-year ML	WMAP 5-year Mean	WMAP+BAO+SN Mean
$100\Omega_b h^2$	2.268	2.273 ± 0.062	2.265 ± 0.059
[1] $\Omega_c h^2$	0.1081	0.1099 ± 0.0062	0.1143 ± 0.0034
Ω_Λ	0.751	0.742 ± 0.030	0.721 ± 0.015
n_s	0.961	$0.963^{+0.014}_{-0.015}$	$0.960^{+0.014}_{-0.013}$
σ_8	0.787	0.796 ± 0.036	0.817 ± 0.026
H_0	72.4 km/s/Mpc	$71.9^{+2.6}_{-2.7}$ km/s/Mpc	70.1 ± 1.3 km/s/Mpc
Ω_b	0.0432	0.0441 ± 0.0030	0.0462 ± 0.0015
Ω_c	0.206	0.214 ± 0.027	0.233 ± 0.013
$\Omega_m h^2$	0.1308	0.1326 ± 0.0063	0.1369 ± 0.0037

[1] Ω_c denotes the cold dark matter component of Ω_m.

(Riess et al. 2004, 2007; Astier et al. 2006; Wood-Vasey et al. 2007), and Baryonic Acoustic Oscillations from the SDSS and 2dFGRS (Percival et al. 2007). In chapter 7 we will refer to these estimates and the corresponding confidence intervals. Combining several cosmological probes improves the constraints significantly because different probes have different parameter degeneracies, which can be broken in a combined analysis. In addition to the cosmological probes described in this chapter, there exist several others, e.g. Sunyaev-Zel'dovich effect or the Lyman alpha forest, which we do not explain in detail. Also cosmic shear is not discussed in this chapter; we postpone this to chapter 3.3. However, all these cosmological probes agree on the cosmological model summarized in table 2.1, and this agreement is truly remarkable. Especially, if we take into account that all the aforementioned experiments test very different physical processes, at very different cosmic time, it is astonishing how well the ΛCDM model combined with

our picture of structure formation is able to explain the observations. We therefore consider this concordance model as a robust theory on which future cosmological projects should be based. Today's open questions mainly address the nature of dark matter and dark energy. Especially dark energy poses one of the most interesting mysteries, for which many future experiments are proposed. For a detailed analysis which of the different methods is most suitable to constrain the dark energy parameters, we refer the reader to Albrecht et al. (2006) and Peacock et al. (2006). Note, that cosmic shear is considered to be one of the most promising methods.

Chapter 3

Gravitational Lensing and Cosmic Shear

Based on Einstein's theory of General Relativity, gravitational lensing describes the behavior of light rays in a gravitational field. It can be used as a direct measure for matter distributions on all scales, starting from small masses like stars and galaxies up to large mass distributions like clusters of galaxies or the Large Scale Structure (LSS) of the Universe. Compared to other methods, gravitational lensing has the advantage to probe all types of matter directly, regardless whether it is dark or luminous. Depending on the mass of the considered lens, gravitational lensing is divided into two regimes. In *strong lensing*, high mass distributions (galaxies or clusters of galaxies) create multiple distorted and magnified images. In contrast, *weak lensing* deals with numerous background sources which are also distorted and magnified but the effect is much smaller compared to strong lensing. For this reason weak lensing must be studied statistically by averaging over a large number of images. This chapter starts with the basic theory of gravitational lensing. We briefly explain strong lensing, but focus in much more detail on weak lensing, in particular on weak gravitational lensing by the LSS, called *cosmic shear*.

3.1 Deflection of light

Consider a light ray which is bend in the gravitational potential of a point mass M; the deflection angle of this ray can be calculated as

$$\hat{\alpha}(\xi) = \frac{4G}{c^2} \frac{M}{\xi} = \frac{2R_s}{\xi} \, , \tag{3.1}$$

where ξ is the impact parameter and R_s the Schwarzschild radius of the mass M. This equation only holds for a small deflection angle $\hat{\alpha} \ll 1$ which is always true in case of a weak Newtonian gravitational potential ($\phi/c^2 \ll 1$). The deflection angle caused by a mass distribution can be calculated by the vectorial sum of the deflections caused by the individual mass elements. Assuming a small deflection angle (*weak field assumption*), and a mass distribution with an extent much smaller than the distances between source, lens and observer (*thin lens approximation*) the mass distribution of the lens is characterized by the *surface mass density*

$$\Sigma(\boldsymbol{\xi}) := \int \mathrm{d}r_3 \, \rho(\xi_1, \xi_2, r_3) \, . \tag{3.2}$$

Here, we choose coordinates such that r_3 points towards the line of sight and $\boldsymbol{\xi} = (\xi_1, \xi_2)$ being a vector in the plane perpendicular to r_3 (see Fig. 3.1). $\rho(\xi_1, \xi_2, r_3)$ is the volume density. Integration over the individual mass elements at $\mathbf{r}' = (\xi_1', \xi_2', r_3)$ and inserting the definition of the surface mass density gives

$$
\begin{aligned}
\hat{\alpha}(\boldsymbol{\xi}) &= \frac{4G}{c^2} \sum dm\, (\xi_1', \xi_2', r_3') \frac{\boldsymbol{\xi} - \boldsymbol{\xi}'}{|\boldsymbol{\xi}' - \boldsymbol{\xi}'|^2} \\
&= \frac{4G}{c^2} \int d^2\xi' \int dr_3'\, \rho(\xi_1', \xi_2', r_3') \frac{\boldsymbol{\xi} - \boldsymbol{\xi}'}{|\boldsymbol{\xi}' - \boldsymbol{\xi}'|^2} \\
&= \frac{4G}{c^2} \int d^2\xi'\, \Sigma(\boldsymbol{\xi}') \frac{\boldsymbol{\xi} - \boldsymbol{\xi}'}{|\boldsymbol{\xi}' - \boldsymbol{\xi}'|^2} .
\end{aligned}
\tag{3.3}
$$

This equation can be applied to galaxies and clusters of galaxies which fulfill the condition of the thin lens approximation, but it becomes inaccurate for cosmic shear.

3.1.1 Lens equation
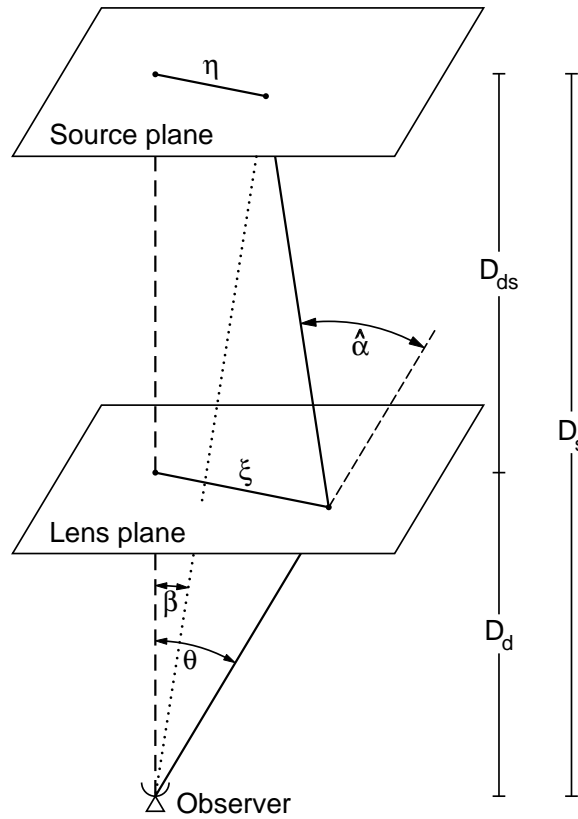
Figure 3.1: *This figure illustrates a typical lensing situation as described in the text. Note that $\boldsymbol{\xi}, \boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\theta}, \hat{\boldsymbol{\alpha}}$ are two-component quantities. (Figure from Schneider et al. 2006)*

Figure 3.1 shows a light source at distance D_s from the observer and a gravitational lens lying on the line of sight at distance D_d from the observer. Looking for a relation between the true position $\boldsymbol{\eta}$ of the light source in the *source plane* and the observed position $\boldsymbol{\xi}$ of the image in the *lens plane* we find

$$\boldsymbol{\eta} = \frac{D_s}{D_d}\boldsymbol{\xi} - D_{ds}\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) \ . \tag{3.4}$$

Using

$$\boldsymbol{\eta} = D_s \boldsymbol{\beta} \qquad \text{and} \qquad \boldsymbol{\xi} = D_d \boldsymbol{\theta} \ , \tag{3.5}$$

we are able to express the lens equation in terms of the angular positions

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \frac{D_{ds}}{D_s}\hat{\boldsymbol{\alpha}}(D_d\boldsymbol{\theta}) \equiv \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta}) \ , \tag{3.6}$$

where $\boldsymbol{\alpha}(\boldsymbol{\theta})$ is the so-called *scaled deflection angle*. From the observer's point of view, $\boldsymbol{\beta}$ is the angular position of the unlensed source and $\boldsymbol{\theta}$ is the angular image position. The deflection angle only depends on surface mass density and impact parameter. It is possible to have multiple images for a source, corresponding to multiple solutions $\boldsymbol{\theta}$ for fixed $\boldsymbol{\beta}$ in the lens equation. The decisive quantity for the occurrencemultiple images is the *critical surface mass density*

$$\Sigma_{cr} = \frac{c^2}{4\pi G}\frac{D_s}{D_d D_{ds}} \ , \tag{3.7}$$

which is used to define the *convergence* κ

$$\kappa(\boldsymbol{\theta}) := \frac{\Sigma(D_d\boldsymbol{\theta})}{\Sigma_{cr}} \ . \tag{3.8}$$

Multiple images become possible, if the surface mass density exceeds the critical surface mass density, which is obviously the case for $\kappa > 1$. Lenses with $\kappa > 1$ are called strong lenses whereas for $\kappa \ll 1$ we are in the weak lensing regime and multiple images are impossible. We write the scaled deflection angle in terms of κ

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{1}{\pi}\int d^2\theta'\kappa(\boldsymbol{\theta}')\frac{\boldsymbol{\theta} - \boldsymbol{\theta}'}{|\boldsymbol{\theta} - \boldsymbol{\theta}'|^2} \ , \tag{3.9}$$

and by introducing the deflection potential ψ,

$$\psi(\boldsymbol{\theta}) = \frac{1}{\pi}\int d^2\theta' \, \kappa(\boldsymbol{\theta}') \, \ln|\boldsymbol{\theta} - \boldsymbol{\theta}'| \ , \tag{3.10}$$

we express $\boldsymbol{\alpha}$ in terms of the deflection potential

$$\boldsymbol{\alpha} = \nabla\psi \ . \tag{3.11}$$

The two-dimensional Poisson equation gives the convergence κ as second derivative of the deflection potential

$$\kappa = \frac{1}{2}\nabla^2\psi \tag{3.12}$$

and can be derived from (3.10) using $\nabla^2 \ln|\boldsymbol{\theta}| = 2\pi\delta_D(\boldsymbol{\theta})$.

3.1.2   Distortion and magnification

Liouville's theorem guarantees conservation of the surface brightness in the absence of absorption or emission of photons. Hence, we deduce $I(\boldsymbol{\theta}) = I^{(s)}[\boldsymbol{\beta}(\boldsymbol{\theta})]$ with $I(\boldsymbol{\theta})$ and $I^{(s)}(\boldsymbol{\beta})$ denoting the radiation intensity of image and source. Assuming that the angular diameter of the source is small compared to the scale on which the density of the lens changes, we can locally linearize the lens mapping and describe the surface brightness of an image around a fixed point $\boldsymbol{\theta}_0$ by

$$I(\boldsymbol{\theta}) = I^{(s)}[\boldsymbol{\beta}(\boldsymbol{\theta}_0) + \mathcal{A}(\boldsymbol{\theta}_0) \cdot (\boldsymbol{\theta} - \boldsymbol{\theta}_0)] \ . \tag{3.13}$$

Distortion and magnification are then given by the Jacobian of the lens equation (3.6)

$$\mathcal{A} = \frac{\partial \boldsymbol{\beta}}{\partial \boldsymbol{\theta}} = \left(\delta_{ij} - \frac{\partial \alpha_i(\boldsymbol{\theta})}{\partial \theta_j} \right) = \left(\delta_{ij} - \frac{\partial^2 \psi(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} \right) . \tag{3.14}$$

We define the *shear* as a complex number $\gamma := \gamma_1 + i\gamma_2 = |\gamma|e^{2i\varphi}$, and the shear components are related to the deflection potential as

$$\gamma_1 = \frac{1}{2} (\partial_1 \partial_1 \psi - \partial_2 \partial_2 \psi) \qquad \text{and} \qquad \gamma_2 = \partial_1 \partial_2 \psi \ , \tag{3.15}$$

Using (3.15) and (3.12) the Jacobian matrix $\mathcal{A}$ can be parameterized through convergence and shear

$$\mathcal{A} = \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix} . \tag{3.16}$$

To illustrate the meaning of κ and γ we decompose $\mathcal{A}$ into a diagonal and a trace-free part

$$\mathcal{A}(\boldsymbol{\theta}) = (1 - \kappa) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \gamma \begin{pmatrix} \cos(2\varphi) & \sin(2\varphi) \\ \sin(2\varphi) & -\cos(2\varphi) \end{pmatrix} . \tag{3.17}$$

The convergence κ magnifies the image isotropically, whereas γ maps a circular source onto an elliptical image (Fig. 3.3). The ratio of the semi-axes of the ellipse is determined by the eigenvalues of $\mathcal{A}$. As mentioned at the beginning of this section the surface brightness is conserved, however the image's shape and size is distorted. This results in a (de-)magnification, which we define as the ratio of image flux S [given by the integral over $I(\boldsymbol{\theta})$] to source flux S_0 [given by the integral over $I^{(s)}(\boldsymbol{\beta})$]. This ratio is calculated to be the inverse of the determinant of $\mathcal{A}$

$$\mu = \frac{S}{S_0} = \frac{1}{\det \mathcal{A}} = \frac{1}{(1 - \kappa)^2 - |\gamma|^2} \ . \tag{3.18}$$

The magnification is given by $|\mu|$ but μ in general can have either sign; the sign expresses the parity of the image with respect to the unlensed source.

3.1.3 Caustics and critical curves

Curves in the lens plane where the determinant of the Jacobian matrix det $\mathcal{A}(\boldsymbol{\theta})$ vanishes are called *critical curves*. Mapping a critical curve to the source plane using the lens equation gives a so-called *caustic*. Whenever a source crosses a caustic two images in the lens plane are either created or destroyed. Every source close to and inside[1] a caustic causes two highly magnified images, one on each side of the critical curve. As proved by Burke (1981) the absolute number of images created by a lens must be odd, although observations of most lens systems show an even number. This is due to one highly de-magnified image in most lens systems which is difficult to observe. In contrast to critical curves which are always smooth, caustics have cusps. A source just inside and very close to such a cusp produces three highly magnified images. In case of an extended source lying exactly on a caustic its images will merge. The resulting image is highly magnified; this effect leads to the giant luminous arcs which we observe for example in the inner region of the galaxy cluster Abell 1689 (Fig. 3.2).

3.2 Weak lensing

The weak lensing regime is characterized by $\kappa \ll 1$ and $|\gamma| \ll 1$. Here, the Jacobian matrix $\mathcal{A}$ is close to the unit matrix, distortions and magnifications are much harder to identify. To detect the weak lensing effect we need a statistical approach, i.e. we have to consider a large sample of galaxy images, from which we determine the shapes. First, we introduce the *reduced shear*

$$g \equiv \frac{\gamma}{1 - \kappa} = \frac{|\gamma|}{1 - \kappa} e^{2i\varphi} , \tag{3.19}$$

which is a complex quantity and describes the degree of distortion. The phase φ gives the orientation of the distorted image. The Jacobian $\mathcal{A}$ can be expressed in terms of g

$$\mathcal{A} = (1 - \kappa) \begin{pmatrix} 1 - g_1 & -g_2 \\ -g_2 & 1 + g_1 \end{pmatrix}. \tag{3.20}$$

If the sources were circular their images would be ellipses with a ratio of axes given by (see Fig. 3.3)

$$\frac{a}{b} = \frac{1 - |g|}{1 + |g|} . \tag{3.21}$$

Unfortunately, the source galaxies are not intrinsically round. The image that we observe must be decomposed into intrinsic ellipticity $\epsilon^{(s)}$ and the distortion as a result from lensing. In order to measure the ellipticity of an image and to relate it to the source ellipticity, we define the center of an object with brightness distribution $I(\boldsymbol{\theta})$ on the sky as

$$\bar{\boldsymbol{\theta}} := \frac{\int \mathrm{d}^2\theta \, I(\boldsymbol{\theta}) \, q_I[I(\boldsymbol{\theta})] \, \boldsymbol{\theta}}{\int \mathrm{d}^2\theta \, I(\boldsymbol{\theta}) \, q_I[I(\boldsymbol{\theta})]} , \tag{3.22}$$

---

[1]"inside" means the side of a caustic where the number of produced images is larger
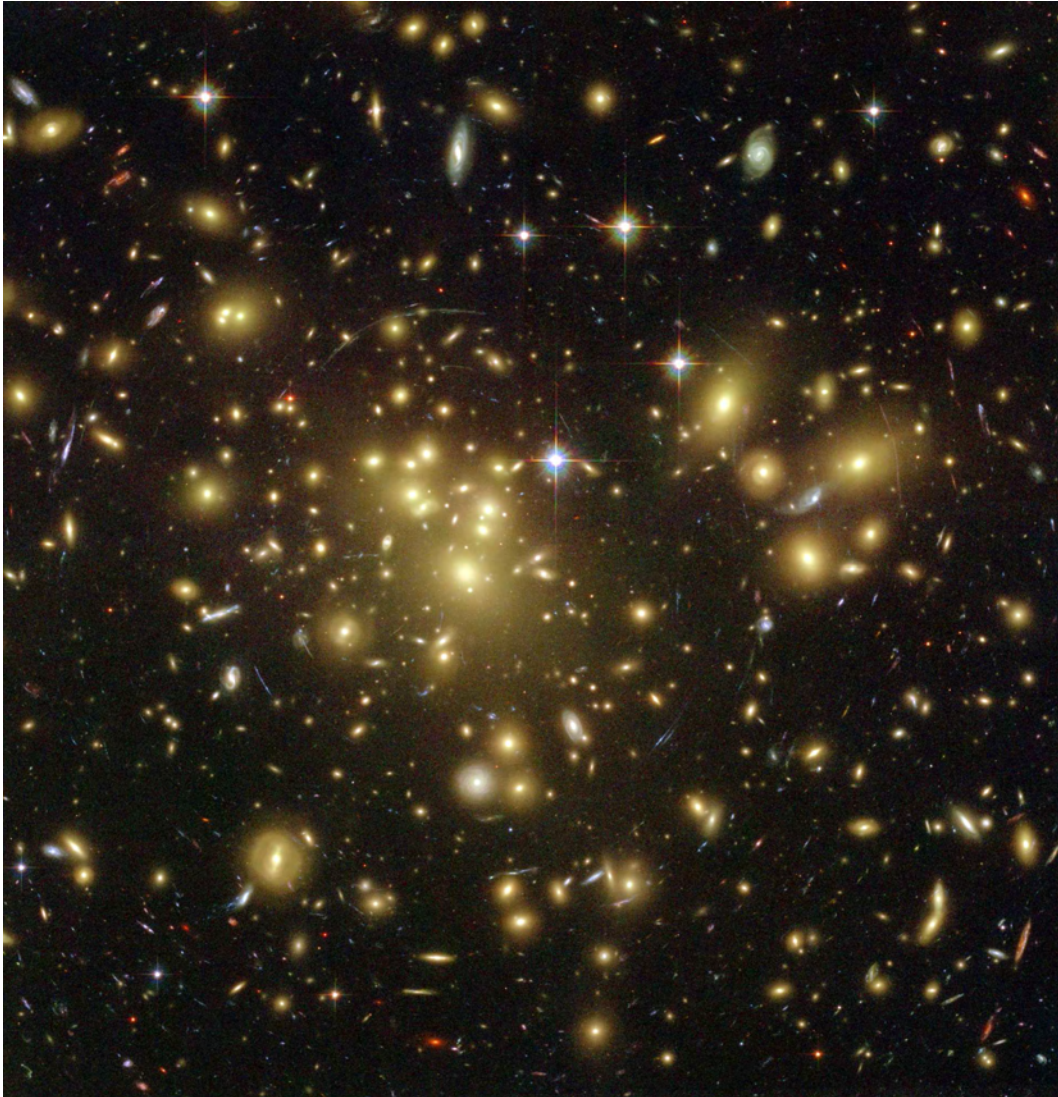
Figure 3.2: *This image of the inner region of Abell 1689 was taken with the HST/ACS camera. One can see the luminous arcs which are highly magnified and distorted images of background galaxies on a caustic. (Figure credits to the Space Telescope Science Institute)*

where $q_I(I)$ is a weight function in order to suppress the noise coming from the brightness of the surrounding area. Furthermore, we define the tensor of second brightness moments

$$Q_{ij} := \frac{\int \mathrm{d}^2\theta \, I(\boldsymbol{\theta}) \, q_I[I(\boldsymbol{\theta})] \, (\theta_i - \bar{\theta}_i)(\theta_j - \bar{\theta}_j)}{\int \mathrm{d}^2\theta \, I(\boldsymbol{\theta}) \, q_I[I(\boldsymbol{\theta})]} \quad , \qquad i, j \in \{1, 2\} \, . \qquad (3.23)$$
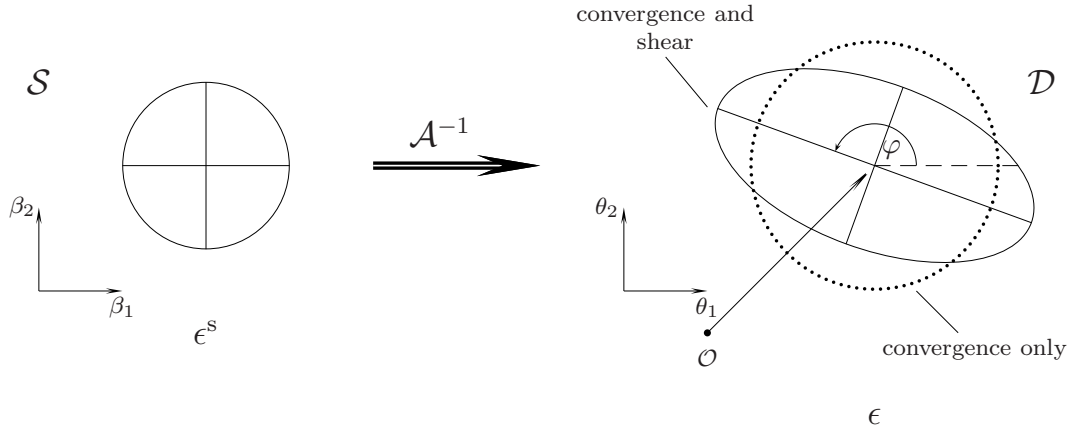
Figure 3.3: *This illustration shows the mapping in case of a circular source. Considering only the effect of the convergence the source is simply magnified to a larger circle. The shear additionally distorts the image to an ellipse depending on the phase of the shear φ. (Figure from Marusa Bradac)*

We can now quantify the complex ellipticity of a galaxy in terms of the Q_{ij} as

$$\epsilon = \epsilon_1 + \mathrm{i}\epsilon_2 := \frac{Q_{11} - Q_{22} + 2\mathrm{i}Q_{12}}{Q_{11} + Q_{22} + 2\left(Q_{11}Q_{22} - Q_{12}^2\right)^{1/2}} \; , \tag{3.24}$$

and

$$\chi = \chi_1 + \mathrm{i}\chi_2 = \frac{Q_{11} - Q_{22} + 2\mathrm{i}Q_{12}}{Q_{11} + Q_{22}} \; , \tag{3.25}$$

which are related to each other by

$$\epsilon = \frac{\chi}{1 + (1 - |\chi|^2)^{\frac{1}{2}}} \; , \qquad \chi = \frac{2\epsilon}{1 + |\epsilon|^2} \; . \tag{3.26}$$

Using the second-order brightness tensor for the unlensed source and the angular coordinates β

$$Q_{ij}^{(\mathrm{s})} = \frac{\int \mathrm{d}^2\beta \; I^{(\mathrm{s})}(\theta) \; q_I[I^{(\mathrm{s})}(\beta)] \; (\beta_i - \bar{\beta}_i)(\beta_j - \bar{\beta}_j)}{\int \mathrm{d}^2\beta \; I^{(\mathrm{s})}(\theta) \; q_I[I^{(\mathrm{s})}(\beta)]} \; , \qquad i, j \in \{1, 2\} \tag{3.27}$$

we define a complex ellipticity for the source $\epsilon^{(\mathrm{s})}$ and $\chi^{(\mathrm{s})}$, similar to (3.24, 3.25). The relation between source and image tensor then reads

$$Q^{(\mathrm{s})} = \mathcal{A}(\bar{\theta}) \, Q \, \mathcal{A}(\bar{\theta}). \tag{3.28}$$

Inserting the above relation into the definition of the ellipticities Seitz & Schneider (1995, 1997) derive the following relation between source and observed ellipticities

$$
\epsilon^{(s)} = \begin{cases} \dfrac{\epsilon - g}{1 - g^*\epsilon} & \text{for} \quad |g| \le 1 \\[2ex] \dfrac{1 - g\epsilon^*}{\epsilon^* - g^*} & \text{for} \quad |g| > 1 \end{cases} \qquad\qquad \chi^{(s)} = \frac{\chi - 2g + g^2\chi^*}{1 + |g|^2 - 2\mathrm{Re}(g\chi^*)} \, . \tag{3.29}
$$

We assume that there is no preferred intrinsic orientation of galaxy ellipticities in the Universe. Therefore, the expectation value of the source ellipticities should be zero

$$
\left\langle \epsilon^{(s)} \right\rangle = 0 = \left\langle \chi^{(s)} \right\rangle . \tag{3.30}
$$

With the above relation the expectation value of the observed ellipticity ϵ can be calculated as (Schramm & Kayser 1995; Seitz & Schneider 1997)

$$
\langle \epsilon \rangle = \begin{cases} g & \text{for} \quad |g| \le 1 \\[2ex] \dfrac{1}{g^*} & \text{for} \quad |g| > 1 \end{cases} . \tag{3.31}
$$

In the weak lensing regime ($\kappa \ll 1$ and $|\gamma| \ll 1$) the expression for the reduced shear (3.19) reduces to $g \approx \gamma$. Similarly, Seitz & Schneider (1995) have shown that in the weak lensing regime $1/2 \langle \chi \rangle = \langle \epsilon \rangle$ holds, which implies that any measured galaxy ellipticity is a (very noisy) measure of the local shear. The noise is given by the dispersion of the *intrinsic ellipticity dispersion*, i.e.

$$
\sigma_\epsilon = \sqrt{\langle \epsilon^{(s)} \epsilon^{(s)} \rangle}. \tag{3.32}
$$

By averaging over N galaxy images, which were all distorted by the same local shear, σ_ϵ is reduced by a factor of $\sqrt{N}$. The value, $\sigma_\epsilon / \sqrt{N}$ denotes the 1-σ deviation of observed mean ellipticity from true shear.

3.3   Cosmic Shear

In contrast to the last sections, where we describe the lensing effect of local mass distributions like galaxies or clusters, cosmic shear is subject to light deflection of the large-scale structure (LSS) in the Universe itself. In particular, the concept of a thin lens fails for cosmic shear and must be replaced by an extended three-dimensional matter distribution. Light bundles emitted from distant galaxies travel through this inhomogeneous matter distribution and are continuously distorted. These distortions in shape and size of the galaxy images can be measured statistically and provide information on the LSS and on cosmological parameters.

In the following we briefly describe the theory of cosmic shear, focussing on its various measures, their interrelations and covariances. Cosmic shear has progressed into a very important tool in observational cosmology during the recent years. We therefore conclude this section with a short review of its main achievements, future prospects and current challenges.
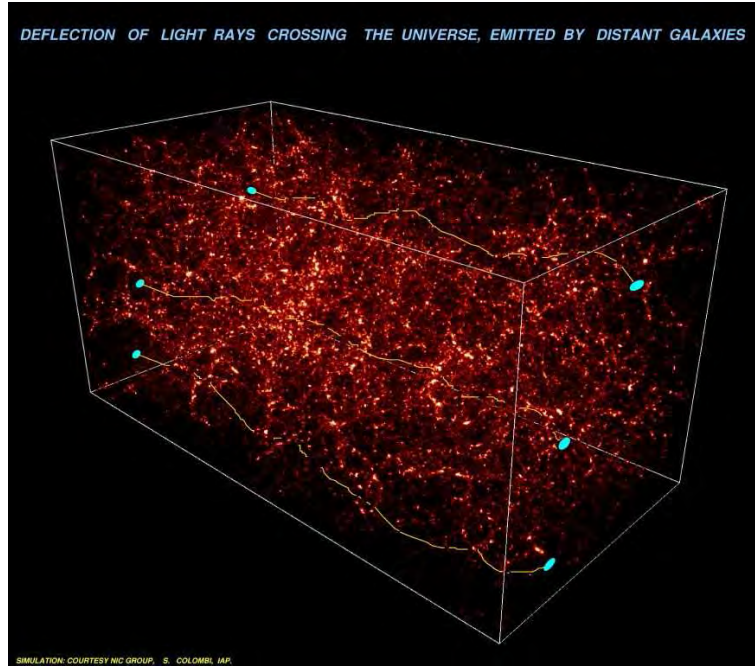
Figure 3.4: *This figure illustrates the principle of cosmic shear. Light rays from source galaxies are continuously deflected by the inhomogeneous matter distribution in the Universe. The observed images are therefore distorted. (Image credit: DESCART project at IAP, France)*

3.3.1 Propagation of light in a 3-D matter distribution

This section reviews the technical details of lensing effects in an inhomogeneous matter distribution; for a more thorough treatment on the derivation we refer the reader to Bartelmann & Schneider (2001) or Schneider et al. (2006).

In order to investigate the propagation of light we consider a situation as illustrated in Fig. 3.5 and examine, how a comoving separation vector $\mathbf{x}(\boldsymbol{\theta}, \chi)$ of two light rays evolves. We only consider weak gravitational perturbations; for this case the propagation equation for $\mathbf{x}(\boldsymbol{\theta}, \chi)$ reads

$$\frac{\mathrm{d}^2 \mathbf{x}}{\mathrm{d}\chi^2} + K\mathbf{x} = -\frac{2}{c^2} \left[\nabla_\perp \Phi(\mathbf{x}(\boldsymbol{\theta}, \chi), \chi) - \nabla_\perp \Phi^{(0)}(\chi) \right] , \tag{3.33}$$

where K is the spatial curvature of the Universe as defined in (2.14), $\nabla_\perp$ is the transverse comoving gradient operator and $\Phi^{(0)}(\chi)$ denotes the Newtonian potential along the fiducial ray. An exact derivation of (3.33) is given in Bartelmann & Schneider (2001). This differential equation can be solved using the Green's function which leads to

$$\mathbf{x}(\boldsymbol{\theta}, \chi) = f_K(\chi)\boldsymbol{\theta} - \frac{2}{c^2} \int_0^\chi \mathrm{d}\chi' f_K(\chi - \chi') \left[\nabla_\perp \Phi(\mathbf{x}(\boldsymbol{\theta}, \chi'), \chi') - \nabla_\perp \Phi^{(0)}(\chi') \right] . \tag{3.34}$$

Inserting the comoving separation vector between two light rays $\boldsymbol{x}$ into (2.21) we obtain the angular separation of the unlensed source and the fiducial ray $\boldsymbol{\beta} = \boldsymbol{x}/f_K(\chi)$. In order to describe
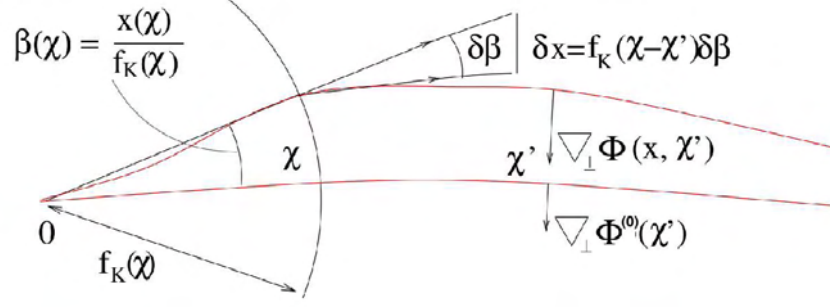
Figure 3.5: *For a given comoving radial coordinate χ, a light ray is separated from the fiducial light ray (lower one) by $\mathbf{x}(\chi)$. This distance corresponds to an angular separation of $\boldsymbol{\beta}$ if there are no further lensing effects on the way of the light ray from w to the observer. Both light rays are constantly influenced by a small Newtonian potential and differences in the deflection refer to different transverse gradients of the potential $\nabla_{\perp}\Phi$. (Figure from Tim Schrabback)*

the lensing effects we define the Jacobian (similar, to the definition in standard lens theory (3.14)) as

$$\mathcal{A}(\boldsymbol{\theta},\chi) = \frac{\partial\boldsymbol{\beta}}{\partial\boldsymbol{\theta}} = \frac{1}{f_{\mathrm{K}}(\chi)}\frac{\partial\mathbf{x}}{\partial\boldsymbol{\theta}} \, . \tag{3.35}$$

Inserting (3.34) into (3.35), expanding $\mathcal{A}$ in powers of Φ and considering only up to linear terms in Φ (*Born approximation*), we derive

$$\mathcal{A}_{ij}(\boldsymbol{\theta},\chi) = \delta_{ij} - \frac{2}{c^2}\int_0^{\chi}\mathrm{d}\chi'\frac{f_{\mathrm{K}}(\chi-\chi')f_{\mathrm{K}}(\chi')}{f_{\mathrm{K}}(\chi)}\frac{\partial^2\Phi_{ij}(\mathbf{x}=f_{\mathrm{K}}(\chi')\boldsymbol{\theta},\chi')}{\partial x_i\partial x_j} \, . \tag{3.36}$$

With the definition of the deflection potential

$$\psi(\boldsymbol{\theta},\chi) := \frac{2}{c^2}\int_0^{\chi}\mathrm{d}\chi'\frac{f_{\mathrm{K}}(\chi-\chi')}{f_{\mathrm{K}}(\chi)f_{\mathrm{K}}(\chi')}\Phi(f_{\mathrm{K}}(\chi')\boldsymbol{\theta},\chi') \, , \tag{3.37}$$

we find the same expression for the distortion matrix as in (3.14)

$$\mathcal{A}_{ij} = \delta_{ij} - \frac{\partial^2\psi}{\partial\theta_i\partial\theta_j}. \tag{3.38}$$

We see that, when using the weak-field metric, lensing by a 3D matter distribution can be described through the deflection potential ψ, similar to standard lens theory. The expressions for κ and γ also correspond to the standard lensing definitions (3.12) and (3.15).

3.3.2   Power spectrum of convergence

In Sect. 2.2 we show that the field of density fluctuations depends strongly on cosmological parameters. To constrain the latter with cosmic shear, we have to relate shear quantities (in

this case κ) to the matter density field. Starting from the definition of κ (3.12) we calculate the 2D-Laplacian of ψ (3.37)

$$\nabla^2\psi(\boldsymbol{\theta},\chi) = \frac{2}{c^2}\int_0^\chi \mathrm{d}\chi' \frac{f_\mathrm{K}(\chi-\chi')f_\mathrm{K}(\chi')}{f_\mathrm{K}(\chi)}\nabla^2\Phi(f_\mathrm{K}(\chi')\boldsymbol{\theta},\chi') \,. \tag{3.39}$$

White & Hu (2000) show that the two-dimensional Laplacian can be extended to a three-dimensional one, without inducing a significant error. Adopting this procedure we apply the Poisson equation in comoving coordinates (2.47) and derive

$$\kappa(\boldsymbol{\theta},\chi) = \frac{3H_0^2\Omega_\mathrm{m}}{2c^2}\int_0^\chi \mathrm{d}\chi' \, \frac{f_K(\chi-\chi')f_K(\chi')}{f_K(\chi)}\frac{\delta(f_K(\chi')\boldsymbol{\theta},\chi')}{a(\chi')} \,. \tag{3.40}$$

As the source galaxies are not located at one redshift but follow a redshift distribution $p_z(z)\mathrm{d}z = p_\chi(\chi)\mathrm{d}\chi$, we have to integrate κ over χ to obtain the effective convergence,

$$\begin{aligned} \kappa(\boldsymbol{\theta}) &= \int \mathrm{d}\chi\, p_\chi(\chi)\kappa(\boldsymbol{\theta},\chi) \\ &= \frac{3H_0^2\Omega_\mathrm{m}}{2c^2}\int_0^{\chi_\mathrm{h}}\mathrm{d}\chi\, g(\chi)f_K(\chi)\frac{\delta(f_K(\chi)\boldsymbol{\theta},\chi)}{a(\chi)}, \end{aligned} \tag{3.41}$$

with

$$g(\chi) = \int_\chi^{\chi_\mathrm{h}}\mathrm{d}\chi'\, p_\chi(\chi')\frac{f_\mathrm{K}(\chi'-\chi)}{f_\mathrm{K}(\chi')}. \tag{3.42}$$

The factor $g(\chi)$ describes the redshift-weighted efficiency of the density fluctuations just like $D_\mathrm{ds}/D_\mathrm{s}$ in "normal" lens theory (Sect. 3.1.1). The upper limit of the integral is the comoving horizon defined in (2.24). We have seen in Sect. 2.2.2.1 that the power spectrum of density fluctuations P_δ contains all second-order information about the matter density field. In (3.41) we related the convergence to the matter density; next we seek a similar relation for the power spectrum.

If δ is an isotropic and homogenous 3-D random field, the projections

$$g_i(\boldsymbol{\theta}) = \int \mathrm{d}\chi\, q_i(\chi)\,\delta(f_K(\chi)\boldsymbol{\theta},\chi) \tag{3.43}$$

are also isotropic and homogenous random fields in 2-D (q_i are weight functions). Limber's equation originally relates the correlation function of δ (3D) to a line-of-sight projected correlation function (2D). Correspondingly there exists a Fourier space version of Limber's equation; Kaiser (1992, 1998) use it to relate P_δ to P_κ

$$P_\kappa(\ell) = \frac{9H_0^4\Omega_\mathrm{m}^2}{4c^4}\int_0^{\chi_\mathrm{h}}\mathrm{d}\chi\frac{g^2(\chi)}{a^2(\chi)}P_\delta\left(\frac{\ell}{f_K(\chi)},\chi\right), \tag{3.44}$$

with $g(\chi)$ explained in (3.42). The above expression is essential for lensing, as it explicitly relates a lensing quantity to cosmological parameters and the power spectrum of density fluctuations (which again depends on the cosmological model). Figure 3.6 shows the power spectrum of convergence P_κ for the same cosmologies as for P_δ in Fig. 2.5. We see that P_κ depends strongly on the underlying cosmological parameters, enabling us to constrain them, once P_κ is determined through observations. In the next section we present several real-space measures which are related to P_κ.

Figure 3.6: *The left plot shows linear (thin lines) and nonlinear (thick lines) power spectra P_κ as a function of the wave vector ℓ. The right side shows the corresponding power spectra in dimensionless form. The cosmologies used in the calculation are similar to those of Fig. 2.5. Furthermore, we assume all sources to be at a constant redshift $z_0 = 1$.*
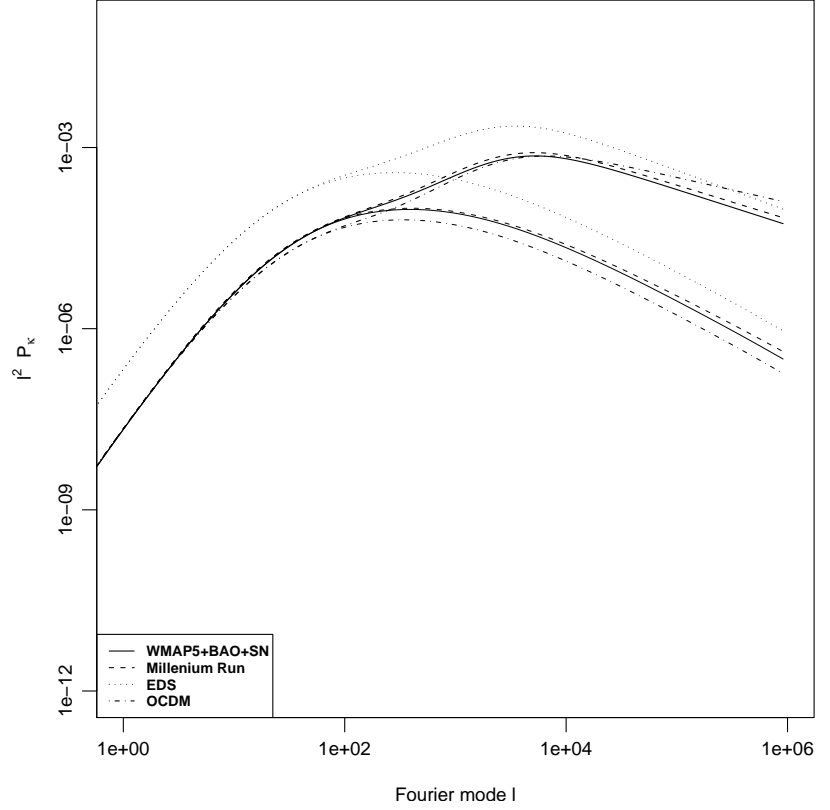
3.3.3   Second-order cosmic shear measures

Considering the Fourier transform of shear and convergence we find the relation

$$\hat{\gamma}(\boldsymbol{\ell}) = e^{2i\phi}\hat{\kappa}(\boldsymbol{\ell})\,, \tag{3.45}$$

which implies

$$\begin{aligned}\langle\hat{\gamma}(\boldsymbol{\ell})\,\hat{\gamma}^*(\boldsymbol{\ell}')\rangle &= \langle\hat{\kappa}(\boldsymbol{\ell})\,\hat{\kappa}^*(\boldsymbol{\ell}')\rangle\\ &= (2\pi)^2\delta_D^{(2)}(\boldsymbol{\ell} - \boldsymbol{\ell}')P_\kappa(\ell)\,, \end{aligned} \tag{3.46}$$

with $\delta_D^{(2)}(\boldsymbol{\ell})$ as the two-dimensional Dirac delta distribution. Hence, the power spectra of the convergence and shear are identical and we can determine P_κ through the measured shear. The shear is a polar quantity, which implies that a rotation of $\phi = 180°$ is an identity transformation. By looking at the shear components in a rotated coordinate frame it is useful to decompose the shear into a *tangential γ_t* and a *cross component $\gamma_\times$*.

$$\gamma_t = -\mathrm{Re}\left(\gamma e^{-2i\phi}\right) = -\gamma_1\cos 2\phi - \gamma_2\sin 2\phi, \tag{3.47}$$

$$\gamma_\times = -\mathrm{Im}\left(\gamma e^{-2i\phi}\right) = \gamma_1\sin 2\phi - \gamma_2\cos 2\phi\,. \tag{3.48}$$

Here, γ_1 and γ_2 refer to cartesian coordinates and ϕ specifies a direction relative to which the shear is measured.

3.3.3.1 Correlation function

In real space the power spectrum corresponds to the two-point correlation function (2PCF). Consider a pair of galaxies at position $\boldsymbol{\vartheta}$ and $\boldsymbol{\vartheta} + \boldsymbol{\theta}$ respectively. Using the above definition of tangential and cross component of the shear we define the 2PCFs as

$$\xi_{\pm}(\theta) = \langle \gamma_t \gamma_t \rangle \pm \langle \gamma_\times \gamma_\times \rangle(\theta) , \tag{3.49}$$

$$\xi_\times(\theta) = \langle \gamma_t \gamma_\times \rangle(\theta) . \tag{3.50}$$

$\xi_\times(\theta)$ vanishes because $\gamma_\times \to -\gamma_\times$ under parity transformation and the LSS is parity-symmetric. Note that the 2PCF only depends on the magnitude of the separation vector but not on the orientation. Practically we can calculate $\xi_\pm$ by measuring the ellipticities of many background galaxy pairs with the same separation. The shear is obtained as described in Sect. 3.2. Furthermore, there is a relation of the 2PCF and the convergence power spectrum P_κ

$$\xi_+(\theta) = \int_0^\infty \frac{\mathrm{d}\ell\,\ell}{2\pi} J_0(\ell\theta) P_\kappa(\ell) , \tag{3.51}$$

$$\xi_-(\theta) = \int_0^\infty \frac{\mathrm{d}\ell\,\ell}{2\pi} J_4(\ell\theta) P_\kappa(\ell) , \tag{3.52}$$

with J_0 as the 0th-order Bessel function and J_4 the 4th-order Bessel function. Recall that there is a similar relation for the matter power spectrum, with only one difference: the matter distribution is three-dimensional, whereas the shear only has two dimensions.

3.3.3.2 Aperture mass dispersion

The so-called *aperture mass* was introduced by Kaiser (1994) and Schneider (1996) as a measure of the local surface mass density around a fixed position $\boldsymbol{\vartheta}$. For an aperture of radius θ the aperture mass is defined as

$$M_{\mathrm{ap}}(\theta, \boldsymbol{\vartheta}) = \int \mathrm{d}^2\vartheta'\, \kappa(\boldsymbol{\vartheta}')\, U_\theta(|\boldsymbol{\vartheta} - \boldsymbol{\vartheta}'|) , \tag{3.53}$$

with U_θ being a filter function satisfying the criterion

$$\int_0^\infty \mathrm{d}\vartheta\, U_\theta(\vartheta)\, \vartheta = 0 . \tag{3.54}$$

Furthermore, Kaiser (1994) and Schneider (1996) show that M_{ap} can be expressed in terms of the tangential shear. For each point inside the aperture, γ_t and $\gamma_\times$ are measured relative to the connecting vector of aperture center to the considered point.

$$M_{\mathrm{ap}}(\theta, \boldsymbol{\vartheta}) = \int \mathrm{d}^2\vartheta'\, Q_\theta(|\boldsymbol{\vartheta} - \boldsymbol{\vartheta}'|)\, \gamma_t(\boldsymbol{\vartheta}') , \tag{3.55}$$

with Q_θ as a weight function related to U_θ by

$$Q_\theta(\vartheta) = \frac{2}{\vartheta^2} \int_0^\vartheta \mathrm{d}\vartheta'\, \vartheta' U_\theta(\vartheta') - U_\theta(\vartheta) . \tag{3.56}$$

The importance of this measure for cosmic shear was discovered later (Schneider et al. 1998) with the dispersion of M_{ap} which is related to P_κ via

$$\langle M_{\mathrm{ap}}^2 \rangle(\theta) = \frac{1}{2\pi} \int_0^\infty \mathrm{d}\ell\, \ell\, P_\kappa(\ell) W_{\mathrm{ap}}(\theta\ell) \,, \tag{3.57}$$

with

$$W_{\mathrm{ap}}(\theta\ell) = \left(\frac{24 J_4(\ell\theta)}{(\ell\theta)^2} \right)^2 . \tag{3.58}$$

3.3.3.3 Shear dispersion

Another second-order measure is the *shear dispersion*. This is obtained by calculating the mean shear inside a circular aperture of radius θ. By averaging over many apertures we find the shear dispersion $\langle |\bar{\gamma}|^2 \rangle$. The relation to the power spectrum is given by

$$\langle |\bar{\gamma}|^2 \rangle(\theta) = \frac{1}{2\pi} \int \mathrm{d}\ell\, \ell\, P_\kappa(\ell) \left(\frac{2 J_1(\ell\theta)}{\ell\theta} \right)^2 . \tag{3.59}$$

When taking a closer look (Fig. 3.8) at the filter functions of the three cosmic shear measures, we just reviewed one sees that ξ_+ and the shear dispersion $\langle |\bar{\gamma}|^2 \rangle$ have very broad filter functions. These imply that for a fixed θ the corresponding measure integrates the over a wide range of the power spectrum. The filter function of ξ_- is more localized, but still broad compared with the filter function of $\langle M_{\mathrm{ap}}^2 \rangle$. The aperture mass dispersion probes the power spectrum very narrowly giving by far the most localized information but it has a low signal. We will discuss advantages and disadvantages of the individual second-order measures in detail in chapter 5. Figure 3.7 shows all second-order cosmic shear measures in comparison.

3.3.3.4 Interrelations

All mentioned two-point statistics depend linearly on the power spectrum. Crittenden et al. (2002) show that they can be related to each other. Using the orthonormality of the Bessel functions we can invert the relation (3.51) and (3.52) and express the power spectrum in terms of the two-point correlation function

$$2\pi \int_0^\infty \mathrm{d}\theta\, \theta\, \xi_+(\theta)\, \mathrm{J}_0(\ell\theta) = P_\kappa(\ell) = 2\pi \int_0^\infty \mathrm{d}\theta\, \theta\, \xi_-(\theta)\, \mathrm{J}_4(\ell\theta) \,. \tag{3.60}$$

Inserting the l.h.s. into (3.52) and the r.h.s. into (3.51), we derive an expression of either correlation function in terms of the other

$$\xi_+(\theta) \;\; = \;\; \xi_-(\theta) + \int_\theta^\infty \frac{\mathrm{d}\vartheta}{\vartheta} \xi_-(\vartheta) \left(4 - 12\frac{\theta^2}{\vartheta^2} \right) \,, \tag{3.61}$$

$$\xi_-(\theta) \;\; = \;\; \xi_+(\theta) + \int_0^\theta \frac{\mathrm{d}\vartheta\, \vartheta}{\theta^2} \xi_+(\vartheta) \left(4 - 12\frac{\vartheta^2}{\theta^2} \right) \,. \tag{3.62}$$
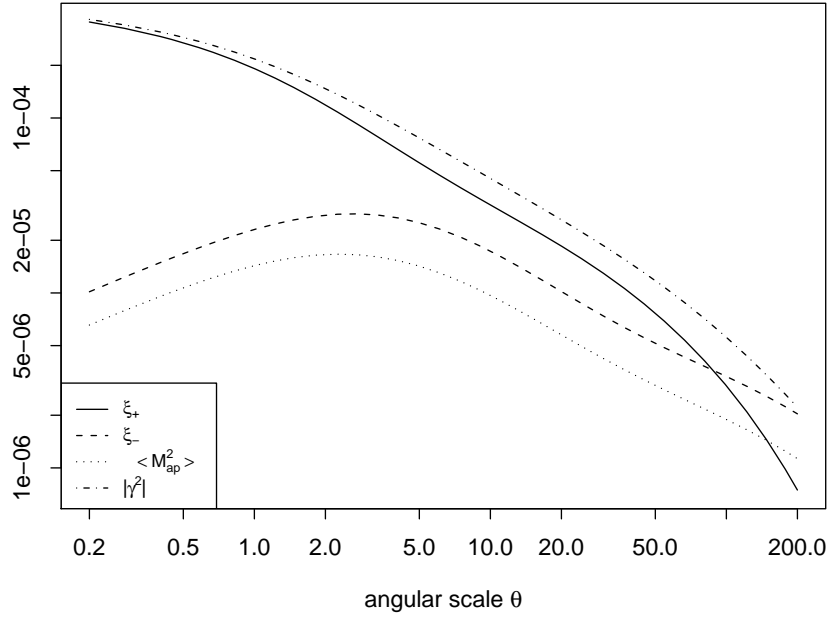
Figure 3.7: *The different 2-point statistics of cosmic shear, calculated via (3.51), (3.52), (3.57), (3.59) for the WMAP5 cosmology (see Tab. 2.1, WMAP+BAO+SN parameters)*

Note that in the first equation the integral does not diverge in case $\vartheta \longrightarrow 0$. The reason for this is that ξ_- goes faster to zero than $1/\vartheta^3$ to infinity, which can be seen from the filter function J_4. Inserting (3.60) into (3.57) we can express $\langle M_{ap}^2 \rangle$ in terms of the correlation function as

$$\langle M_{ap}^2 \rangle(\theta) = \int_0^{2\theta} \frac{d\vartheta\, \vartheta}{\theta^2} \xi_+(\vartheta) T_+\left(\frac{\vartheta}{\theta}\right) = \int_0^{2\theta} \frac{d\vartheta\, \vartheta}{\theta^2} \xi_-(\vartheta) T_-\left(\frac{\vartheta}{\theta}\right) \quad . \tag{3.63}$$

From the above formula one sees that in order to calculate $\langle M_{ap}^2 \rangle$ on an aperture of radius θ, the 2PCF must be measured over a range $[0; 2\theta]$. Both correlation functions can be used for this calculation. The filter functions $T_\pm$ are derived in Schneider et al. (2002a).

$$T_+(x) = \left\{ \frac{6(2-15x^2)}{5} \left[1 - \frac{2}{\pi} \arcsin\left(\frac{x}{2}\right)\right] \right.$$
$$\left. + \frac{x\sqrt{4-x^2}}{100\pi}(120 + 2320x^2 - 754x^4 + 132x^6 - 9x^8) \right\} H(2-x) \quad , \tag{3.64}$$

$$T_-(x) = \frac{192}{35\pi} x^3 \left(1 - \frac{x^2}{4}\right)^{7/2} H(2-x) \quad . \tag{3.65}$$

Similarly, we can express the shear dispersion in terms of the correlation function but with different filter functions

$$\langle |\bar{\gamma}|^2 \rangle(\theta) = \int_0^{2\theta} \frac{d\vartheta\, \vartheta}{\theta^2} \xi_+(\vartheta) S_+\left(\frac{\vartheta}{\theta}\right) = \int_0^{\infty} \frac{d\vartheta\, \vartheta}{\theta^2} \xi_-(\vartheta) S_-\left(\frac{\vartheta}{\theta}\right) , \tag{3.66}$$

where $S_{\pm}$ is given as

$$
\begin{aligned}
S_{+}(x) &= \frac{1}{\pi}\left[4\arccos\left(\frac{x}{2}\right) - x\sqrt{4 - x^2}\right]H(2 - x) \quad, \\
S_{-}(x) &= \frac{x\sqrt{4 - x^2}(6 - x^2) - 8(3 - x^2)\arcsin(x/2)}{\pi x^4}H(2 - x) \\
&\quad + \frac{4(x^2 - 3)}{x^4}H(x - 2) \quad.
\end{aligned}
$$

In either case of the filter function $T_{\pm}$ and $S_{\pm}$, H denotes the Heaviside step function. The filter functions $T_{\pm}$ and $S_{\pm}$ are plotted in Fig. 3.8. It is quite important to mention that all second-order measures can be evaluated from the correlation functions as $\xi_{\pm}$ is easiest to measure. It is independent of gaps on real data fields, coming from satellite traces, bright stars or CCD-defects. All this affects the aperture mass dispersion as well as the shear dispersion if they are directly obtained from the data and not calculated from $\xi_{\pm}$. Furthermore, it is worth to mention that $\langle M_{\mathrm{ap}}^2\rangle$ can be calculated from both, ξ_{+} and ξ_{-}, on a finite interval, which means to calculate $\langle M_{\mathrm{ap}}^2\rangle(\theta)$, the 2PCF must be measured on an angular scale $[0, 2\theta]$. In case of the shear dispersion, this is only possible for ξ_{+}. The reason for this are again the filter functions, the one of $\langle M_{\mathrm{ap}}^2\rangle$ is very narrow, hence all information needed can be obtained from ξ_{+} and ξ_{-}, which have much wider filter functions. However, this is different in case of $\langle|\bar\gamma|^2\rangle$ which also has a broad filter function and therefore also probes the very power specrum on small ℓ (large θ). Due to its narrower filter function ξ_{-} cannot provide this information; it probes the power specrum more locally, therefore one has to "measure" it to infinity to calculate the shear dispersion. The broad filter function of ξ_{+} can still provide this large-scale information.

3.3.3.5   E-modes and B-modes

The shear can be calculated from the deflection potential ψ via (3.15). Noting that γ is a complex quantity whereas ψ is a scalar field, the two components of the shear cannot be independent of each other. This can be shown by calculating the gradient of κ

$$
\nabla\kappa = \begin{pmatrix} \partial_1\gamma_1 + \partial_2\gamma_2 \\ \partial_2\gamma_1 - \partial_1\gamma_2 \end{pmatrix} \equiv \mathbf{u} \quad. \tag{3.67}
$$

Taking again the derivative, we expect $\nabla\times\mathbf{u}$ to vanish, providing a relation between the two shear components in terms of their second derivatives. However, if $\mathbf{u}$ is measured from a data field, most cosmic shear surveys measure a non-gradient component in their signal, the so-called B-modes (Fig. 3.9). B-modes are considered to be a contamination of the pure lensing signal, their origin is not fully explained. The limited validity of the Born approximation (Jain et al. 2000) or redshift source clustering (Schneider et al. 2002b) can also create B-modes, although these effects are small. Intrinsic alignment of source galaxies is another possible source of B-modes. Predictions coming from numerical simulations differ on the impact of these effects (e.g. Heavens et al. 2000; Crittenden et al. 2001; Jing 2002). However, the observed B-mode amplitude is higher than one would expect from these explanations. Most likely, B-modes indicate remaining
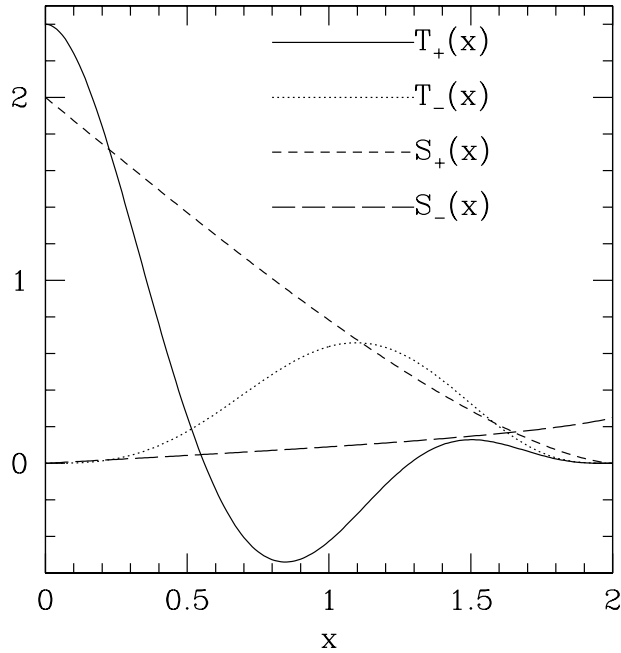
Figure 3.8: *The filter functions T_+, T_-, S_+ and S_-. x gives the ratio of (θ/ϑ). All filter functions except S_- become zero for x > 2. (Figure from Schneider et al. 2006)*

systematics in the observation and data analysis (e.g. insufficient PSF-correction). We further outline possible contaminations to the shear signal in Sect. 3.5.

To distinguish between the curl and the gradient parts of **u** (E-modes), we consider the complex surface mass density $\kappa = \kappa^E + i\kappa^B$ and define

$$\nabla^2 \kappa^E = \nabla \cdot \mathbf{u} \tag{3.68}$$

$$\nabla^2 \kappa^B = \nabla \times \mathbf{u} = \partial_1 u_2 - \partial_2 u_1 . \tag{3.69}$$

Similarly we define two components of the deflection potential via the Poisson equation

$$\nabla^2 \psi^{E,B} = 2\kappa^{E,B} . \tag{3.70}$$

In general, both ψ and κ are expressed as complex quantities.

$$\psi = \psi^E + i\psi^B , \tag{3.71}$$

$$\kappa = \kappa^E + i\kappa^B . \tag{3.72}$$

Considering second-order statistics, the decomposition into E-and B-modes can also be applied to the power spectrum. From (3.46) together with (3.72) Schneider et al. (2002b) calculate

$$\langle \hat{\kappa}_E(\boldsymbol{\ell})\hat{\kappa}_E^*(\boldsymbol{\ell}') \rangle = (2\pi)^2 \delta_D^{(2)}(\boldsymbol{\ell} - \boldsymbol{\ell}') P_E(\ell) , \tag{3.73}$$

$$\langle \hat{\kappa}_B(\boldsymbol{\ell})\hat{\kappa}_B^*(\boldsymbol{\ell}') \rangle = (2\pi)^2 \delta_D^{(2)}(\boldsymbol{\ell} - \boldsymbol{\ell}') P_B(\ell) , \tag{3.74}$$

$$\langle \hat{\kappa}_E(\boldsymbol{\ell})\hat{\kappa}_B^*(\boldsymbol{\ell}') \rangle = (2\pi)^2 \delta_D^{(2)}(\boldsymbol{\ell} - \boldsymbol{\ell}') P_{EB}(\ell) . \tag{3.75}$$
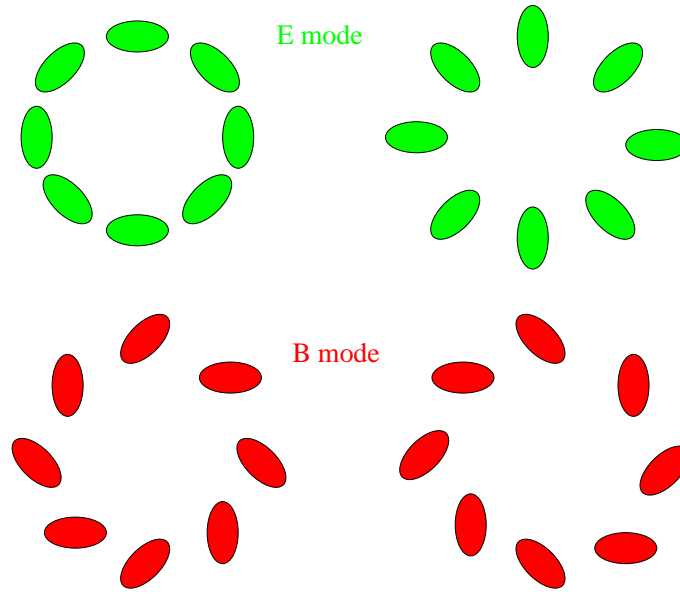
Figure 3.9: *The upper part of this figure shows a typical E-mode pattern which can be expressed as the gradient part of $\nabla\kappa$. The left panel shows the mapping of galaxies due to a matter overdensity and the right images are produced by an underdensity. In the lower part we see the corresponding curl component of the shear signal (B-modes). (Figure from van Waerbeke & Mellier 2003)*

The cross power spectrum P_{EB} is expected to vanish for a statistically parity-invariant shear field. Similar to (3.51) and (3.52), we can now express the correlation function in terms of the decomposed power spectrum

$$\xi_+(\theta) = \int_0^\infty \frac{d\ell\,\ell}{2\pi} J_0(\ell\theta) \left[P_E(\ell) + P_B(\ell))\right] , \tag{3.76}$$

$$\xi_-(\theta) = \int_0^\infty \frac{d\ell\,\ell}{2\pi} J_4(\ell\theta) \left[P_E(\ell) - P_B(\ell)\right] . \tag{3.77}$$

Again, similar to the case of E-modes only, the above relations can be inverted

$$P_{E,B}(\ell) = \pi \int_0^\infty d\theta\,\theta \left[\xi_+(\theta) J_0(\ell\theta) \pm \xi_-(\theta) J_4(\ell\theta)\right] . \tag{3.78}$$

In order to constrain parameters we want to distinguish between the power spectrum resulting from E-modes and the one resulting from B-modes. Practically this cannot be done using (3.78), because this requires a measured correlation function with arguments from 0 to infinity. Instead, one uses the aperture mass dispersion to separate E-modes and B-modes. More precisely, one can show that $\langle M_{\rm ap}^2 \rangle$ is sensitive only to E-modes and $\langle M_\perp^2 \rangle$ provides a measure of the B-mode

only and vanishes in its absence

$$\langle M_{\mathrm{ap}}^2\rangle(\theta) \;\; = \;\; \frac{1}{2\pi}\int_0^\infty \mathrm{d}\ell\,\ell\,P_{\mathrm{E}}(\ell)W_{\mathrm{ap}}(\theta\ell)\,, \tag{3.79}$$

$$\langle M_\perp^2\rangle(\theta) \;\; = \;\; \frac{1}{2\pi}\int_0^\infty \mathrm{d}\ell\,\ell\,P_{\mathrm{B}}(\ell)W_{\mathrm{ap}}(\theta\ell)\,. \tag{3.80}$$

Again, we can calculate $\langle M_{\mathrm{ap}}^2\rangle$ and $\langle M_\perp^2\rangle$ directly from the much easier obtainable correlation functions

$$\langle M_{\mathrm{ap}}^2\rangle(\theta) \;\; = \;\; \frac{1}{2}\int_0^{2\theta}\frac{\mathrm{d}\vartheta\,\vartheta}{\theta^2}\left[\xi_+(\vartheta)T_+\left(\frac{\vartheta}{\theta}\right)+\xi_-(\vartheta)T_-\left(\frac{\vartheta}{\theta}\right)\right]\,, \tag{3.81}$$

$$\langle M_\perp^2\rangle(\theta) \;\; = \;\; \frac{1}{2}\int_0^{2\theta}\frac{\mathrm{d}\vartheta\,\vartheta}{\theta^2}\left[\xi_+(\vartheta)T_+\left(\frac{\vartheta}{\theta}\right)-\xi_-(\vartheta)T_-\left(\frac{\vartheta}{\theta}\right)\right]\,. \tag{3.82}$$

In the absence of B-modes (3.82) vanishes, hence ξ_+ can be expressed through ξ_-, implying that for the case of a pure lensing signal (only E-modes) we can express $\langle M_{\mathrm{ap}}^2\rangle$ in (3.81) only either via ξ_+ or ξ_-.

3.4 Cosmological parameters from cosmic shear

In recent years cosmic shear has become a valuable tool to constrain cosmological parameters. In this section we briefly outline the most important results and also explain the main difficulties which cosmic shear parameter estimation faces today. We conclude with a short outlook to future surveys.

Cosmic shear is an extremely powerful method to determine the normalization of the power spectrum σ_8 in case Ω_{m} can be obtained from a different cosmological probe. Table 3.1 summarizes the most important surveys and their results for σ_8 since the detection of cosmic shear in the year 2000 (Bacon et al. 2000; Kaiser et al. 2000; van Waerbeke et al. 2000; Wittman et al. 2000). Although the obtained values for σ_8 deviate slightly, they are in good agreement with each other and results coming from other cosmological experiments (see Sect 2.3.6). The latest cosmic shear analysis of Fu et al. (2008) is suited best to illustrate the high quality of parameter estimation, which can be done with cosmic shear. In this analysis the raw data of the CFHTLS 3-years data release was reduced using two independent pipelines. The estimated shear signals were in good agreement. The 2PCF was detected out to a scale of 460′ which by far exceeds the scales on which former analyses were able to measure a signal. From the 2PCF several 2-point cosmic shear statistics were calculated, such as the aperture mass dispersion, ξ_{E}, and the top-hat shear dispersion. Figure 3.10 illustrates the results for the three aforementioned measures; we see that the E-mode signal is significantly larger compared to possible contaminations indicated by the amplitude of the B-mode signal. Furthermore, the latter is consistent with zero on all scales except around 60 arcmin. In chapter 6 we calculate the ring statistics from the same set of 2PCF as used in Fu et al. (2008) to check their findings. The remarkable results cosmic shear has achieved so far will improve in the future with large upcoming surveys like Pan-STARRS,
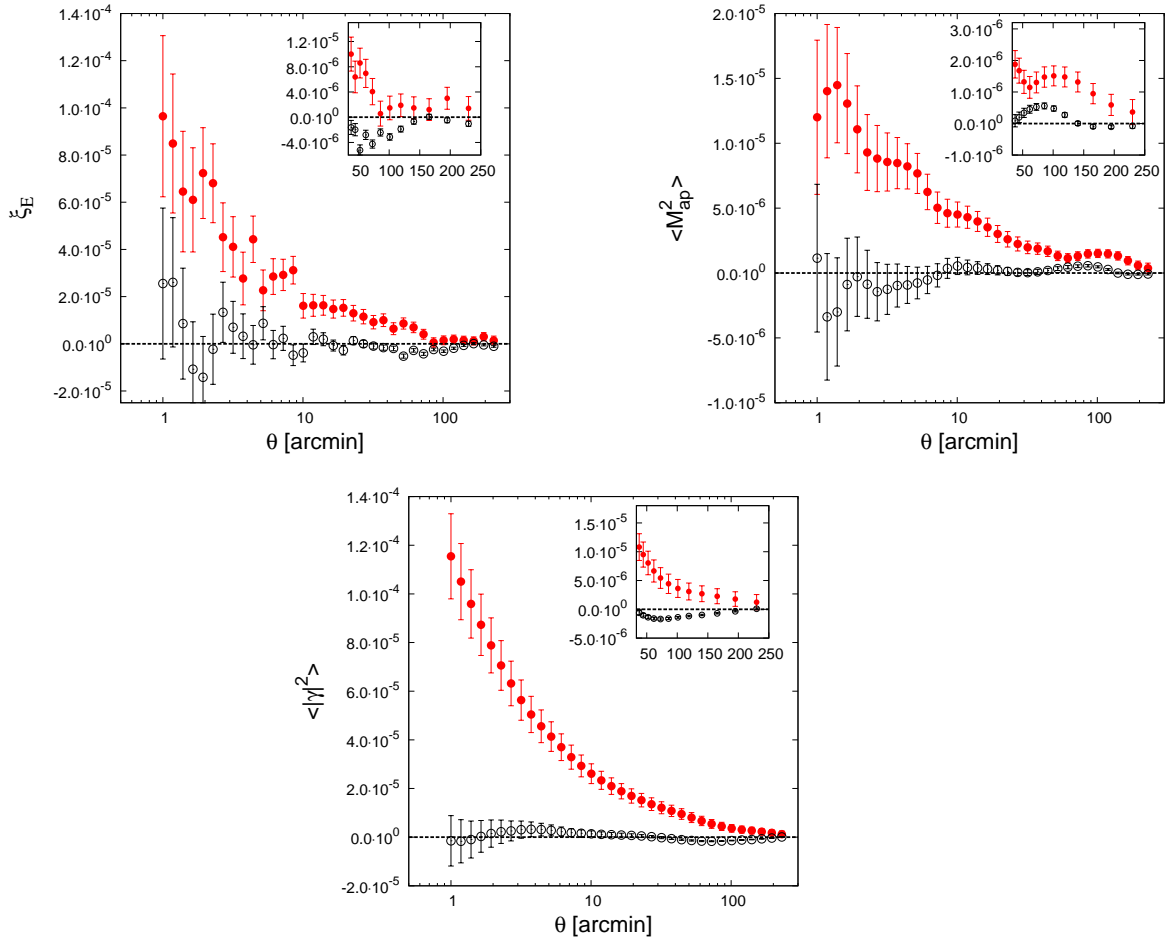
Figure 3.10: *This figure shows three second-order measures (ξ_E, $\langle M_{ap}^2 \rangle$ and $\langle |\gamma|^2 \rangle$) of the latest CFHTLS analysis. The red dots correspond to the E-mode signal, the black circles to the B-mode signal. Note that the latter is almost everywhere consistent with zero. (Figure from Fu et al. 2008)*

KIDS, DES, Euclid or LSST. Already the first two surveys, starting in 2009, provide data which enables us to estimate the shear signal with less than 1% statistical error. These small statistical errors make cosmic shear an ideal tool for future dark energy studies (see Peacock et al. 2006; Albrecht et al. 2006), we examine this more closely in chapter 8. Still, for precision cosmology cosmic shear has to solve several problems, which we outline in the next section.

3.5   Future challenges - precision cosmology with cosmic shear

Despite past successes and positive future prospects, there are unsolved systematics and uncertainties which affect parameter estimation with cosmic shear. These issues, though of minor importance for the current generation of weak lensing surveys, must be solved for the next survey

Table 3.1: *Summary of weak lensing surveys with constraints on σ_8 – assuming a flat universe with $n_s = 1$. This table is an extension of a similar table in Peacock et al. (2006)*

Survey	Sky coverage	n_{gal}/arcmin^2	σ_8 ($\Omega_m = 0.3$)	Ref.
VLT-Descart	0.65 deg^2	21	1.05±0.05	Maoli et al. (2001)
Groth Strip	0.05 deg^2	23	$0.90^{+0.25}_{-0.30}$	Rhodes et al. (2001)
MDS	0.36 deg^2	23	0.94±0.17	Refregier et al. (2002)
RCS	16.4 + 7.6 deg^2	9	$0.81^{+0.14}_{-0.19}$	Hoekstra et al. (2002b)
Virmos-Descart	8.5 deg^2	15	0.98 ±0.06	Van Waerbeke et al. (2002)
RCS	45.4 + 7.6 deg^2	9	$0.87^{+0.09}_{-0.12}$	Hoekstra et al. (2002a)
COMBO-17	1.25 deg^2	32	0.72 ±0.09	Brown et al. (2003)
Keck + WHT	0.6 + 1.0deg^2	27.5/15	0.93 ±0.13	Bacon et al. (2003)
CTIO	75 deg^2	7.5	$0.71^{+0.06}_{-0.08}$	Jarvis et al. (2003)
SUBARU	2.1 deg^2	32	$0.78^{+0.55}_{-0.25}$	Hamana et al. (2003)
COMBO-17	1.25 deg^2	R	0.67 ±0.10	Heymans et al. (2004)
FIRST	10000 deg^2	0.01	1.0 ±0.2	Chang et al. (2004)
GEMS	0.22 deg^2	60	0.68 ±0.13	Heymans et al. (2005)
WHT + COMBO-17	4.0 + 1.25deg^2	15 / 32	1.02 ±0.15	Massey et al. (2005)
Virmos-Descart	8.5 deg^2	12.5	0.83 ±0.07	van Waerbeke et al. (2005)
CTIO	75 deg^2	7.5	$0.71^{+0.06}_{-0.08}$	Jarvis et al. (2006)
CFHTLS Deep	2.1 deg^2 +	22	0.89 ±0.06	Semboloni et al. (2006)
CFHTLS Wide	22 deg^2	13	0.86 ±0.05	Hoekstra et al. (2006)
GaBoDS	15 deg^2	12.5	0.80 ±0.10	Hetterscheidt et al. (2007)
ACS parallel + GEMS + GOODS	0.018 + 0.027 deg^2	63/96	$0.52^{+0.13}_{-0.17}$	Schrabback et al. (2007)
COSMOS	1.64 deg^2	40	$0.866^{+0.085}_{-0.068}$	Massey et al. (2007b)
CFHTLS Wide (3years)	34.2 deg^2	13.3	0.785±0.043 (for $\Omega_m = 0.25$)	Fu et al. (2008)
CFHTLS Wide (3years)	34.2 deg^2	13.3	0.698±0.038 (for $\Omega_m = 0.3$)	rescaled Fu et al. (2008)

generation in order to extract the full information from their high-quality data. In the following we summarize the main issues, with a focus on brevity in the explanation; each of the points can easily be subject of a complete thesis.

PSF correction and shape measurements   An insufficient PSF correction causes a bias in the amplitude of the lensing signal. The strength of this contamination changes with size, surface brightness, and intrinsic ellipticity of the considered galaxy and is therefore redshift dependent. The Shear TEsting Program (STEP) has significantly improved on this issue (for latest results see Heymans et al. 2006; Massey et al. 2007a); still the accuracy of of the ellipticity measurements does not satisfy the requirements for precision cosmology. The STEP program is ongoing; the problem is even discussed outside the lensing community (Bridle et al. 2008).

Astrophysical contaminations   In Sect. 3.2 we show that each measured ellipticity is a very noisy measure of the local shear, $\epsilon_i = \epsilon_i^{(s)} + \gamma(\boldsymbol{\theta}_i)$. When calculating the correlator of two ellipticities the following terms arise (assume that $z_1 \leq z_2$)

$$\langle \epsilon_1 \epsilon_2^* \rangle = \underbrace{\left\langle \epsilon_1^{(s)} \epsilon_2^{(s)*} \right\rangle}_{\text{intrinsic alignment}} + \underbrace{\left\langle \epsilon_1^{(s)} \gamma_2^* \right\rangle}_{\text{shape−shear correlation}} + \underbrace{\left\langle \gamma_1 \epsilon_2^{(s)*} \right\rangle}_{\equiv 0} + \underbrace{\left\langle \gamma_1 \gamma_2^* \right\rangle}_{\text{cosmic shear signal}} \tag{3.83}$$

Intrinsic alignment   If galaxies are physically close their ellipticities are not necessarily oriented randomly but can align according to the filamentary structure which is present at their location. This effect mimics a shear and thereby contaminates the signal. This contamination mainly affects shallow surveys (e.g. SDSS) and becomes less important the higher the mean redshift of the source galaxies. However, King & Schneider (2003) show how to separate the cosmic shear signal from intrinsic alignments contaminations if redshift information is available.

Shape-shear correlation   Besides intrinsic alignment, Hirata & Seljak (2004) outline that a the cosmic shear signal can be affected by an additional contamination, namely that the shear of a background galaxy and the ellipticity of a foreground galaxy can have a non-zero correlation. More precisely, in case the foreground galaxy is hosted in a large-scale structure filament, which is also responsible for the shear of the background galaxy, the corresponding term in (3.83) can give a non-zero contribution. The effect was detected by Mandelbaum et al. (2006). Recently Joachimi & Schneider (2008) developed a method to remove the shape-shear contribution, by introducing a redshift-dependent filter function which excludes the contaminated part of the measured 2PCF. Similar to the intrinsic alignment removal, this method requires accurate redshift information.

Source/ lens clustering   As outlined in Schneider et al. (2002b) clustering of source galaxies induces a B-mode signal as well as an additional contribution to the E-mode. However, the contamination is only present on scales smaller than one arcminute, which can be discarded in order to exclude the effect. In addition, the clustering of sources/ lenses affects the amplitude of higher-order cosmic shear measures (Bernardeau 1998). Remote sources are lensed by high

density regions, creating a high shear signal. These high-density regions serve as lenses, but in addition they host many galaxies which are not sheared by this high-density region. This preferred accumulation of galaxies results in a different shear signal, than if the galaxies were distributed uniformely. Note that once again this effect can be removed employing redshift information, more precisely the effect is negligible in case the redshift distribution is sufficiently narrow.

Redshift information   As mentioned before precise redshift information is necessary for precision cosmology in order to account properly for astrophysical contaminations. In addition, one can enormously improve the constraints on cosmological parameters if the shear signal is examined as a function of the sources' redshift (shear tomography) instead of considering 2D-lensing only. For the large sample of galaxies which are observed in future survey, spectroscopic redshifts are hardly obtainable. Instead, one has to refer to photometric redshifts (e.g. Heymans et al. 2005), which improve in accuracy the more filter bands are available.

Theoretical tasks   Theoretical issues of future cosmic shear surveys can be structured into three main groups.

- precise lensing predictions (Born approximation, thin lens approximation, inaccuracy of Limber's equation or the flat-sky approximation),

- precise cosmological predictions (accurate models for the non-linear power spectrum or non-Gaussian covariances, models for higher order moments of the density/shear field)

- accurate methods for inference of cosmological parameters (robust estimators of cosmic shear measures, precise model for the likelihood function, derivation of the covariance)

In this thesis we focus on the latter two theoretical tasks. In particular, we want to improve on the issues of cosmic shear estimators and covariances. In chapter 5 we analyze the quality of second-order cosmic shear estimators with respect to their information content and robustness against B-mode contamination. We develop a new data vector which combines the advantages of the most commonly used second-order cosmic shear measures and test this new data vector using ray-tracing simulations. As a second contribution to the topic of estimators, we examine the so-called ring statistics (chapter 6), which is the most recently developed second-order cosmic shear measure (Schneider & Kilbinger 2007). Finally, we employ the ring statistics to measure the shear signal from the CFHTLS 3 years data, which marks the first application of the ring statistics to real data.
As a second topic we examine the impact of covariances on the parameter constraints from cosmic shear. In particular, we show that cosmic shear covariances are cosmology dependent, quantify the impact on parameter constraints when neglecting this effect and provide a method to incorporate cosmology dependent covariances into a likelihood analysis (chapter 7). In the last chapter we analyze the impact on parameter constraints when using non-Gaussian covariances. Here, we focus on the estimation of dark energy parameters as this will play an important role in future cosmic shear surveys.

Chapter 4

Parameter estimation

At the end of the last chapter we summarized the recent achievements of cosmic shear in constraining cosmology and its future perspectives. In this chapter we introduce the main concepts of parameter estimation, which we extensively use in the next chapters. It is a very common situation in science to measure data from an experiment and, in the framework of an underlying theory, to infer information on the parameters which determine this theory. For the case of cosmic shear this means: we measure the galaxy ellipticities, therefrom calculate the 2PCF (or other cosmic shear measures), and finally constrain cosmological parameters. The underlying theory was outlined in the last two chapters; it is given by the cosmological model (if not stated otherwise, we assume the ΛCDM model) and the concepts of gravitational lensing.

We illustrate the concepts of parameter estimation using the basic two-point cosmic shear statistics, i.e. the 2PCF, as a familiar example. Therefore we first explain how estimators and covariances of the 2PCF are obtained, either from the measured galaxy ellipticities or from simulations and/or analytic expressions. The estimators and similarly the covariances are binned (linearly or logarithmically); we conclude this chapter with an analysis of binning effects and their impact on parameter inference.

4.1 Estimators and covariances for the 2PCF

In this section we define the estimators and covariances of the 2PCF. Similar quantities for other second-order statistics are introduced when needed.

Consider a sample of galaxies with angular positions θ_i. For each pair of galaxies we define the connecting vector $\theta = \theta_i - \theta_j$ and determine tangential and cross-components of the ellipticities (ϵ_t and $\epsilon_\times$) with respect to this connecting vector. From these ellipticities we estimate the 2PCF in linear or logarithmic bins in ϑ with bin width $\Delta\vartheta$ (Schneider et al. 2002a). If the bin width is small enough an unbiased estimator for $\xi_\pm(\vartheta)$ is given by

$$\hat{\xi}_\pm(\vartheta) = \frac{1}{N_p(\vartheta)} \sum_{ij} (\epsilon_{it}\epsilon_{jt} \pm \epsilon_{i\times}\epsilon_{j\times}) \Delta_\vartheta(|\theta_i - \theta_j|) \,, \tag{4.1}$$

with $N_{\mathrm{p}}(\vartheta) = \sum_{ij} \Delta_\vartheta(|\theta_i - \theta_j|)$ as the number of galaxy pairs inside a bin, and $\Delta_\vartheta(|\theta_i - \theta_j|)$ is 1 if $|\theta_i - \theta_j|$ lies inside bin ϑ, 0 otherwise.

Obtaining appropriate covariances is an important issue in the context of a precision cosmology likelihood analysis. For the 2PCF its general definition reads

$$C_\xi\left(\vartheta_i, \vartheta_j\right) := \left\langle \left(\xi_\pm(\vartheta_i) - \hat{\xi}_\pm(\vartheta_i)\right)\left(\xi_\pm(\vartheta_j) - \hat{\xi}_\pm(\vartheta_j)\right)\right\rangle. \tag{4.2}$$

As one already sees from (4.2) the 2PCF has four different covariances, denoted as C_{++}, C_{+-}, C_{-+}, C_{--}. Only three of them are independent since $C_{+-}(\vartheta_i, \vartheta_j) = C_{-+}(\vartheta_j, \vartheta_i)$. We neglect the index ξ in (4.2) from now on, as we only consider 2PCF covariances in this chapter.

Several methods to derive a covariance are suggested in the literature and have been applied to cosmic shear data. An analytic expression assuming a Gaussian shear field is derived in Schneider et al. (2002a) and confirmed in Joachimi et al. (2008) who use a power spectrum approach which significantly reduces the computational effort in the calculation. This analytic expression has been used for parameter estimation in many surveys (e.g. van Waerbeke et al. 2005; Semboloni et al. 2006; Hoekstra et al. 2006) and it is also used extensively in this thesis. Schneider et al. (2002a) and Joachimi et al. (2008) decompose the covariance into three terms, namely the cosmic variance term (V), the pure shot noise term (S), and the mixed term (M)

$$C_{++}(\vartheta_i, \vartheta_j) = V_{++} + M_{++} + S, \tag{4.3}$$

$$C_{--}(\vartheta_i, \vartheta_j) = V_{--} + M_{--} + S, \tag{4.4}$$

$$C_{+-}(\vartheta_i, \vartheta_j) = V_{+-} + M_{+-}. \tag{4.5}$$

The pure shot noise term vanishes in case of C_{+-} and only contributes to the diagonal of C_{++} and C_{--}. It can be calculated as

$$S = \frac{\sigma_\epsilon^4}{2\pi\vartheta_i\Delta\vartheta_i A\bar{n}^2} \delta_{\vartheta_i\vartheta_j}, \tag{4.6}$$

where A denotes the area the data field, σ_ϵ is the intrinsic ellipticity dispersion, and $\bar{n}$ the number density of source galaxies. The cosmic variance term (V) and the mixed term (M) can be either calculated via the power spectrum or via the 2PCF. According to Joachimi et al. (2008) the power spectrum expression reads

$$V_{\pm\pm} = \frac{1}{\pi A} \int_0^\infty \mathrm{d}\ell\, \ell \mathrm{J}_{0/4}(\ell\vartheta_i)\mathrm{J}_{0/4}(\ell\vartheta_j) P_{\mathrm{E}}^2(\ell), \tag{4.7}$$

$$M_{\pm\pm} = \frac{\sigma_\epsilon^2}{\pi A\bar{n}} \int_0^\infty \mathrm{d}\ell\, \ell \mathrm{J}_{0/4}(\ell\vartheta_i)\mathrm{J}_{0/4}(\ell\vartheta_j) P_{\mathrm{E}}(\ell). \tag{4.8}$$

However, on small scales the assumption of a Gaussian shear field breaks down; according to Kilbinger & Schneider (2005) and Semboloni et al. (2007) non-Gaussian effects already become important at angular scales $\lesssim 10$ arcmin. To account for this non-Gaussianity, Semboloni et al. (2007) introduce a calibration factor which is derived from a comparison of Gaussian to ray-tracing covariances. An application of this method to real data can be found in Fu et al. (2008). An alternative approach is to derive the covariance matrix from the data (e.g. Hetterscheidt et al. 2007; Massey et al. 2007b). There, the covariance is calculated via field-to-field variation which

involves a separation of the data set into many independent subsamples. This may lead to a loss of information on large scales if the survey is not sufficiently large. A third approach is to estimate the covariance matrix from ray-tracing simulations, a method which circumvents the aforementioned loss in information. Although in this method the covariance is again derived via field-to-field variation, we can choose a sufficiently large numerical simulation to create many independent subsamples of adequate size.

Note that the last two methods involve an estimation process in the determination of the covariance matrix, which means that the inverse is biased and one has to correct for this effect (Anderson 2003; Hartlap et al. 2007). We return to this issue in the next chapter. Nevertheless, deriving covariance matrices from ray-tracing simulations seems to be a promising method as it preserves all the information in the data and additionally takes the non-Gaussianity of the shear field into account. However, the analytic expression and the ray-tracing covariance assume a specific cosmological model in their derivation. So far, cosmic shear likelihood analyses treat the covariance matrix as constant with respect to cosmology, hence its underlying cosmological model is assumed not to influence the parameter constraints. It is the intention of chapter 7 to check for this assumption and in case it does not hold, to present an improved likelihood formalism for future surveys.

4.2 Basic theory of statistical inference

In the last section we described how to derive estimators and covariances for the 2PCF, which are the basic ingredients needed in a cosmic shear likelihood analysis. We now turn to the question of how to infer cosmological parameters from these ingredients. There are two different schools for parameter inference, the *frequentist* and the *Bayesian*. In many cases the differences hardly affect the mathematical treatment, but rather the interpretation of the results. Often both theories can be used complementary. In the following we briefly outline the main differences and the implications for cosmic shear.

Frequentists define probability as *the frequency of occurrence of an observed random variable (the data) during repeated experiments*. Prior information is not considered in their ansatz. In contrast, Bayesians start from prior knowledge on the underlying parameters, assume a probability distribution for the observed data and calculate a posterior probability via Bayes theorem (see 4.10). This posterior probability expresses the *degree of belief that a specific parameter set is true, given the prior information and the observed data*. Obviously, the Bayesian probability definition ("degree of belief") is somewhat subjective, equally unpleasant is the fact that the analytic form of the likelihood must be assumed. In case the likelihood of the data is not known, this assumption depends on the data analyst. In order to derive the analytic form of a likelihood objectively, Bayesians introduced the *maximum entropy principle* (for details see Loredo 1990, and references therein). There is also a connection to frequentist theory, namely for the case that an experiment can be repeated many times, we can infer the Bayesian likelihood from the frequentist probability definition. In particular this applies to the cosmic shear case, where we measure the ellipticities of many galaxies, which, in the absence of intrinsic alignment or shape-shear correlation, can be treated as independent experiments. Frequentists try to avoid the aforementioned

deficits, but face other difficulties. These can be illustrated best, when looking at the concepts of *confidence intervals* (frequentist) compared to *credible regions* (Bayesian). The frequentist probability definition only allows for statements about the long term behavior of an repeated experiment. Consider an experiment with true underlying mean μ, which is repeated n times, obtaining the means $\bar{x}_i$. Then, the frequentist e.g. 68% confidence interval denotes the region around μ which in 68% of the repetitions overlaps with the possible outcomes of the experiment $\bar{x}_i$. In case the experiment can be performed only once, which is true for many astronomical observations, the above interpretation of confidence intervals is meaningless. In contrast, the corresponding Bayesian probability definition is based on only one data set obtained from one experiment. The corresponding 68% credible region defines a range in parameter space which contains the underlying mean with a probability of 68%; this second interpretation is also what one intuitively is interested in.

However, in case the data follows a Gaussian likelihood with $\bar{x}$ being an unbiased estimator of μ, one can show (e.g. Loredo 1990) that both methods are mathematically equivalent, hence the intervals constrain the same parameter space. As we explain below, there is good reason to approximate the 2PCF likelihood to be a multivariate normal distribution, therefore one can, in principle, base a cosmic shear data analysis on frequentist theory and use the χ^2-statistics to constrain parameters. Nevertheless, this makes the inclusion of prior information more difficult (strictly speaking: prior information is not considered in frequentist theory); in addition the χ^2-statistics cannot be used in the extended likelihood analysis, which we outline in chapter 7. We therefore follow Bayesian concepts and interpretations, throughout this thesis.

Consider a cosmic shear data vector consisting of the 2PCF, both ξ_+ and ξ_-

$$\xi = \begin{pmatrix} \xi_+ \\ \xi_- \end{pmatrix} \quad \text{with} \quad \xi_+ = \begin{pmatrix} \xi_+(\vartheta_1) \\ \vdots \\ \xi_+(\vartheta_{d/2}) \end{pmatrix}, \; \xi_- = \begin{pmatrix} \xi_-(\vartheta_1) \\ \vdots \\ \xi_-(\vartheta_{d/2}) \end{pmatrix}, \tag{4.9}$$

hence the data vector has dimension d. From these data we want to infer information on the cosmological parameters, subsumed in the parameter vector π. According to Bayes theorem the posterior likelihood $p(\pi|\xi)$ is defined as

$$p(\pi|\xi) = \frac{p(\xi|\pi)}{p(\xi)} \, p(\pi), \tag{4.10}$$

where $p(\pi)$ denotes the prior probability density, $p(\xi|\pi)$ is the likelihood and $p(\xi)$ the so-called evidence. The prior usually contains knowledge on the parameter vector π coming from former experiments. If not stated otherwise, we assume flat priors with cutoffs, which means $p(\pi)$ is constant for all parameters inside a fixed interval and $p(\pi) = 0$ else. The exact functional form of the likelihood function is subject to ongoing research. It is state of the art in a cosmic shear likelihood analysis to assume a multivariate normal likelihood function. This assumption is partly justified if the 2PCF is measured from many independent patches of the sky and if the estimator of a bin is calculated from a sufficient number of galaxies. Then, the *central limit theorem* suggests that $\hat{\xi}(\vartheta_i)$ is normally distributed (Kendall & Stuart 1979). If in addition the individual $\hat{\xi}(\vartheta_i)$ were independent, it follows that $\hat{\xi}$ is distributed as a multivariate normal distribution. Note, that

the latter assumption does not hold; the fact that we observe off-diagonal terms in the covariance matrix implies that the individual 2PCFs estimators are correlated (and therefore not independent of each other). The *Independent Component Analysis* (ICA) is a method to check whether the assumption of a multivariate normal distribution for the 2PCF data vector holds, and if not it will improve on this issue. Our (multivariate normal) likelihood $p(\boldsymbol{\xi}|\boldsymbol{\pi})$ reads

$$p(\boldsymbol{\xi}|\boldsymbol{\pi}) = \frac{\exp\left[-\frac{1}{2}\left((\boldsymbol{\xi}_\pi - \hat{\boldsymbol{\xi}})^\mathrm{t}\, \mathbf{C}^{-1}\, (\boldsymbol{\xi}_\pi - \hat{\boldsymbol{\xi}})\right)\right]}{(2\pi)^{d/2}\, |\mathbf{C}|^{\frac{1}{2}}}, \tag{4.11}$$

where $\hat{\boldsymbol{\xi}}$ denotes the mean data vector and $\boldsymbol{\xi}_\pi$ the model data vector. As these vectors have dimension d, $|\mathbf{C}|$ is the determinant of a $d \times d$ 2PCF covariance matrix (see Sect. 4.1) with the structure

$$\mathbf{C} = \left(\begin{array}{c|c} \mathbf{C}_{++} & \mathbf{C}_{+-} \\ \hline \mathbf{C}_{+-}^\mathrm{t} & \mathbf{C}_{--} \end{array} \right). \tag{4.12}$$

If the $|\mathbf{C}|$ in (4.11) is independent of the parameters (which we examine in chapter 7) all cosmological information is contained in the χ^2-function, which is defined as

$$\chi^2(\boldsymbol{\xi}, \boldsymbol{\pi}) = (\boldsymbol{\xi}_\pi - \hat{\boldsymbol{\xi}})^\mathrm{t}\, \mathbf{C}^{-1}\, (\boldsymbol{\xi}_\pi - \hat{\boldsymbol{\xi}}). \tag{4.13}$$

The evidence is a normalization obtained by integrating the likelihood over the considered parameter space

$$p(\boldsymbol{\xi}) = \int \mathrm{d}\boldsymbol{\pi}\, \frac{\exp\left[-\frac{1}{2}\left((\boldsymbol{\xi}_\pi - \hat{\boldsymbol{\xi}})^\mathrm{t}\, \mathbf{C}^{-1}\, (\boldsymbol{\xi}_\pi - \hat{\boldsymbol{\xi}})\right)\right]}{(2\pi)^{d/2}\, |\mathbf{C}|^{\frac{1}{2}}}. \tag{4.14}$$

We express the result of a likelihood analysis through two-dimensional contour plots. We choose likelihood contours representing the aforementioned credible regions which contain the true parameter set with 68%, 95%, 99,9% probability. In addition, we quantify the size of these credible regions through the determinant of the second-order moment of the posterior likelihood (see Kilbinger & Schneider 2004)

$$Q_{ij} \equiv \int \mathrm{d}^2\boldsymbol{\pi}\, p(\boldsymbol{\pi}|\boldsymbol{\xi})\, (\pi_i - \pi_i^\mathrm{f})(\pi_j - \pi_j^\mathrm{f}), \tag{4.15}$$

with π_i and π_j as the varied parameters, π_i^f as the parameter of the fiducial model. The determinant is given by

$$q = \sqrt{|Q_{ij}|} = \sqrt{Q_{11}Q_{22} - Q_{12}^2}. \tag{4.16}$$

Smaller credible regions in parameter space correspond to a smaller value of q. Note, that the q can only be applied to two-dimensional parameter spaces; when considering more free parameters one first has to marginalize over all but two parameters in order to apply 4.16.

4.3 The Fisher information matrix

There are several methods to sample the posterior likelihood in a parameter space. In this thesis we calculate the likelihood on a fixed grid of parameters. Note, that there are more sophisticated methods, e.g. Monte Carlo methods, importance sampling, or Gibbs sampling (see Gelman et al. 2004, for an introduction and references). The Fisher matrix is an approximation to these methods which is valuable because of its low computational costs.

The Cramér-Rao inequality states that the Fisher matrix gives lower bounds on the error bars in parameter space; we use it in chapter 7 in order to quantify the accuracy with which we infer parameters from a data set. The definition reads (Kendall & Stuart 1979; Tegmark et al. 1997)

$$\mathbf{F}_{ij} = \left\langle \frac{\partial^2 L}{\partial \pi_i \, \partial \pi_j} \right\rangle = \left(\frac{\partial^2 L}{\partial \pi_i \, \partial \pi_j} \right)_{\pi=\pi_{\mathrm{ML}}} \, , \tag{4.17}$$

where $L = -\ln p(\boldsymbol{\pi}|\boldsymbol{\xi})$, $\boldsymbol{\pi} = (\pi_1, ..., \pi_n)$ describes the underlying (cosmological) parameters, and $\boldsymbol{\pi}_{\mathrm{ML}}$ denotes the maximum likelihood parameter vector. If we Taylor-expand the log-likelihood function in parameter space around $\boldsymbol{\pi}_{\mathrm{ML}}$ we derive

$$L(\boldsymbol{\pi}) = L(\boldsymbol{\pi}_{\mathrm{ML}}) + 0 + \frac{1}{2}(\boldsymbol{\pi} - \boldsymbol{\pi}_{\mathrm{ML}})^{\mathrm{t}} \, \mathbf{T}^{-1}(\boldsymbol{\pi} - \boldsymbol{\pi}_{\mathrm{ML}}) + O(\Delta_\pi^3) \, , \tag{4.18}$$

with

$$(\mathbf{T}^{-1})_{ij} = \left(\frac{\partial^2 L}{\partial \pi_i \, \partial \pi_j} \right)_{\pi=\pi_{\mathrm{ML}}} \, , \tag{4.19}$$

with $\mathbf{T}$ being the parameter covariance matrix. The first-order term in (4.18) vanishes since $(\partial L/\partial \boldsymbol{\pi})_{|\pi_{\mathrm{ML}}}$ is zero, hence (4.18) is dominated by the second-order term. Comparing (4.19) and (4.17) it turns out that the Fisher matrix is the expectation value of the inverse parameter covariance matrix. More illustratively, it measures the local curvature of the likelihood function around the maximum likelihood parameter set. To obtain likelihood contours we rewrite (4.18) replacing $\mathbf{T}^{-1}$) by the Fisher matrix. Hence, for a given fisher matrix we are able to calculate lower bounds on the likelihood contours. For the case that $p(\boldsymbol{\xi}|\boldsymbol{\pi})$ is Gaussian, which is a good approximation at least close to the maximum likelihood parameter vector ($p(\boldsymbol{\xi}|\boldsymbol{\pi}) \propto \exp(-L)$), one can directly express the Fisher matrix in terms of the data vectors and the data covariance matrix (e.g. Tegmark et al. 1997)

$$\mathbf{F}_{ij} = \frac{1}{2} tr \left[\mathbf{C}^{-1} \mathbf{C}_{,i} \mathbf{C}^{-1} \mathbf{C}_{,j} + \mathbf{C}^{-1} \mathbf{M}_{ij} \right] \, , \tag{4.20}$$

where $\mathbf{C}_{,i} \equiv \partial \mathbf{C}/\partial \pi_i$ denotes the derivative of the covariance matrix with respect to the ith component of the parameter vector and $\mathbf{M}_{ij} \equiv \hat{\boldsymbol{\xi}}_{,i} \hat{\boldsymbol{\xi}}_{,j}^{\mathrm{t}} + \hat{\boldsymbol{\xi}}_{,j} \hat{\boldsymbol{\xi}}_{,i}^{\mathrm{t}}$. The first term of (4.20) vanishes in case the covariance matrix is constant in parameter space, the second term vanishes in case of a constant mean data vector.

4.4 Impact of binning effects

As we describe in Sect. 4.1 the 2PCF data vectors are binned versions of a continuous quantity. In this context several questions arise: how much does the binning influence the parameter constraints? Is there an optimal binning? Does it make any difference if one uses linear or logarithmic bins? It is sensible to assume that there is a minimum threshold of bins, below which the information contained in the data is not fully transferred to the data vector. When going above this threshold, there must be a saturation limit in information at some bin number; including more bins cannot contribute further information.

We consider the angular range $[\vartheta_{\min} = 1'; \vartheta_{\max} = 180']$ and change the number of bins inside this interval. Note that whenever we refer to "number of bins", this is equivalent to the number of data points in only ξ_+. The full data vector consisting of ξ_+ and ξ_- contains twice as many bins, similarly any statement about e.g. 20 bins implies that we consider a 40×40 covariance matrix. We first examine the impact on the trace of the covariance and its inverse and second on the likelihood contours and values of q. Finally, we study the stability of the covariance during the inversion process as a function of bin width and give suggestions how to avoid/correct for possible numerical artifacts.

4.4.1 Binning effects on the trace of the covariance

The most obvious impact of binning on the covariances can be seen in (4.6), where the bin width enters explicitly in the shot noise term. In addition to the shot noise, the full covariance matrix consists of cosmic variance and mixed term (4.6) - (4.8). The bin width does not enter in the latter two terms, but the arguments in their Bessel functions are different when changing the bin size. From the analytic expression of $\mathbf{C}$ we expect its trace to increase when increasing the number of bins. This is clearly seen in Fig. 4.1, similar for linear (left panel) and logarithmic (right panel) binning. The effect is much stronger in the latter case for the reason that the $\Delta\vartheta$ for small ϑ are extremely small (such that the shot noise term dominates) in case of logarithmic binning. More of interest is the trace of the inverse covariance matrix. It is difficult to infer the behavior of $\mathrm{tr}\,\mathbf{C}^{-1}$ from the analytic expressions, because the different covariance terms mix during the inversion process. For the case of a pure shot noise matrix one can show that the trace of the inverse is independent of the bin width. Assume that the data vector is binned linearly with $\Delta\vartheta = (\vartheta_{\max} - \vartheta_{\min})/N$ and the $i - th$ bin can be expressed as $\vartheta_i = \vartheta_{\min} + \frac{2i-1}{2}\Delta\vartheta$. Then we can
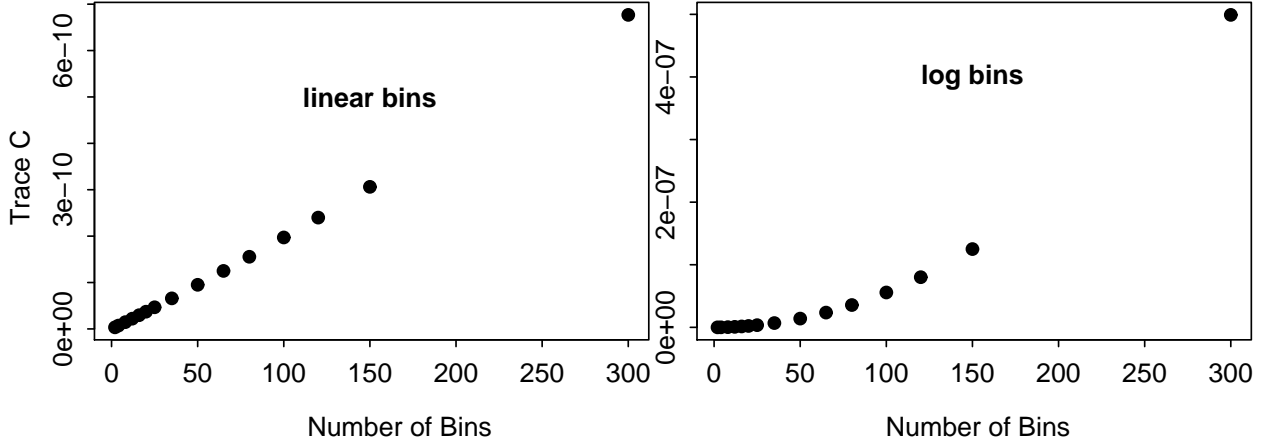
Figure 4.1: *The trace of the covariance* $\mathbf{C}_\xi$. *We compare linear (left panel) and logarithmic (right panel) binning for the following survey parameters:* $A = 512 \, deg^2$, $\sigma_\epsilon = 0.3$, *and* $\bar{n} = 15/arcmin^2$.

calculate the trace of the inverse shot noise covariance $\mathbf{S}$ as

$$
\begin{aligned}
tr\left(\mathbf{S}^{-1}\right) &= \sum_{i=1}^{N} \frac{2\pi A \bar{n}^2}{\sigma_\epsilon^4} \Delta\vartheta \, \vartheta_i \\
&= \frac{2\pi A \bar{n}^2}{\sigma_\epsilon^4} \frac{(\vartheta_{\max} - \vartheta_{\min})}{N} \sum_{i=1}^{N} \left[\vartheta_{\min} + \frac{(2i-1)}{2} \frac{(\vartheta_{\max} - \vartheta_{\min})}{N} \right] \\
&= \frac{2\pi A \bar{n}^2}{\sigma_\epsilon^4} \vartheta_{\min}(\vartheta_{\max} - \vartheta_{\min}) + \frac{2\pi A \bar{n}^2}{\sigma_\epsilon^4} \left(\frac{(\vartheta_{\max} - \vartheta_{\min})}{N} \right)^2 \underbrace{\sum_{i=1}^{N} \frac{2i-1}{2}}_{\frac{N^2}{2}} \\
&= \frac{\pi A \bar{n}^2}{\sigma_\epsilon^4} \left(\vartheta_{\max}^2 - \vartheta_{\min}^2 \right) ,
\end{aligned}
\tag{4.21}
$$

which shows that the trace of the inverse shot noise covariance depends only on minimum and maximum angular separation of the 2PCF.

In general, the covariance matrix is non-diagonal, therefore the structure of $\mathbf{C}^{-1}$ cannot be predicted, and we have to refer to numerical methods. Figure 4.2 shows the trace of $\mathbf{C}^{-1}$ as a function of bin number for linear (left panels) and logarithmic binning (right panels). Furthermore, we consider two different sets of survey parameters in order to quantify the impact of noise. The survey parameters in the upper row read $A = 512 \, \mathrm{deg}^2$, $\sigma_\epsilon = 0.3$, and $\bar{n} = 15/\mathrm{arcmin}^2$ and correspond to the values we use for the analysis in chapter 8. In the lower row we choose $A = 1 \, \mathrm{deg}^2$, $\sigma_\epsilon = 0.45$, and $\bar{n} = 10/\mathrm{arcmin}^2$, which results in a much higher noise contribution to the covariance, in particular the shot noise term increases. In general, all plots show a strong increase of $\mathrm{tr} \, \mathbf{C}^{-1}$ at small bin numbers, which slows down significantly when going to narrower binning.
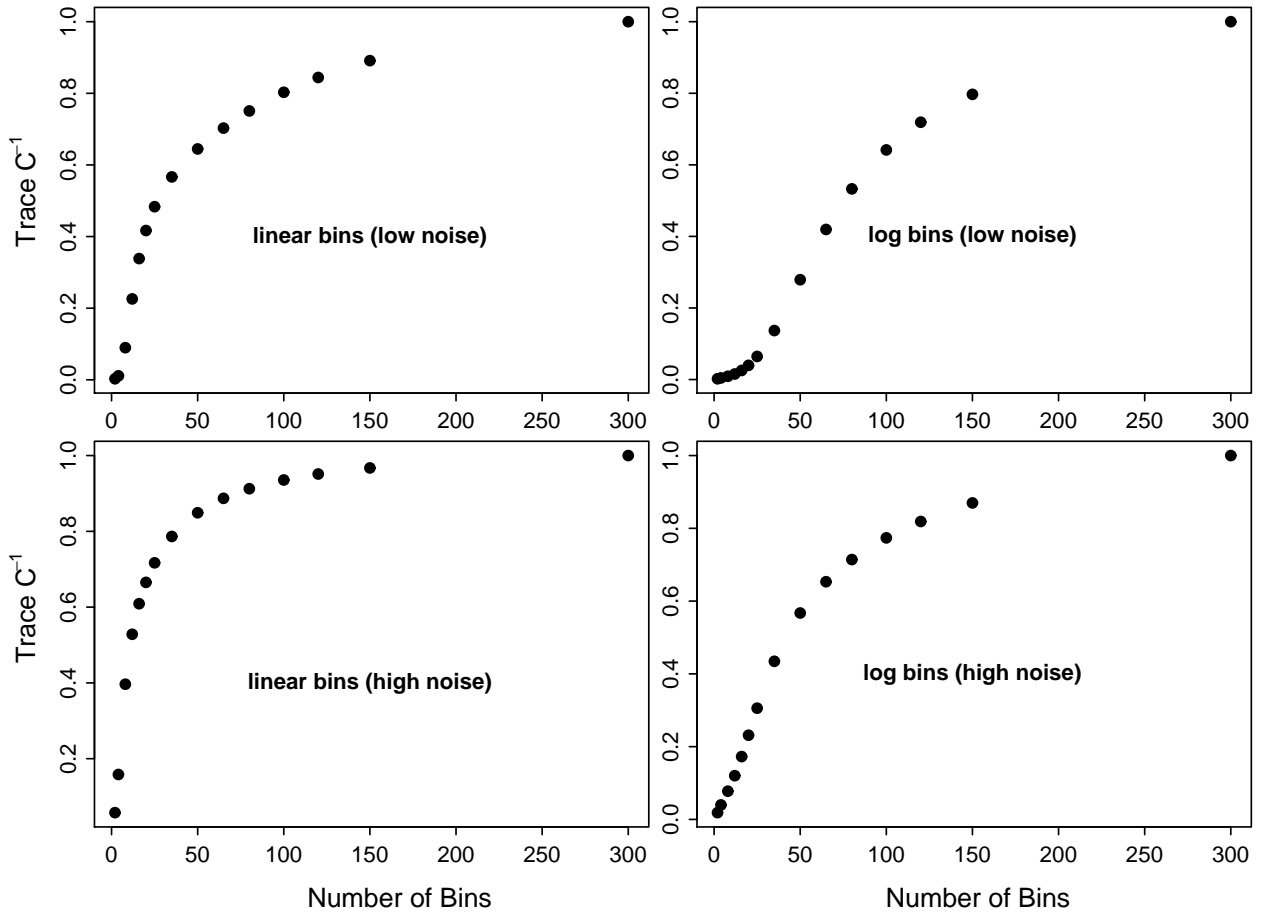
Figure 4.2: *The trace of the inverse covariance* $\mathbf{C}^{-1}$. *We compare linear (left panel) and logarithmic (right panel) binning for two different survey parameters:* $A = 512\ deg^2$, $\sigma_\epsilon = 0.3$, *and* $\bar{n} = 15/\mathrm{arcmin}^2$ *in the upper panels and* $A = 1\ deg^2$, $\sigma_\epsilon = 0.45$, *and* $\bar{n} = 10/\mathrm{arcmin}^2$ *in the lower panels. The inverse traces are normalized by tr* $\mathbf{C}^{-1}$ *for 300 bins*

We can explain this behavior by the fact that a small $\Delta\vartheta$ implies a relatively larger shot noise term compared to the mixed and cosmic variance term. Therefore, the full covariance matrix tends to behave as a pure shot noise matrix when arbitrarily increasing the number of bins. For the latter we have proven in (4.21), that its trace is independent of the binning. This explanation is supported by the fact that tr $\mathbf{C}^{-1}$ reaches its saturation limit earlier when assuming a high shot noise contribution (lower row of Fig. 4.2) compared to a low shot noise term (upper row of Fig. 4.2).

4.4.2   Binning effects on the likelihood contours

In order to quantify the impact of binning on the likelihood contours we consider the χ^2-function (4.13). There is good reason to assume that for the χ^2-function there exists a saturation limit, similar to that of tr $\mathbf{C}^{-1}$, where a further increase in the binning cannot result in a higher value of χ^2. We can prove this analytically for a simple case, namely when going from one bin to two, hence $\Delta\vartheta_1 = 2\Delta\vartheta_2$ (linear binning). We denote the shot noise term as S and the sum of the mixed and the cosmic variance term as $K = M + V$. The χ^2-value for one bin reads

$$\chi^2_{\Delta\vartheta_1} = \Delta^2_\xi (K + S)^{-1} \, , \tag{4.22}$$

with $\Delta_\xi = \xi_\pi - \hat{\xi}$. For the case of two bins the corresponding χ^2-value can be calculated via

$$\chi^2_{\Delta\vartheta_2} = (\Delta_{\xi_1} , \Delta_{\xi_2}) \left(\begin{array}{c|c} K_{11} + S_{11} & K_{12} \\ \hline K_{21} & K_{22} + S_{22} \end{array} \right)^{-1} \left(\begin{array}{c} \Delta_{\xi_1} \\ \Delta_{\xi_2} \end{array} \right) \tag{4.23}$$

The inverse of the covariance matrix can be calculated using

$$\left(\begin{array}{c|c} K_{11} + S_{11} & K_{12} \\ \hline K_{21} & K_{22} + S_{22} \end{array} \right) \times \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array} \right) = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & 1 \end{array} \right) . \tag{4.24}$$

For the reason that the covariance is symmetric we denote $A_{12} = A_{21} = A$; solving the corresponding system of equations (using that $K_{12} = K_{21} = K$) we calculate

$$A = -\frac{K}{(K_{22} + S_{22})(K_{11} + S_{11}) - K^2} \tag{4.25}$$

$$A_{11} = \frac{(K_{22} + S_{22})}{(K_{22} + S_{22})(K_{11} + S_{11}) - K^2} \tag{4.26}$$

$$A_{22} = \frac{(K_{11} + S_{11})}{(K_{22} + S_{22})(K_{11} + S_{11}) - K^2} . \tag{4.27}$$

Using the above expressions, the χ^2-value can be calculated as

$$\chi^2_{\Delta\vartheta_2} = \underbrace{\frac{\Delta^2_{\xi_1}(K_{22} + S_{22}) + \Delta^2_{\xi_2}(K_{11} + S_{11})}{(K_{22} + S_{22})(K_{11} + S_{11}) - K^2}}_{\text{diagonal terms}} - \underbrace{\frac{2\Delta_{\xi_1}\Delta_{\xi_2}K}{(K_{22} + S_{22})(K_{11} + S_{11}) - K^2}}_{\text{off-diagonal terms}} \tag{4.28}$$

We explicitly choose the above notation to stress that the χ^2-value is a trade-off between diagonal and off-diagonal terms. If we assume that the bin $\Delta\vartheta_1$ is already very narrow, increasing the binning further, hardly affects the bin estimators. In this limit we can approximate $\hat{\xi} \approx \hat{\xi}_1 \approx \hat{\xi}_2$. Intuitively, this is equivalent to the case where no additional information can be gained from the estimators. Further implications of this approximation are

$$K_{11} = K_{22} = K \qquad \text{and} \qquad S_{11} = S_{22} = 2S \, , \tag{4.29}$$

Table 4.1: *The values of q in units of 10^{-4} depending on the number of bins for the Ω_{m} vs σ_8 likelihood contours.*

# bins	4	8	12	16	20	25	35	50	65	80	100	120	150
q (lin bins)	5.69	1.77	1.28	1.17	1.15	1.13	1.08	1.01	0.97	0.95	0.94	0.93	0.93
q (log bins)	1.28	0.97	0.93	0.93	0.92	0.92	0.92	0.92	0.92	0.91	0.91	0.91	0.91

where the last equation results from the fact that $\Delta\vartheta_1 = 2\Delta\vartheta_2$ and the shot noise term scales with the inverse bin width. Using these approximations we arrive at

$$\chi^2_{\Delta\vartheta_2} \;=\; \frac{2\Delta_\xi^2(K + 2S) - 2\Delta_\xi^2 K}{(K + 2S)^2 - K^2} \tag{4.30}$$

$$\;=\; \Delta_\xi^2 (K + S)^{-1}, \tag{4.31}$$

which is equivalent to (4.22). Although one might argue that this calculation is somewhat artificial for the reason that we consider the limit $\Delta\vartheta \longrightarrow 0$, it nicely illustrates how off-diagonal terms compensate for the diagonal terms, in case the binning becomes very narrow. For arbitrary number of bins we again refer to numerical methods and perform a likelihood analysis employing the above examined covariances and their corresponding 2PCF data vectors. Figure 4.3 shows a sample of contour plots; we see a small difference for the linear case when going from 8 bins to 12 and further to 50. Going from 50 to 150 bins does not give any tighter constraints. For the case of logarithmic binning there is a small deviation between 4 and 8 bins but further differences cannot be seen by eye. We illustrate all results using the values of q (see table 4.1 and Fig. 4.4). For both cases we find that the q (and therefore the size of the likelihood contours) hardly changes when going beyond 35 bins. Note, that the relative change in q for the case of logarithmic binning is much smaller compared to linear bins, hence the first samples the information of the underlying cosmological parameters much better than the latter binning. For this reason we use logarithmic binning throughout the whole thesis.

4.4.3 Binning effects on the stability of C under inversion

There is another important aspect in the context of binning effects. If one uses too many bins one might run into numerical difficulties during the inversion process of the covariance matrix. The reason for this is that a high number of bins, implying a small bin width, results in high correlation terms in the covariance. At some point the covariance matrix becomes unstable under numerical transformations, so-called ill-conditioned. Note this does not necessarily imply that the covariance matrix is not invertible; an ill-conditioned covariance matrix can have a unique inverse, which can be calculated through various methods (e.g. LU decomposition, Gaussian elimination, etc). Nevertheless, this inverse is dominated by numerical artifacts. As a result, the likelihood analysis no longer fits the parameters according to the information provided by the
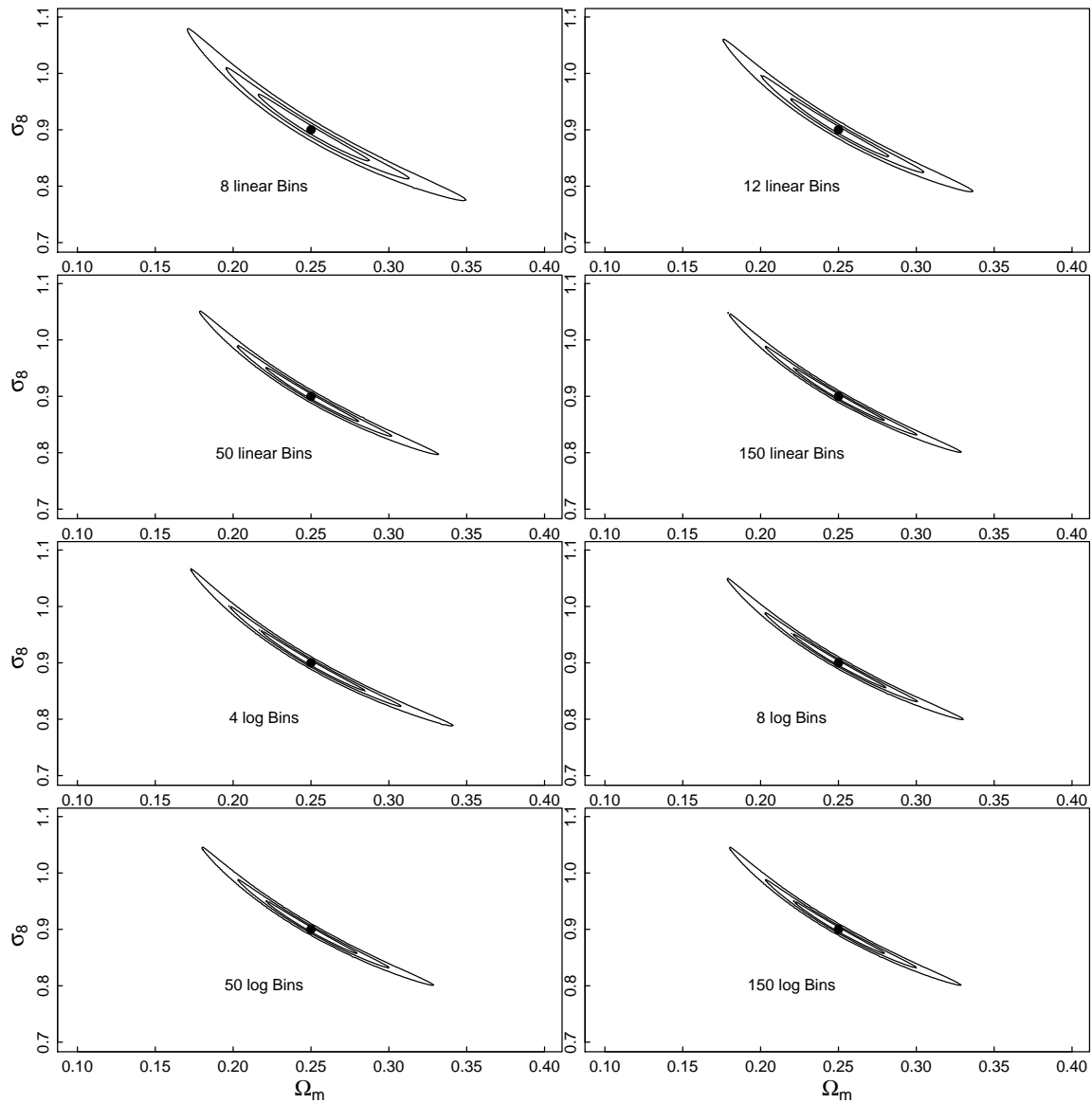
Figure 4.3: *The likelihood contours when increasing the number of bins inside the fixed interval* $[\vartheta_{\min} = 1'; \vartheta_{\max} = 180']$. *The upper four panels show the result for linear binning, the lower correspond to logarithmic binning. The survey parameters are* $A = 512\ deg^2$, $\sigma_\epsilon = 0.3$, *and* $\bar{n} = 15/\mathrm{arcmin}^2$.
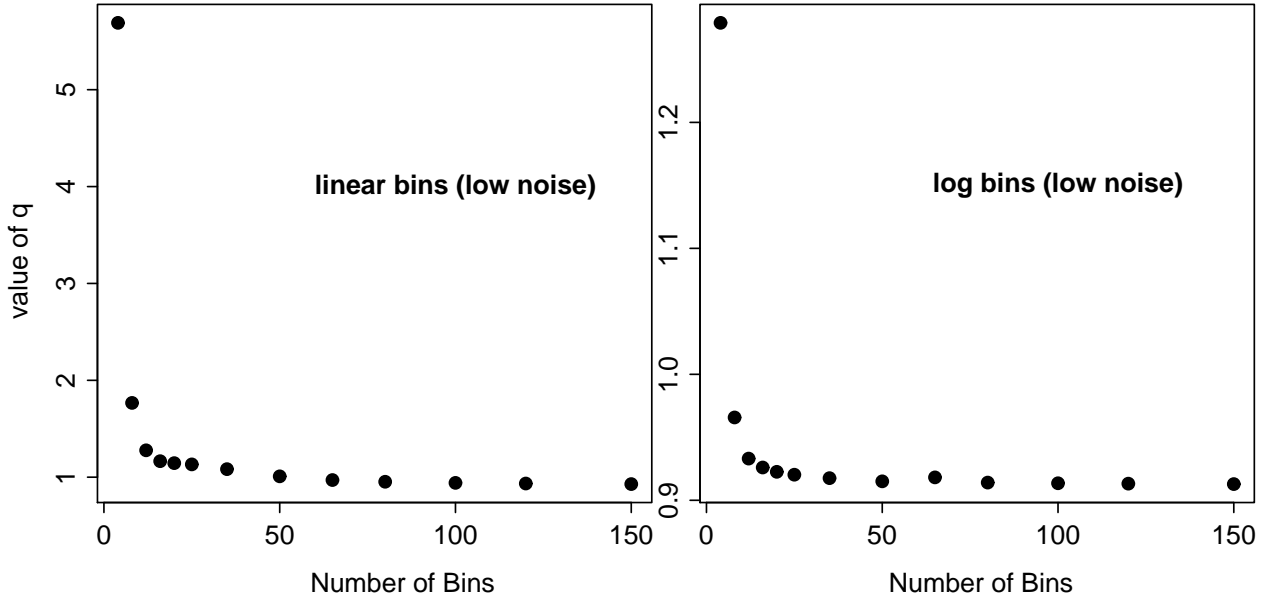
Figure 4.4: *This figure shows the values of q in units of 10^{-4} depending on the number of bins for linear (left panel) and logarithmic (right panel) binning. The q quantify the size of the likelihood contours in the considered Ω_m vs σ_8 parameter space (see Fig. 4.3)*

data, but rather fits the numerical artifacts which occur in the inversion process. In order to check for this, we perform a *singular value decomposition (SVD)*[1] and examine the *condition number*, which is defined as the ratio of maximum to minimum singular value, $c = \sigma_{max}/\sigma_{min}$. Note that the condition number has the same value for the covariance matrix and its inverse. In case the condition number exceeds the inverse of the numerical precision (in our case double precision) the matrix is certainly ill-conditioned. We require $c < 10^{10}$, which means that roundoff errors in the last digit of σ_{max} would only affect σ_{min} at the 0.0001 percent level.

Figure 4.5 shows c depending on the number of bins; we see that the condition number is much larger for logarithmic binning. However, in all cases the condition number is below the required threshold, hence we are confident that our inverse covariance matrix is not affected by noise artifacts.

As an additional cross check we consider the eigenvalues of the covariance matrix. For the case of a positive definite and symmetric matrix an eigenvalue decomposition is equivalent to an SVD. If a deviation between eigenvalues and singular values occurs, the eigenvalue decomposition is most likely affected by numerical artifacts (according to Press et al. 1992, the SVD is stabler). This comparison of eigenvalues and singular values should be performed for the in-

---

[1]An SVD decomposes a $m \times n$ matrix A (m,n are arbitrary) into $\mathbf{A} = \mathbf{U\Sigma V^t}$, where $\mathbf{U}(m \times m)$ and $\mathbf{V}(n \times n)$ are orthogonal matrices ($\mathbf{U^t} = \mathbf{U^{-1}}$). $\mathbf{\Sigma}(m \times n)$ has zero entries everywhere except on the diagonal which contains the singular values σ in descending order. Note that "diagonal" for the case of $m \neq n$ refers to the diagonal of the [$\min(m,n) \times \min(m,n)$] submatrix. For more information see Trefethen & Bau (1997). This relation also holds for complex $\mathbf{A}$, where the orthogonal $\mathbf{U}, \mathbf{V}$ are replaced by unitary matrices.
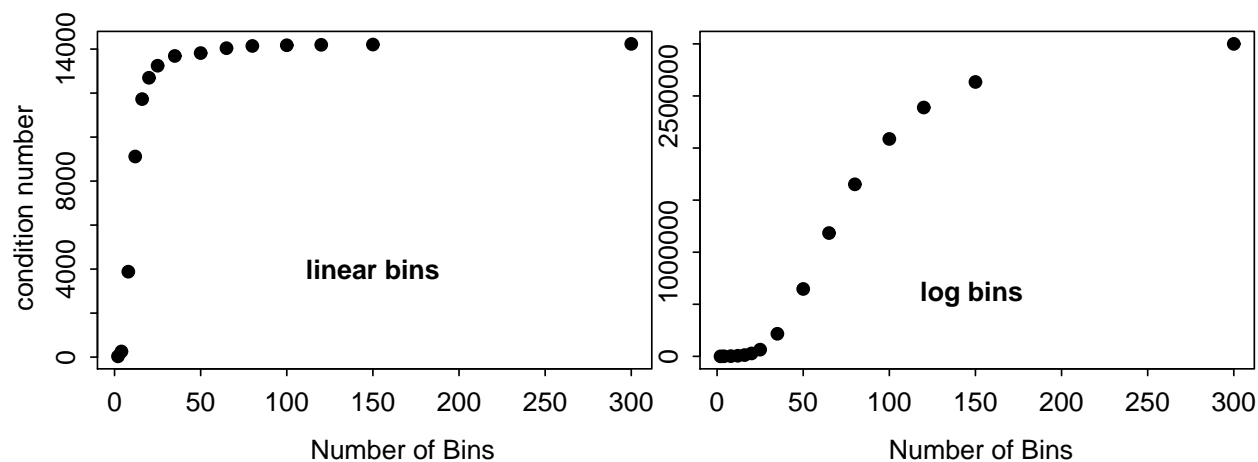
Figure 4.5: *This figure shows the condition number $c = \sigma_{max}/\sigma_{min}$ as a function of bin number. The left panel illustrates the results for linear and the right panel for logarithmic binning.*

verse covariance matrix. If one finds a deviation between both, this is a good indication that the inversion process causes numerical artifacts, which suggests that the original covariance was ill-conditioned.

In this thesis, this problem occurs in chapter 8, when we compare covariances with and without shot noise. Noise free covariances become ill-conditioned at much lower bin number compared to covariances with shot noise. Still, we can derive parameter constraints by excluding the singular values which are dominated by numerical artifacts. In order to determine the latter, we compare eigenvalues and singular values, and considered those as contaminated where a deviation occurs. We then calculate the inverse covariance using a singular value decomposition and set all contaminated inverse singular values to 0. This methods gives the so-called pseudo-inverse of **C**. Keeping the contaminated inverse singular values is no alternative; in general these have a huge impact on the likelihood analysis. The reason for this is that these formerly small singular values dominate the inverse covariance.

Chapter 5

Improvement of cosmic shear data vectors

Currently, most cosmic shear surveys only consider second-order shear statistics, for which all information is contained in the power spectrum of the convergence (P_κ). Although P_κ is not directly measureable, it is linearly related to second-order cosmic shear measures (e.g. the two-point correlation function and the aperture mass dispersion), which can be estimated from the distorted ellipticities of the observed galaxies. More precisely, all second-order measures are filtered versions of P_κ, and the corresponding filter functions determine how the information content of P_κ is sampled (see Sect. 3.3.3).

In this chapter we compare several data vectors of cosmic shear measures and create an optimal data vector with high information content, largely uncorrelated data points and only little sensitivity to a possible B-mode contamination. We prove a general statement that a data vector consisting of 2PCF data points ($\boldsymbol{\xi}$) always gives tighter constraints on cosmological models compared to a data vector consisting of $\langle M_{\mathrm{ap}}^2 \rangle$ data points ($\langle \boldsymbol{M_{\mathrm{ap}}^2} \rangle$) and we confirm this by a likelihood analysis of ray-tracing simulations. The results of this chapter are published in Eifler et al. (2008a).

5.1 Estimators and covariances

5.1.1 The aperture mass dispersion $\langle M_{\mathrm{ap}}^2 \rangle$

For the case of the 2PCF we already introduced estimators and covariances in Sect. 4.1. For the general case of E- and B-modes and a continuous 2PCF, the aperture mass dispersion can be calculated via (3.81). Similarly, an unbiased estimator of $\langle M_{\mathrm{ap}}^2 \rangle$ can be calculated from $\hat{\xi}_\pm(\vartheta)$ via

$$\mathcal{M}(\theta_k) = \sum_{i=1}^{I} \frac{\Delta \vartheta_i \vartheta_i}{2 \, \theta_k^2} \left[\hat{\xi}_+(\vartheta_i) \, T_+\left(\frac{\vartheta_i}{\theta_k}\right) + \hat{\xi}_-(\vartheta_i) \, T_-\left(\frac{\vartheta_i}{\theta_k}\right) \right], \tag{5.1}$$

where I must be chosen such that the upper limit of the Ith bin equals twice the value of θ_k. Using (5.1) we can calculate the covariances of the aperture mass dispersion $\text{C}_{\mathcal{M}}$ in terms of C_ξ:

$$
\begin{aligned}
\text{C}_{\mathcal{M}}(\theta_k, \theta_l)) \;=\; & \frac{1}{4} \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{\Delta\vartheta_i \Delta\vartheta_j}{\theta_k^2 \theta_l^2}\, \vartheta_i \vartheta_j \\
& \times \left[\sum_{m,n=+,-} T_m\!\left(\frac{\vartheta_i}{\theta_k}\right) T_n\!\left(\frac{\vartheta_j}{\theta_l}\right) \text{C}_{mn}(\vartheta_i, \vartheta_j) \right].
\end{aligned}
\tag{5.2}
$$

Similar to (5.1), I (J) are chosen such that the upper limit of the Ith (J^{th}) bin equals twice of θ_k (θ_l). The exact expressions for the T-functions are given in (3.64) and (3.65).

5.1.2 The new data vector $\mathcal{N}$

We first compare the information content of the two-point correlation function (2PCF) and aperture mass dispersion ($\langle M_{\text{ap}}^2 \rangle$). For the 2PCF the data vector reads

$$
\boldsymbol{\xi} = \begin{pmatrix} \boldsymbol{\xi}_+ \\ \boldsymbol{\xi}_- \end{pmatrix} \quad \text{with} \quad \boldsymbol{\xi}_+ = \begin{pmatrix} \xi_+(\vartheta_1) \\ \vdots \\ \xi_+(\vartheta_m) \end{pmatrix}, \; \boldsymbol{\xi}_- = \begin{pmatrix} \xi_-(\vartheta_1) \\ \vdots \\ \xi_-(\vartheta_m) \end{pmatrix}
\tag{5.3}
$$

In case of the aperture mass dispersion we consider

$$
\langle \boldsymbol{M_{\text{ap}}^2} \rangle = \begin{pmatrix} \langle M_{\text{ap}}^2 \rangle(\theta_1) \\ \vdots \\ \langle M_{\text{ap}}^2 \rangle(\theta_n) \end{pmatrix}.
\tag{5.4}
$$

The relation (5.1) can also be written in terms of data vectors and an $n \times 2m$ *transfer matrix* $\mathbf{A}$

$$
\langle \boldsymbol{M_{\text{ap}}^2} \rangle = \underbrace{\left(\; \mathbf{A}_+ \; | \; \mathbf{A}_- \; \right)}_{\mathbf{A}} \begin{pmatrix} \boldsymbol{\xi}_+ \\ \boldsymbol{\xi}_- \end{pmatrix},
\tag{5.5}
$$

with $\mathbf{A}_+$ denoting the part of $\mathbf{A}$ referring to $\boldsymbol{\xi}_+$ and $\mathbf{A}_-$ denotes the corresponding part referring to $\boldsymbol{\xi}_-$. Eq. (5.5) implies that the information content of $\langle M_{\text{ap}}^2 \rangle$ is less than or equal to $\boldsymbol{\xi}$. The amount of information can be equal if and only if the rank of $\mathbf{A}$ equals the dimension of $\boldsymbol{\xi}$, hence rank $\mathbf{A} = 2m$. We explicitly prove these statements in the Appendix A. For the case of $\boldsymbol{\xi}$ and $\langle M_{\text{ap}}^2 \rangle$, $n \leq m$ holds, which can be seen from (5.1). Therefore the relation (5.5) is not invertible and the information content of $\langle M_{\text{ap}}^2 \rangle$ is smaller compared to $\boldsymbol{\xi}_\pm$. That $\boldsymbol{\xi}_\pm$ contains more information on cosmological parameters can also be explained when looking at the filter functions J_0, J_4, and W_{ap} relating the corresponding second-order shear measures to the underlying power spectrum. The 2PCF, especially ξ_+, probes the power spectrum over a broad range of Fourier modes and also collects information on scales larger than the survey size. In contrast, the aperture mass dispersion provides a highly localized probe of the power spectrum and does not contain this

large-scale information. Hence, due to the limited field size of a survey, the information content of $\langle M_{\mathrm{ap}}^2 \rangle$ is smaller compared to $\xi_\pm$. Nevertheless $\langle M_{\mathrm{ap}}^2 \rangle$ has important advantages. First, it can be used to separate E-modes and B-modes (Crittenden et al. 2002; Schneider et al. 2002b); more precisely $\langle M_{\mathrm{ap}}^2 \rangle$ is only sensitive to E-modes. Second, thanks to its narrow filter function, two different $\langle M_{\mathrm{ap}}^2 \rangle$ data points are much less correlated compared to the 2PCF. Third, $\langle M_{\mathrm{ap}}^2 \rangle$ can be easier extended to higher-order statistics (Schneider et al. 2005). These advantages are valuable and should be maintained, but the information content should be improved. Hence, we extend the $\langle M_{\mathrm{ap}}^2 \rangle$ data vector by one data point of $\xi_+(\theta_0)$, which provides the large-scale information of P_κ and call this new data vector $\boldsymbol{N}$

$$\boldsymbol{N} = \begin{pmatrix} \langle M_{\mathrm{ap}}^2 \rangle(\theta_1) \\ \vdots \\ \langle M_{\mathrm{ap}}^2 \rangle(\theta_n) \\ \xi_+(\theta_0) \end{pmatrix}. \tag{5.6}$$

The corresponding covariance matrix reads

$$\mathbf{C}_{\mathcal{N}} = \left(\begin{array}{ccc|c} \mathrm{C}_{\mathcal{M}_{11}} & \cdots & \mathrm{C}_{\mathcal{M}_{1n}} & \mathrm{C}(\mathcal{M}_1, \xi_+) \\ \vdots & \ddots & \vdots & \vdots \\ \mathrm{C}_{\mathcal{M}_{1n}} & \cdots & \mathrm{C}_{\mathcal{M}_{nn}} & \mathrm{C}(\mathcal{M}_n, \xi_+) \\ \hline \mathrm{C}(\xi_+, \mathcal{M}_1) & \cdots & \mathrm{C}(\xi_+, \mathcal{M}_n) & \mathrm{C}(\xi_+, \xi_+) \end{array} \right). \tag{5.7}$$

The upper left $n \times n$ matrix is exactly $\mathbf{C}_{\mathcal{M}}$ and the entry for $\mathrm{C}(\xi_+, \xi_+)$ is taken from the corresponding covariance matrix of the correlation function. The cross terms can be calculated using (5.1) to read

$$\begin{aligned} \mathrm{C}(\mathcal{M}(\theta_k), \hat{\xi}_+(\theta_0)) &= \frac{1}{2} \sum_{i=1}^{I} \frac{\Delta \vartheta_i}{\theta_k^2} \, \vartheta_i \left[T_+\left(\frac{\vartheta_i}{\theta_k}\right) \mathrm{C}_{++}(\vartheta_i, \theta_0) \right. \\ &\quad \left. + T_-\left(\frac{\vartheta_i}{\theta_k}\right) \mathrm{C}_{-+}(\vartheta_i, \theta_0) \right]. \end{aligned} \tag{5.8}$$

Similar to $\mathbf{C}_{\mathcal{M}}$, $\mathbf{C}_{\mathcal{N}}$ is almost diagonal, hence data points of different angular scales are weakly correlated. For the corss terms derived through 5.8, the strength of this correlation varies with the considered scale θ_k, This property vanishes in case we add more than one data point of the 2PCF. Due to their broad filter function J_0, two or more 2PCF data points would be strongly correlated, resulting in large off-diagonal terms in $\mathbf{C}_{\mathcal{N}}$. The main intention, namely to include the information of the power spectrum on large angular scales, is already fulfilled by adding one data point of ξ_+. For small ℓ the sampling of the power spectrum by $\xi_+(\theta)$ hardly depends on θ, therefore the gain in information on those scales by adding more than one 2PCF data point would be rather small. In the following we perform a likelihood analysis for $\boldsymbol{N}$, examine its ability to constrain cosmological parameters, and compare it to the two aforementioned data vectors. Note, that we consider only E-modes until Sect. 5.3.

5.2 Calculating data vectors and covariances

For the numerical comparison we calculate the data vectors $\boldsymbol{\xi}, \langle M_{\mathrm{ap}}^2 \rangle, \mathcal{N}$ directly from the power spectrum of density fluctuations P_δ using (3.44) to obtain P_{E} and then applying either (3.51), (3.52), or (3.57) depending on the desired cosmic shear measure. To derive P_δ we assume an initial Harrison-Zeldovich power spectrum ($P_i(k) \propto k_{\mathrm{s}}^n$ with $n_{\mathrm{s}} = 1$). The transition to today's power spectrum employs the transfer function described in Efstathiou et al. (1992), and we use the fitting formula of Smith et al. (2003) to calculate the non-linear evolution. In contrast, the covariances are obtained from ray-tracing simulations. The N-body simulation used for the ray-tracing experiment was carried out by the Virgo Consortium (Jenkins et al. 2001a); for details of the ray-tracing algorithm see Ménard et al. (2003). Then, $\mathbf{C}_\xi$ is calculated by field-to-field variation of 36 ray-tracing realizations, where each field has a sidelength of 4.27 degrees. The intrinsic ellipticity noise is $\sigma_\epsilon = 0.3$ and the number density of source galaxies is given by $\bar{n} = 25/\mathrm{arcmin}^2$. From $\mathbf{C}_\xi$ we calculate $\mathbf{C}_\mathcal{M}$ and $\mathbf{C}_\mathcal{N}$ according to (5.2) and (5.7). Our fiducial cosmological model is determined by the cosmology of the ray-tracing simulations, i.e. a flat ΛCDM model with $\Omega_{\mathrm{m}} = 0.3$, $\sigma_8 = 0.9$, $h = 0.7$, and $\Gamma = 0.172$. Furthermore, the ray-tracing simulations assume all source galaxies to be at the same redshift, i.e. $z_0 = 0.98$. Using a redshift distribution instead would not change our results markedly. The data vectors are calculated from P_{E}, and for a given redshift distribution, one can find a characteristic z_0 such that P_{E} is almost similar independent of using the redshift distribution or choosing all sources to be at z_0.

5.2.1 Difficulties with covariances

5.2.1.1 Underestimation of $\mathbf{C}_\mathcal{M}$

Kilbinger et al. (2006) have shown that $\langle M_{\mathrm{ap}}^2 \rangle(\theta)$ is biased for small θ when calculated from the 2PCF using (5.1). This is due to the lack of 2PCF data points on very small angular scales, which causes a small-scale cutoff in the integral of (3.63). In our specific case the $\langle M_{\mathrm{ap}}^2 \rangle$ data vector is not affected by this bias because we calculate it directly from the power spectrum P_{E}. However, since $\mathbf{C}_\mathcal{M}$ and $\mathbf{C}_\mathcal{N}$ are calculated from the covariance of the 2PCF, they are certainly affected by this problem. In this subsection we determine the θ-range on which we can calculate $\mathbf{C}_\mathcal{M}$ with sufficient accuracy; the corresponding data vector of the aperture mass dispersion will be restricted to this range. Figure 5.1 shows $\langle M_{\mathrm{ap}}^2 \rangle$ calculated directly from the power spectrum using (3.57) compared with $\langle M_{\mathrm{ap}}^2 \rangle$ calculated from $\xi_\pm$ using (5.1). We assume that the deviation shown here yields a measure of the bias in $\mathbf{C}_\mathcal{M}$ and we require an accuracy of 5 % to accept a θ-value for the $\langle M_{\mathrm{ap}}^2 \rangle$ data vector. The deviation on large angular scales in Fig. 5.1 is due to the limited range of 2PCFs $0\rlap{.}'2 - 200\rlap{.}'0$ from which $\langle M_{\mathrm{ap}}^2 \rangle$ (circles) is calculated. In order to assure that $\mathbf{C}_\mathcal{M}$ (5.8) is calculated properly we restrict the data vector to a θ-range of $2\rlap{.}'25 - 100\rlap{.}'0$.

5.2.1.2 Inversion of the covariance matrix

A second difficulty in the context of covariance matrices is outlined in Hartlap et al. (2007). The fact that an inversion of an estimated unbiased covariance matrix leads to a biased result can
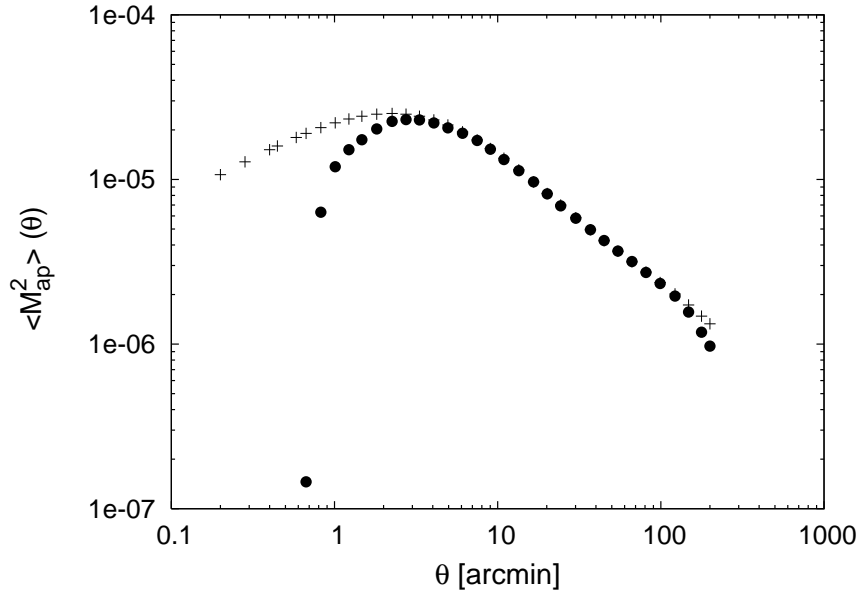
Figure 5.1: $\langle M^2_{ap} \rangle$ *calculated directly from the power spectrum (crosses) compared to* $\langle M^2_{ap} \rangle$ *calculated from* $\xi_{\pm}$ *(circles).*

be overcome by applying a correction factor. According to Hartlap et al. (2007), the correction factor depends on the ratio of number of bins (B) to number of independent realizations (N) from which the covariance matrix is estimated. An unbiased estimate of the inverse covariance matrix is

$$\mathbf{C}^{-1}_{\text{unbiased}} = \frac{N - B - 2}{N - 1} \mathbf{C}^{-1} = \left[1 - \frac{B + 1}{N - 1} \right] \mathbf{C}^{-1}. \tag{5.9}$$

Hartlap et al. (2007) prove the validity of this correction factor for the case of Gaussian errors and statistically independent data vectors. These two assumptions are violated when estimating the covariance matrix from ray-tracing simulations. As a check of whether the correction factor corrects the error in our ray-tracing covariance matrices, we perform the following experiment. We add different Gaussian noise to the ellipticities of the galaxies, which are taken from the 36 independent realizations of the ray-tracing simulations and thereby increase the number of independent realizations. We hold the binning of the matrices constant, calculate covariances for 36, 108, 216, 360, 720, 1080, 1440, 1800 realizations and plot $1/\text{tr}\ \mathbf{C}^{-1}$ depending on the ratio B/N (Fig. 5.2). Note that this method only creates multiple realizations of Gaussian noise on the galaxy ellipticities and does not increase the number of realizations that determine the cosmic variance part of the covariance matrix. Therefore, this method only partly checks for the non-Gaussianity of the errors in a ray-tracing covariance matrix; nevertheless, the impact of statistically dependent data vectors is fully taken into account. We find the same linear behavior of the bias as Hartlap et al. (2007), so are confident that the correction factor is able to unbias our covariance matrices. By using the corrected inverse covariance matrix we assure that the log-likelihood is also unbiased; nevertheless, any non-linear transformation of the log-likelihood will again introduce a bias that influences the results and must be examined.
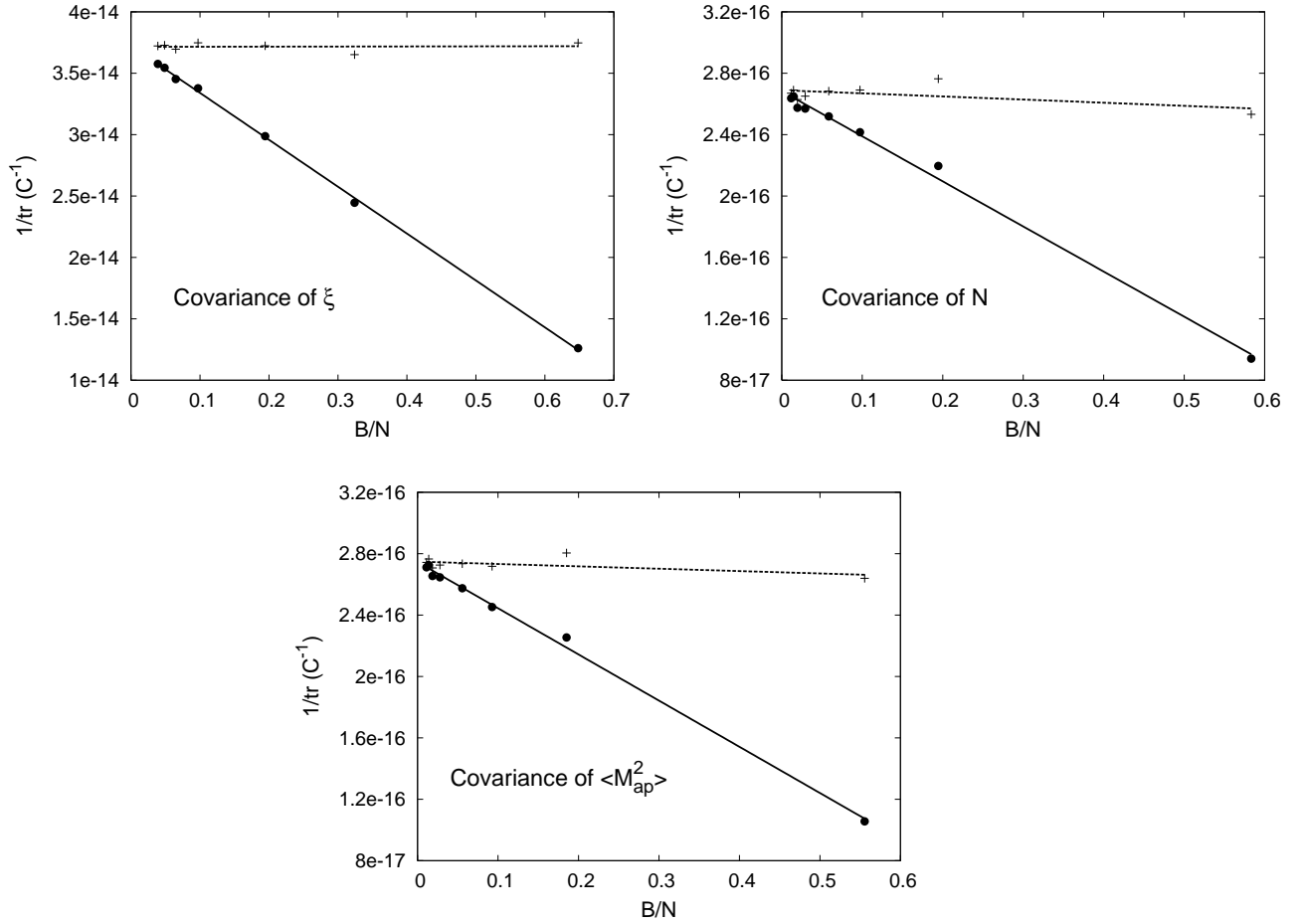
Figure 5.2: *Illustration of how to correct for the bias occurring in the inverted covariance matrices. We plot 1/trace ($\mathbf{C}^{-1}$) against the ratio B/N, where B is the number of bins in the covariance matrix and N the number of independent realizations of the ray-tracing simulations. We show the correction for the inverse of $\mathbf{C}_\xi$ (top left), $\mathbf{C}_\mathcal{N}$ (top right), and $\mathbf{C}_\mathcal{M}$ (bottom); each plot shows the corrected (crosses) and uncorrected (circles) values with the corresponding linear fits through the data points.*

5.2.2   Likelihood analysis

In this section we compare the information content of the three aforementioned data vectors, using a covariance derived from ray-tracing simulations. To quantify the information content we perform likelihood analyses for several parameter combinations in two-, three- and four-dimensional parameter space. In addition we calculate the values of q which we introduced in Sect. 4.2. As can be seen from (4.15) and (4.16), q is non-linear in the log-likelihood, therefore it is also affected from the bias described in the last section. In order to quantify the strength of this bias, we carry out the same analysis as for the covariance matrices in Sect. 5.2.1.2. For six different numbers of independent realizations we perform a likelihood analysis in a two-
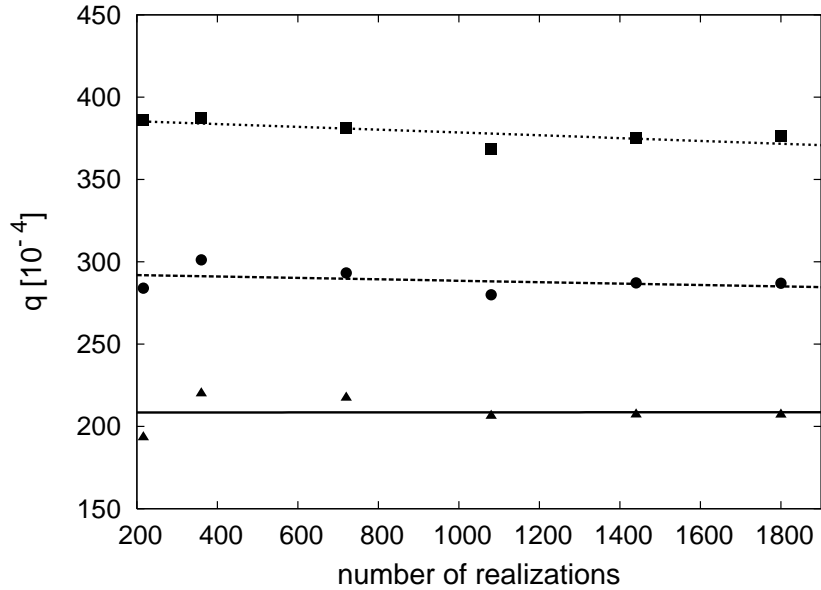
Figure 5.3: *Dependence of q calculated for ξ (triangles), N (circles), $\langle M_{ap}^2 \rangle$ (squares) on the numbers of independent realizations of the ray-tracing simulation. The values of q are calculated for the case of varying Ω_m vs. σ_8 in parameter space. The lines are a linear fit through data points (solid for ξ, dashed for N, dotted for $\langle M_{ap}^2 \rangle$). Note that the deviation of q belonging to different numbers of realizations is much smaller than the difference of q of different data vectors.*

parameter space (Ω_m vs. σ_8) and calculate q for all three cosmic shear measures. The result is plotted in Fig. 5.3. One clearly sees that the q dependence on the number of realizations is much weaker compared to the difference between q of different cosmic shear measures. Therefore, the bias is small and we can confidently use q to compare the relative information content of the different data vectors.

5.2.3 Variation of two parameters

The likelihood analysis in this section was performed in a two-dimensional parameter space; all other cosmological parameters were fixed to the fiducial values. Before comparing the three data vectors we optimized N with respect to the θ_0-value of the added 2PCF data point. We added 35 different $\xi_+(\theta_0)$ covering a range $\theta_0 \in [0.'2\text{-}200.'0]$ and calculated q. Figure 5.4 illustrates the results of this optimization for 3 different pairs of parameters (Γ vs. Ω_m, σ_8 vs. Ω_m, z_0 vs. Ω_m). For all parameter combinations considered, the optimal θ_0 is close to $10'$. This can be explained from the behavior of the covariance matrix. For small angular scales the covariance is dominated by shot noise, whereas the signal of ξ_+ becomes very small for large angular scales. In both cases the signal-to-noise ratio is lower than on medium angular scales, where we find the minimum of q. In our later analysis we always chose the optimal 2PCF data point for the combined data vector. The results are illustrated by contour plots (Fig. 5.5) and the corresponding values of q are summarized in Table 5.1. Here, we also list the results for two additional parameter
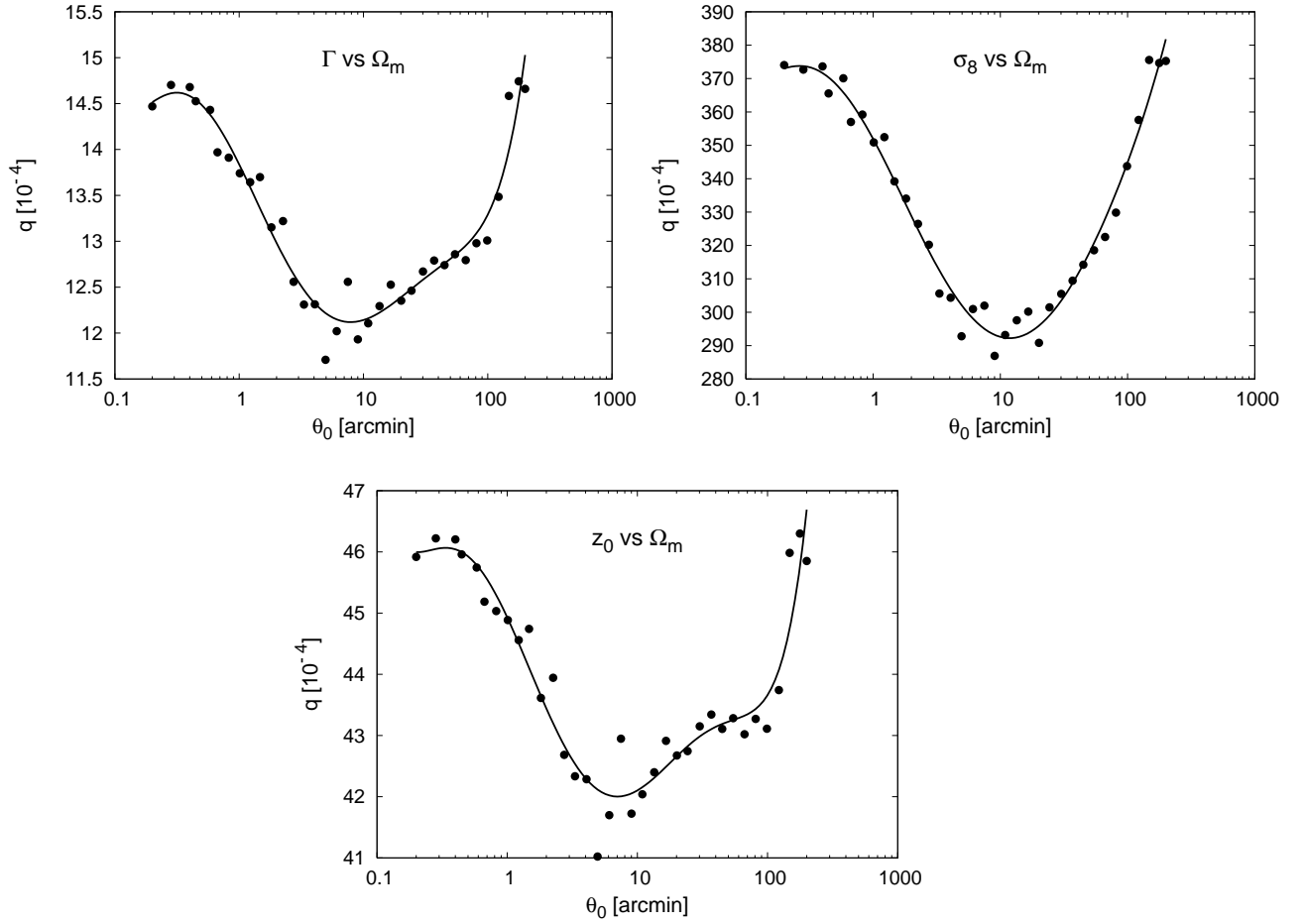
Figure 5.4: *Values of q of the combined data vector $\mathcal{N}$ depending on the angular separation (θ_0) of the additional ξ_+ data point. We consider three different parameter spaces Γ vs. Ω_m (top left), σ_8 vs. Ω_m (top right), and z_0 vs. Ω_m (bottom), and for each case we calculated q for 35 different ϑ_0. The solid lines indicate a polynomial fit through the data; the minima of these fits and therefore the optimal values for θ_0 are $7\!'\!.8$ (Γ vs. Ω_m), $12\!'\!.9$ (σ_8 vs. Ω_m), and $7\!'\!.0$ (z_0 vs. Ω_m).*

combinations, σ_8 vs. Γ and z_0 vs. σ_8, not shown in Fig. 5.5. One clearly sees that the 2PCF data vector gives the tightest constraints on cosmological parameters, whereas constraints from the aperture mass dispersion are weaker. Although not quite matching the amount of information of ξ, the combined data vector is a substantial improvement over $\langle M_{\mathrm{ap}}^2 \rangle$. This result is consistent for all parameter combinations we examine; nevertheless, the amount of the improvement varies. We calculated the difference in information of ξ and $\mathcal{N}$ relative to $\langle M_{\mathrm{ap}}^2 \rangle$ and denote these values
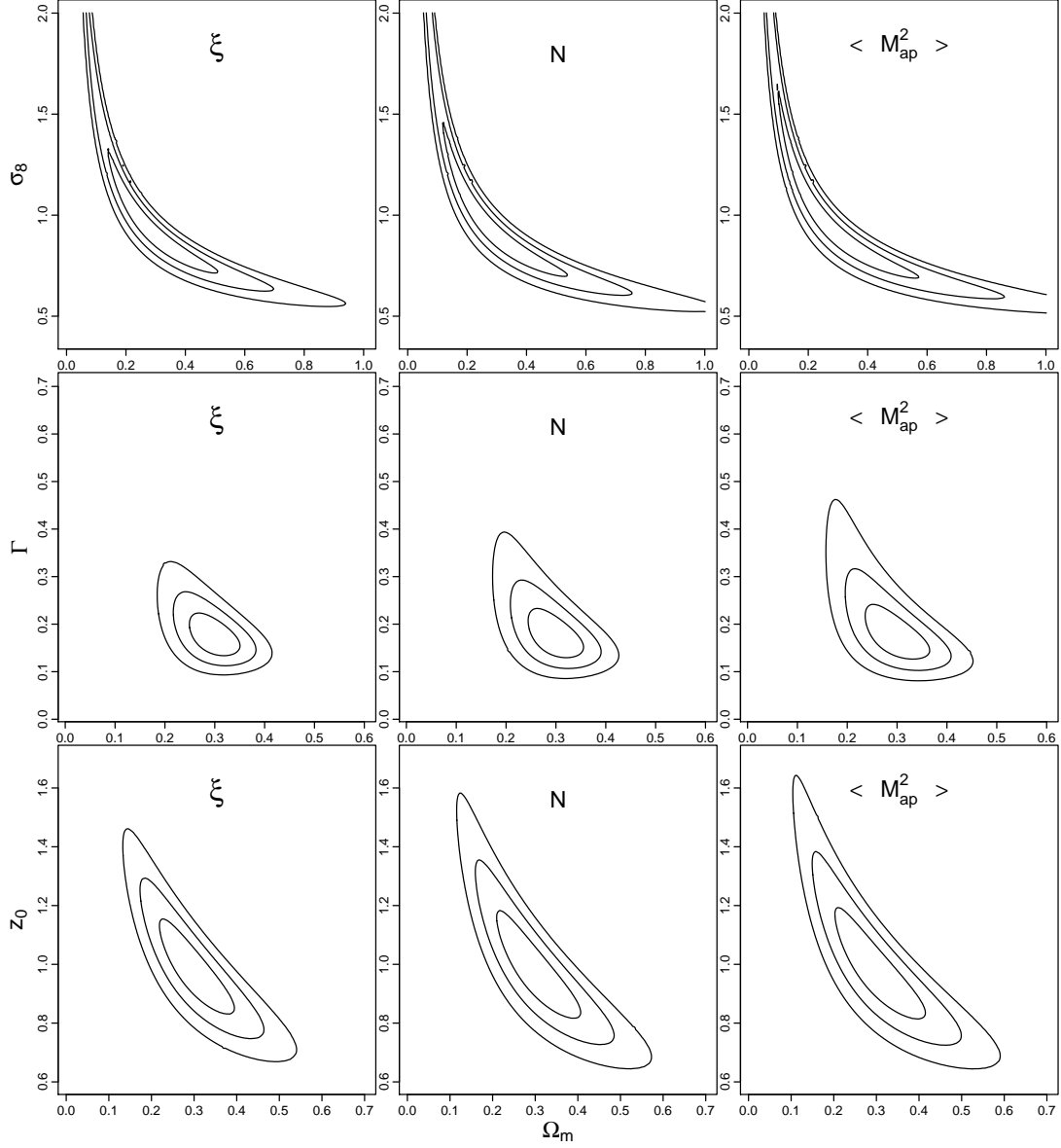
Figure 5.5: *Likelihood contours when varying only two parameters, while the others are fixed to the fiducial values. The contours contain 68.3 %, 95.4 %, 99.73 % of the posterior likelihood. We consider 3 parameter spaces, from top to bottom: σ_8 vs. Ω_m, Γ vs. Ω_m, z_0 vs. Ω_m. The constraints of ξ are shown on the left, N is plotted in the middle, and the results of $\langle M_{ap}^2 \rangle$ are shown on the right.*

Table 5.1: *Values of q for ξ, $\mathcal{N}$, and $\langle M_{ap}^2 \rangle$ considering various parameter spaces.*

Parameter space	$q(\langle M_{ap}^2 \rangle)$	$q(\mathcal{N})$	$q(\xi)$	$\Delta q(\mathcal{N})$	$\Delta q(\xi)$
Γ vs. Ω_m	14.7	11.7	9.1	20.4 %	38.1 %
σ_8 vs. Γ	23.1	19.0	14.6	17.8 %	36.8 %
σ_8 vs. Ω_m	376.3	286.9	207.1	23.7 %	45.0 %
z_0 vs. Ω_m	46.4	41.0	32.9	11.6 %	29.1 %
z_0 vs. σ_8	95.3	91.4	73.2	4.1 %	23.2 %
σ_8 vs. Ω_m (z_0)	416.9	313.4	230.0	25.8 %	44.8 %
σ_8 vs. Ω_m (Γ)	780.5	720.9	527.0	7.6 %	32.5 %
Γ vs. Ω_m (σ_8)	93.7	77.6	61.6	17.2 %	34.3 %
σ_8 vs. Ω_m (Γ, z_0)	983.8	850.6	623.5	13.5 %	36.6 %

Note: Parameters over which we marginalize are mentioned in brackets, the q are given in units of 10^{-4}. The quantities $\Delta\xi$ and $\Delta\mathcal{N}$ are explained in (5.10), (5.11) respectively.

as

$$\Delta\xi = \frac{q(\langle M_{ap}^2 \rangle) - q(\xi)}{q(\langle M_{ap}^2 \rangle)}, \qquad (5.10)$$

$$\Delta\mathcal{N} = \frac{q(\langle M_{ap}^2 \rangle) - q(\mathcal{N})}{q(\langle M_{ap}^2 \rangle)}. \qquad (5.11)$$

The results are summarized in Table 5.1. The parameter combination σ_8 vs. Ω_m shows a relative improvement of $\Delta\mathcal{N} = 26.4\%$, whereas the improvement is much less for the case z_0 vs. σ_8 ($\Delta\mathcal{N} = 4.1\%$). The amount of new information contributed by $\xi_+(\theta_0)$ depends on two main issues. First, ξ_+ integrates over a very broad range of the power spectrum and it can happen that, although P_E is sensitive to the parameters considered, the integral over P_κ is much less. For example, if one varies Γ, the power spectrum is tilted and looks significantly different, whereas the corresponding $\xi_+(\theta_0)$ might be very similar. Second, $\langle M_{ap}^2 \rangle$ does not contain information on small Fourier modes, whereas $\mathcal{N}$ gains information about these modes from the data point $\xi_+(\theta_0)$. However, in case these modes of the power spectrum are not sensitive to the parameters considered, the information that is contributed by $\xi_+(\theta_0)$ is mainly redundant, hence $\Delta\mathcal{N}$ is low. For example, varying σ_8 or Ω_m changes P_E similarly, i.e. increasing Ω_m or σ_8 increases the amplitude of P_E on all Fourier modes. Therefore, the integration over P_E is as sensitive to parameter variations as P_E itself. Furthermore, the deviation of power spectra with different values in σ_8 and Ω_m becomes much more significant for small Fourier modes. Information on these scales is not included in $\langle M_{ap}^2 \rangle$ but is contributed by $\xi_+(\theta_0)$, resulting in a large $\Delta\mathcal{N}(26,4 \%)$. In contrast to this, a variation in z_0 changes the power spectrum very little, especially the dependence is weak on low ℓ-scales. Accordingly, the gain in information for the cases z_0 vs. Ω_m and z_0 vs. σ_8 is rather small.
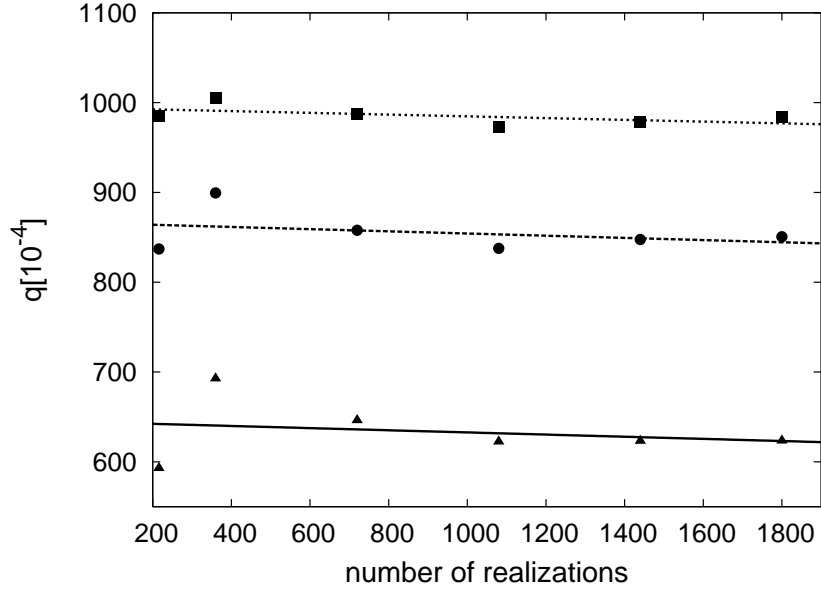
Figure 5.6: *Dependence of q calculated for ξ (triangles), $\mathcal{N}$ (circles), and $\langle M_{\mathrm{ap}}^2 \rangle$ (squares) on the numbers of realizations for the case of a marginalized posterior likelihood. The parameter space is σ_8 vs. Ω_{m} (marginalized over Γ and z_0). The lines are a linear fit through data points (solid for ξ, dashed for $\mathcal{N}$, dotted for $\langle M_{\mathrm{ap}}^2 \rangle$). Similar to Fig. 5.3, the deviation of q belonging to different numbers of realizations is much smaller compared to the deviation of q of different measures.*

5.2.4 Variation of three and four parameters - marginalization

In this section we perform a likelihood analysis in three- and four-dimensional parameter space. To illustrate the results in two-dimensional contour plots, we define the marginalized posterior likelihood

$$P_{\mathrm{mPL}}(\boldsymbol{\pi}_{12}|\boldsymbol{\xi}_{\pm}) = \int d\pi_3 \int d\pi_4 \, P_{\mathrm{PL}}(\boldsymbol{\pi}_{1234}|\boldsymbol{\xi}_{\pm}) \,, \tag{5.12}$$

which is obtained by integrating over the posterior likelihood of the marginalized parameters. The marginalized likelihood is also biased due to its non-linearity in the log-likelihood. To examine whether this bias affects our results significantly, we performed the same experiment as for q in two-dimensional parameter space. We calculated q for our three different measures depending on the number of realizations. The results are shown in Fig. 5.6; again, the bias due to the process of marginalization is small compared to the difference in q of our three data vectors showing that in the marginalized case we can also use q to compare the information content. We also optimize the combined data vector, similar to Sect. 5.2.3 and summarize the results in Table 5.2. For the same reasons as in the previous section, the optimal angular scale of the added ξ_+ data point is again around $10'$, and we choose this optimized $\mathcal{N}$ for the likelihood analysis in three- and four-dimensional parameter space. The results of the likelihood analysis are comparable to those obtained in two-dimensional parameter space. The q (see Table 5.1) are larger and the contours (see Fig. 5.7) broader. Again, the relative improvement $\Delta\mathcal{N}$ depends on the parameter
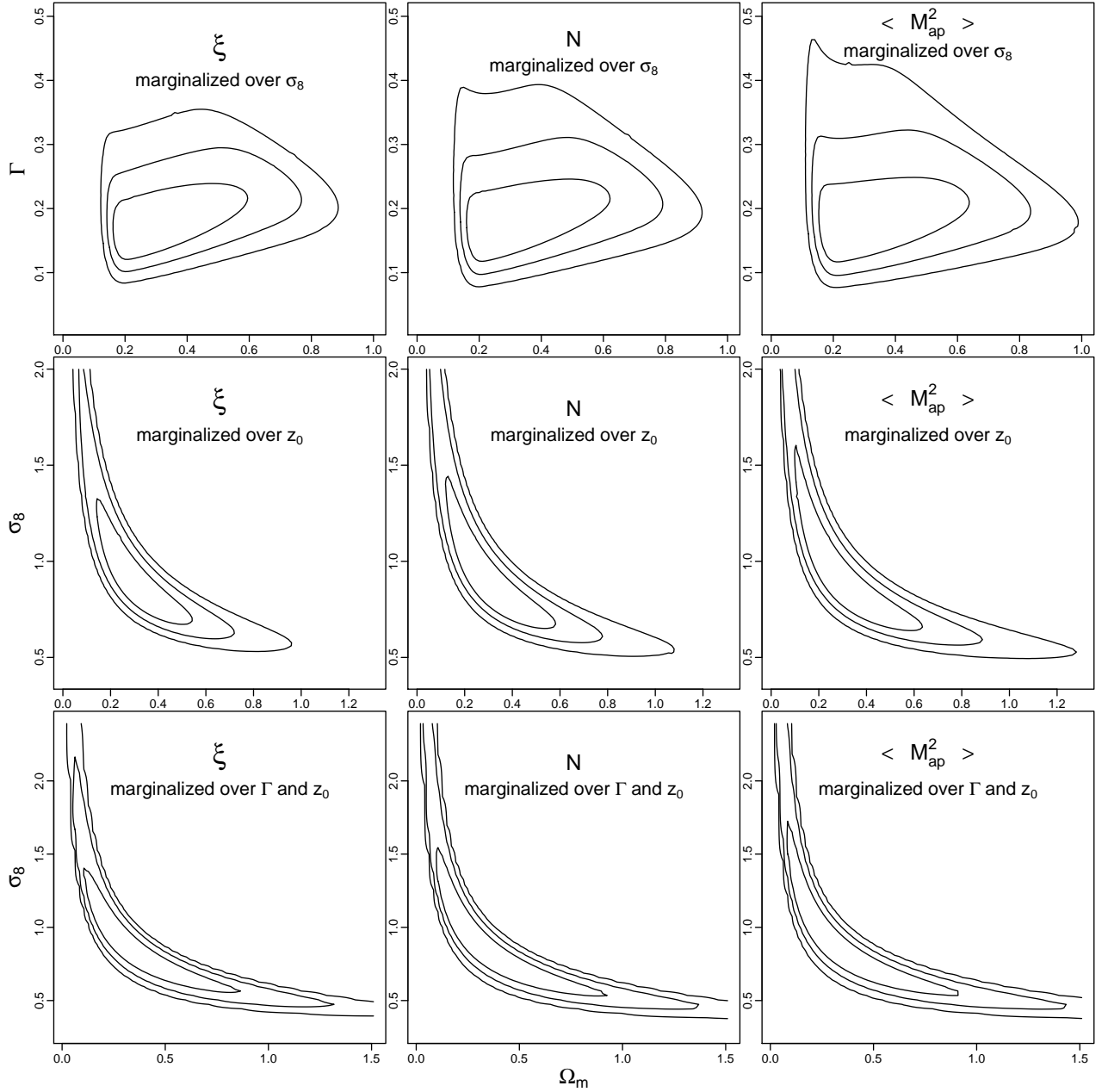
Figure 5.7: *Likelihood contours of ξ, $\mathcal{N}$, and $\langle M_{ap}^2 \rangle$ in three- and four-dimensional parameter space. From top to bottom we see Γ vs. Ω_m marginalized over σ_8, σ_8 vs. Ω_m marginalized over z_0, and σ_8 vs. Ω_m marginalized over Γ and z_0. The contours contain 68.3 %, 95.4 %, 99.73 % of the marginalized posterior likelihood. The small scatter of the contours in the last plot is due to a lower resolution of the grid in four-dimensional parameter space compared to the grids in two- and three-dimensional parameter space.*

Table 5.2: *The optimal angular separation θ_0 of the added ξ_+ data point in $\mathcal{N}$ considering three- and four-dimensional parameter space.*

Parameter space[*]	Optimal value θ_0
Γ vs. Ω_m (σ_8)	$\theta_0 = 9'.1$
σ_8 vs. Ω_m (z_0)	$\theta_0 = 13'.0$
σ_8 vs. Ω_m (Γ and z_0)	$\theta_0 = 12'.0$

[*]We marginalize over the parameters mentioned in parenthesis.

space considered. For σ_8 vs. Ω_m marginalized over z_0, the improvement is very high (25.8 %) but becomes much lower for σ_8 vs. Ω_m marginalized over Γ. This can be explained by looking how P_E changes with respect to the variation in parameter space. For the combination σ_8 vs. Ω_m, we already explained this in Sect. 5.2.3 and the influence of z_0 on P_E is quite similar. Increasing z_0 also increases P_E, although the effect is not very large. Therefore, the improvement of σ_8 vs. Ω_m marginalized over z_0 is comparable to the non-marginalized case. When varying the shape parameter Γ, P_E is tilted and this dependence of P_E on Γ is different compared to the other three parameters. Scales of P_E which are most sensitive to Γ differ from scales sensitive to σ_8, Ω_m and z_0 and the same argument holds for the scales of the added $\xi_+(\theta_0)$. Therefore, the optimal θ_0 for the case σ_8 vs. Ω_m marginalized over Γ is a compromise and the relative improvement is much lower (7.6 %) compared to σ_8 vs. Ω_m marginalized over z_0 (25.8 %).

5.3 Simulation of a B-mode contamination on small angular scales

In this section we simulate a B-mode contamination of ξ, $\mathcal{N}$, and $\langle M_\mathrm{ap}^2 \rangle$ on small angular scales. At present there is no model available to describe B-modes; taking into account that B-modes most likely occur on small angular scales (e.g. Hoekstra et al. 2002b; van Waerbeke et al. 2005; Massey et al. 2007b), we use the following arbitrary model for a B-mode power spectrum

$$P_\mathrm{B}(\ell) = 0.2\, P_\mathrm{E}(\ell)\, \mathrm{e}^{-\ell_\mathrm{B}/\ell}, \tag{5.13}$$

where ℓ_B defines a scale below which the B-mode contamination decreases quickly. In this work we choose $\ell_\mathrm{B} = 1000$; we postpone a detailed study of how other ℓ_B change the results to future work. The B-mode contribution to ξ can be calculated from (3.51) and (3.52) by assuming $P_\mathrm{E} = 0$. To calculate the covariance $\mathbf{C}_\mathrm{B}$, we assume that the probability distribution of B-modes can be described by a Gaussian random field. This assumption enables us to calculate the covariance directly in terms of the power spectrum P_B (Joachimi et al. 2008). The covariance of the 2PCF
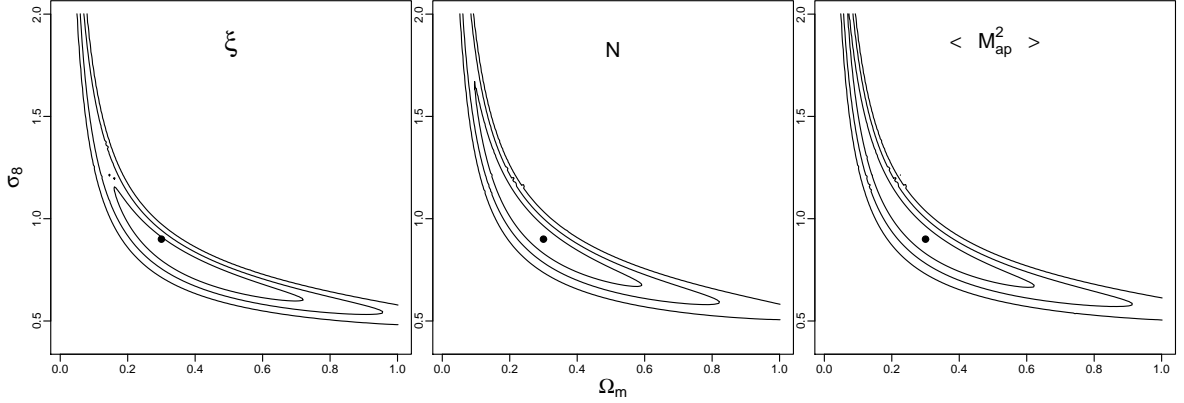
Figure 5.8: *The likelihood contours for the case that the shear signal is contaminated with B-modes. We only consider a two-dimensional parameter space (σ_8 vs. Ω_m) and the contours again contain 68.3 %, 95.4 %, 99.73 % of the posterior likelihood. The black dot in each plot indicates the fiducial model.*

corresponding to the B-mode contribution is given by

$$
C_{B,ij}^{++} = \frac{1}{A\pi} \int d\ell\ell J_0\left(\ell\vartheta_i\right) J_0\left(\ell\vartheta_j\right)\left(P_B^2(\ell) + P_B(\ell)\frac{\sigma_\epsilon^2}{\bar{n}}\right),
$$

$$
C_{B,ij}^{--} = \frac{1}{A\pi} \int d\ell\ell J_4\left(\ell\vartheta_i\right) J_4\left(\ell\vartheta_j\right)\left(P_B^2(\ell) + P_B(\ell)\frac{\sigma_\epsilon^2}{\bar{n}}\right),
$$

$$
C_{B,ij}^{+-} = -\frac{1}{A\pi} \int d\ell\ell J_0\left(\ell\vartheta_i\right) J_4\left(\ell\vartheta_j\right)\left(P_B^2(\ell) + P_B(\ell)\frac{\sigma_\epsilon^2}{\bar{n}}\right),
$$

where A defines the area of the survey, σ_ϵ the intrinsic ellipticity noise, and n the number density of the source galaxies. According to the corresponding values of the ray-tracing simulations we choose $\sigma_\epsilon = 0.3$ and $\bar{n} = 25/\text{arcmin}^2$. Note that $C_{B,ij}^{-+} = C_{B,ji}^{+-}$. The pure shot noise term of $C_B^{\pm\pm}$ is contained in $C_E^{\pm\pm}$, but in case of C_B^{+-} this term vanishes anyway. We further assume that the contamination is independent of the lensing signal, meaning there is no correlation between E- and B-modes. This assumption does not hold in case the B-mode signal is caused by insufficient PSF-correction or other systematics, which we will comment on at the end of this section. For the case that B-modes are created independently from E-modes, we can define a combined E/B-mode covariance matrix as

$$
\mathbf{C}_{tot} = \mathbf{C}_E + \mathbf{C}_B . \tag{5.14}
$$

Having obtained $\mathbf{C}_\xi$ as described above, we therefrom calculate $\mathbf{C}_M$ and $\mathbf{C}_N$ and perform a likelihood analysis similar to Sects. 5.2.3 and 5.2.4. We only show the results for the Ω_m vs. σ_8 plane, which are illustrated in Fig. 5.8 and Table 5.3. As expected, $\langle M_{ap}^2 \rangle$ is not affected by the contamination at all. According to (3.57), there is no contribution of P_B to the individual data points of $\langle M_{ap}^2 \rangle$, and the same holds for $\mathbf{C}_M$. Therefore, it is no surprise that q and contour plots of $\langle M_{ap}^2 \rangle$ are similar to those in Sect. 5.2.3. In contrast to this, the 2PCF data vector is

Table 5.3: *The results of the likelihood analysis in case the data vectors are contaminated by B-modes. The fiducial model corresponds to $\Omega_m = 0.3$ and $\sigma_8 = 0.9$.*

Data vector	q (with B-modes)	q (without B-modes)	Δq^*	Best-fit parameter set
$\langle M_{ap}^2 \rangle$	376.3	376.3	0.0	$\sigma_8 = 0.90$, $\Omega_m = 0.30$
$\mathcal{N}$	314.4	286.9	9.6 %	$\sigma_8 = 0.90$, $\Omega_m = 0.30$
ξ	275.0	207.1	32.8 %	$\sigma_8 = 0.76$, $\Omega_m = 0.39$

    * The relative difference of the q with and without B-mode contamination.

strongly affected by the contamination leading to a q that is 32.8 % higher compared to the case when only E-modes are present. Furthermore, there is a significant deviation between the best-fit parameter set ($\sigma_8 = 0.76$, $\Omega_m = 0.39$) and the fiducial cosmological model ($\sigma_8 = 0.90$ and $\Omega_m = 0.30$). Compared to ξ the combined data vector $\mathcal{N}$ is much less contaminated (9.6 %), and its best-fit parameter set still matches the fiducial model exactly. Considering the q, one might argue that ξ still gives tighter constraints on the parameters, but this result is biased in favor of ξ due to considering a parameter space only up to $\Omega_m = 1.0$. Both measures have different best-fit parameter sets and the likelihood contours are cut off at the limits of the considered parameter space. In such a case, an extension of the parameter space might change the result of comparing the q qualitatively. Compared to $\langle M_{ap}^2 \rangle$, the information content of $\mathcal{N}$ is still significantly higher, although the relative improvement decreases to $\Delta \mathcal{N} = 16.4\%$, whereas we obtained $\Delta \mathcal{N} = 23.7\%$ when only E-modes were present (see Sect. 5.2.3). This decrease, due to the contaminated $\xi_+(\theta_0)$ data point in $\mathcal{N}$, is another reason not to include more data points of ξ_+. In the presence of B-modes, additional 2PCF data points would bias the parameter constraints and weaken the results even more, leading to similar deficits to those obtained from the 2PCF data vector itself. As already mentioned above, the assumption of B-modes being independent of the E-mode signal does not always hold. In case the contamination affects both, E-mode and B-mode signal, the impact on the parameter constraints of the different measures is hard to quantify. When one measures a B-modes signal, it is a common approach to assume that the E-mode signal is contaminated in a similar way, hence one correspondingly increases its error bars. Although this assumption is sensible, there are possible scenarios where the amount of contamination in E- and B-mode differs and the E-mode contamination cannot be quantified at all. Under the assumption that B-modes trace the scales of the E-mode contamination, it is reasonable to exclude those scales from the likelihood analysis. This can be done using $\langle M_{ap}^2 \rangle$ or $\mathcal{N}$, but ξ cannot avoid the contamination due to its broad filter functions.

5.4 Conclusions and summary of the comparison of cosmic shear data vectors

Although the 2PCF and the aperture mass dispersion are both filtered versions of the power spectrum, the first contains more information on P_E than the latter. The reason for this is that ξ sam-

ples the power spectrum over a much broader range and also collects information on scales that are larger than the size of the survey. The data vector $\langle M_{ap}^2 \rangle$ lacks this large-scale information, but yields highly localized information on P_E. Nevertheless, $\langle M_{ap}^2 \rangle$ has other advantages. First, due to its narrow filter function, the data points are much less correlated compared to the 2PCF data points. This leads to a mainly diagonal covariance matrix, which is numerically stabler during the inversion process in a likelihood analysis. Second, when considering higher-order statistics, $\langle M_{ap}^3 \rangle$ is much easier to handle than the three-point correlation function (Schneider et al. 2005), and third, the aperture mass dispersion is only sensitive to E-modes. Based on these considerations we create the combined data vector $\mathcal{N}$, which preserves the advantages of $\langle M_{ap}^2 \rangle$ and additionally provides large-scale information on P_E. This data vector can be optimized with respect to the angular scale of the added data point $\xi_+(\theta_0)$, but this optimization very likely depends on the survey geometry and must be performed for each survey separately. We compared the three data vectors in a detailed likelihood analysis and find that the combined data vector is a strong improvement over $\langle M_{ap}^2 \rangle$ in information content. However, the amount of improvement depends on the parameter space considered, more precisely, on the dependence of P_E on variations in those parameters. The combined data vector $\mathcal{N}$ also maintains the other advantages of the aperture mass dispersion. Its covariance matrix is almost diagonal, and even the cross terms $C(\mathcal{M}(\theta_k), \hat{\xi}_+(\theta_0))$ are much smaller compared with the off-diagonal terms of $\mathbf{C}_\xi$. Comparing the information content of ξ and $\mathcal{N}$, ξ gives tighter constraints if the shear signal only consists of E-modes. In the more realistic case, when B-modes are also present, the parameter constraints of ξ are significantly weakened and, even worse, biased. The data vector $\mathcal{N}$ is much less affected by the contamination and still gives tighter constraints on cosmological parameters than does $\langle M_{ap}^2 \rangle$.

Chapter 6

Ring statistics

In chapter 3 we outline the concept of E- and B-modes, where only the former are created through gravitational lensing and the latter indicate remaining systematics in the shear signal. Therefore, decomposing the shear field into E- and B-modes is an important check for systematics in cosmic shear. The most commonly used methods for E- and B-mode decomposition require the 2PCF to be known down to arbitrary small or up to arbitrary large angular separations. This is not possible in practice, as a consequence the corresponding methods do not separate E- and B-modes properly on small angular scales. This issue is further outlined in Kilbinger et al. (2006). The ring statistics (Schneider & Kilbinger 2007) provides a new method to perform an E-/B-mode decomposition using a 2PCF measured on a finite interval. In this chapter we improve the ring statistics with respect to its signal and examine its potential to constrain cosmological parameters. In addition, we measure the shear signal using the ring statistics with data from the Canada-France-Hawaii Telescope Legacy Survey (CFHTLS). This marks the first measured shear signal obtained from the ring statistics. From this measurement we derive constraints on σ_8, namely $\sigma_8(\Omega_m/0.25)^{0.49} = 0.82^{+0.02}_{-0.04}$.

6.1 Methods to decompose E-modes and B-modes

Schneider et al. (2002b) outline the following concepts to decompose E- and B-modes:

- E/B mode power spectra

$$P_{E/B}(\ell) = \pi \int_0^\infty d\vartheta \, \vartheta \left[\xi_+(\vartheta) J_0(\ell\vartheta) \pm \xi_-(\vartheta) J_4(\ell\vartheta) \right] \tag{6.1}$$

- E/B mode two-point correlation function (in the "$\pm$" notation below the "+" corresponds to ξ_E, the "-" corresponds to ξ_B.)

$$\xi_{E/B,+}(\theta) = \frac{1}{2}\left[\xi_+(\theta) \pm \xi_-(\theta) \pm \int_\theta^\infty \frac{d\vartheta}{\vartheta} \xi_-(\vartheta)\left(4 - 12\frac{\theta^2}{\vartheta^2}\right) \right] \tag{6.2}$$

$$\xi_{E/B,-}(\theta) = \frac{1}{2}\left[\xi_+(\theta) \pm \xi_-(\theta) + \int_0^\theta \frac{d\vartheta \, \vartheta}{\theta^2} \xi_+(\vartheta)\left(4 - 12\frac{\vartheta^2}{\theta^2}\right) \right] \tag{6.3}$$

Figure 6.1: *The aperture mass dispersion calculated from the power spectrum, compared to the case when $\langle M_{ap}^2 \rangle$ is calculated from the 2PCF with $\vartheta_{min} = 0.01$. The vertical dotted lines mark 3 different θ_i, i.e. 0.1, 0.2, and 0.8 which are examined more closely, with respect to the filter functions $T_{\pm}$.*
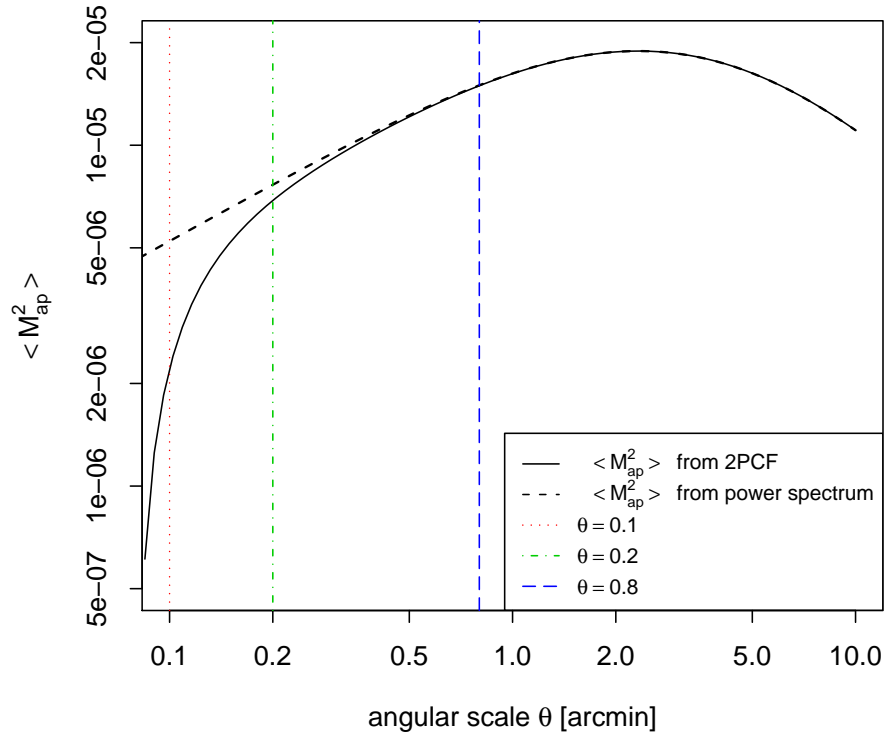
- Aperture mass dispersion

$$\langle M_{ap/\perp}^2 \rangle(\theta) = \frac{1}{2} \int_0^{2\theta} \frac{d\vartheta\, \vartheta}{\theta^2} \left[\xi_+(\vartheta) T_+\left(\frac{\vartheta}{\theta}\right) \pm \xi_-(\vartheta) T_-\left(\frac{\vartheta}{\theta}\right) \right] \tag{6.4}$$

The above quantities can be calculated from the 2PCF, however each method requires that the 2PCF is either measured down to arbitrary small or large angular separation, respectively. For the case of the aperture mass dispersion we already mentioned this problem in Sect. 5.2.1.1; we will now examine it in more detail. Figure 6.1 is similar to Fig. 5.1, but restricts the range of the aperture mass dispersion to angular scales of $\theta \in [0.1; 10.0]$. We compare $\langle M_{ap}^2 \rangle$ calculated from the power spectrum via (3.57) and $\langle M_{ap}^2 \rangle$ calculated from a set of 2PCFs with $\vartheta_{min} = 0.01$. The calculation of $\langle M_{ap}^2 \rangle$ from the 2PCF fails on small angular scales; this behavior can be explained when looking at the filter function $T_{\pm}$ (3.64), (3.65). We mark three different scales θ_i in Fig. 6.1, i.e. 0.1, 0.2, and 0.8. The argument of the T-functions is defined as $x = \vartheta/\theta$, which implies that when calculating $\langle M_{ap}^2 \rangle$ at θ_i, the T-functions are cut off at $x_{min} = \vartheta_{min}/\theta_i$. The corresponding x_{min} are shown in Fig. 6.2; we see that in case of T_- the cut-off has almost no impact on the value of $\langle M_{ap}^2 \rangle$ as the filter function approaches zero for small x. However, in case of T_+ this cut-off is important, as T_+ has its maximum amplitude at $x = 0$ and, in addition, the amplitude of ξ_+ increases when approaching ϑ_{min}. When trying to calculate $\langle M_{ap}^2 \rangle$ from the 2PCF on angular scales smaller than 0.05, its values will even drop below zero. For the three scales we consider here the relative difference between $\langle M_{ap}^2 \rangle$ calculated from the power spectrum and $\langle M_{ap}^2 \rangle$ calculated from the 2PCF is 54.6% for $\theta = 0.1$, 9.5% for $\theta = 0.5$, and less than 1% for $\theta = 0.8$. The considered
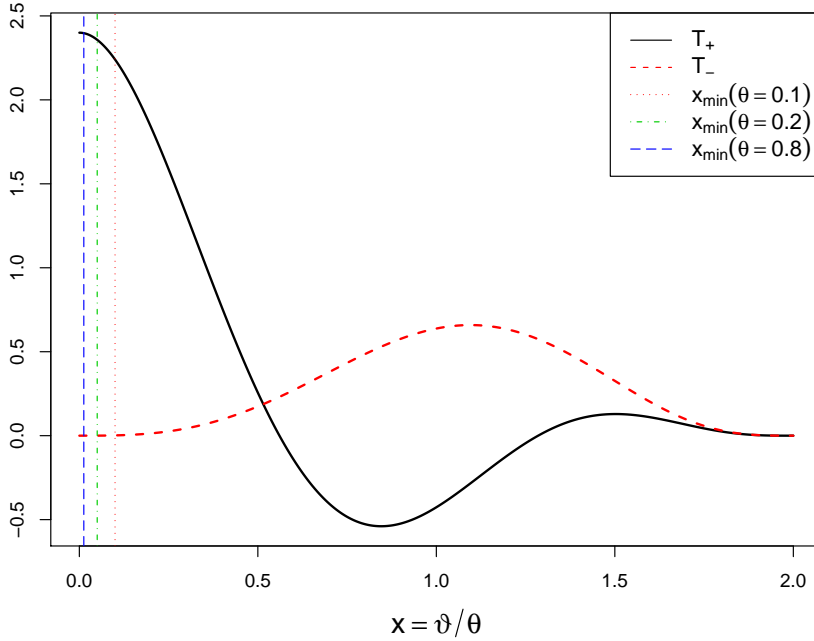
Figure 6.2: *The filter functions T_+ and T_- used in the calculation of $\langle M_{\mathrm{ap}}^2 \rangle$ via the 2PCF; x is defined as $x = \vartheta/\theta$. The vertical lines mark the x_{min} which correspond to the marked θ in Fig. 6.1; a further explanation can be found in the text.*

value $\vartheta_{\mathrm{min}} = 0\overset{\prime\prime}{.}01$ is optimistic; note that a higher ϑ_{min} increases the inaccuracy when calculating $\langle M_{\mathrm{ap}}^2 \rangle$ from the 2PCF on small scales.

Most cosmic shear analyses use the aperture mass dispersion and the E/B-mode 2PCF to decompose E-and B-modes. For example, Massey et al. (2007b) and Fu et al. (2008) account for the unknown 2PCF in the corresponding integrals by simulating 2PCFs using a theoretical model for P_κ. This ansatz is problematic, since one explicitly assumes that scales larger than the size of the survey are free from B-modes. In addition, the assumed cosmology in the theoretical power spectrum can bias the results. In the next sections we describe the ring statistics which provides a method to separate E- and B-modes using a 2PCF on a finite interval.

6.2 From circle to ring statistics

In this section we briefly summarize the basic definitions of circle and ring statistics. For more details on this topic the reader is referred to Schneider & Kilbinger (2007).

Given a circle of radius θ around the origin, we define the mean tangential and cross component of the shear on this circle as

$$C(\theta) = C_{\mathrm{t}}(\theta) + iC_\times(\theta) = \frac{1}{2\pi} \int_0^{2\pi} \mathrm{d}\varphi \, (\gamma_{\mathrm{t}} + i\gamma_\times)(\theta, \varphi) \,. \tag{6.5}$$

According to Crittenden et al. (2002) and Schneider et al. (2002b), C_{t} is sensitive only to E-modes, whereas $C_\times$ only measures B-modes. The correlators of C_{t} and $C_\times$ can be expressed as an integral over φ. To derive an easy expression through the 2PCF, one transforms the integration

variable φ into ϑ and obtains

$$\langle C_t(\theta_1)C_t(\theta_2)\rangle = \int_{\theta_2-\theta_1}^{\theta_1+\theta_2} \frac{\mathrm{d}\vartheta}{2\vartheta} \left[\xi_+(\vartheta)Y_+\left(\frac{\vartheta}{\theta_2},\frac{\theta_1}{\theta_2}\right) + \xi_-(\vartheta)Y_-\left(\frac{\vartheta}{\theta_2},\frac{\theta_1}{\theta_2}\right) \right] \tag{6.6}$$

$$\langle C_\times(\theta_1)C_\times(\theta_2)\rangle = \int_{\theta_2-\theta_1}^{\theta_1+\theta_2} \frac{\mathrm{d}\vartheta}{2\vartheta} \left[\xi_+(\vartheta)Y_+\left(\frac{\vartheta}{\theta_2},\frac{\theta_1}{\theta_2}\right) - \xi_-(\vartheta)Y_-\left(\frac{\vartheta}{\theta_2},\frac{\theta_1}{\theta_2}\right) \right], \tag{6.7}$$

where the functions Y_+ and Y_- are given by

$$Y_+(x,\eta) = \frac{x^2 \left[(1-x^2)^2 - 2\eta^2 x^2 + \eta^4 \right]}{\eta^2 \pi \sqrt{(1+\eta)^2 - x^2} \sqrt{x^2 - (1-\eta^2)}} \tag{6.8}$$

$$Y_-(x,\eta) = \frac{(1-x^2)^2 - 2\eta^2(2-x^2) + \eta^4(6+2x^2+x^4) - 2\eta^6(2+x^2) + \eta^8}{x^2\eta^2\pi \sqrt{(1+\eta)^2 - x^2} \sqrt{x^2 - (1-\eta^2)}}. \tag{6.9}$$

The circle statistics already offers a method for an E-and B-mode decomposition using 2PCFs which are measured over a finite range. Unfortunately, the filter functions Y_+, Y_- suffer from two singularities at the integration boundaries in (6.6) and (6.7), which originate from the variable transformation of φ into ϑ. This property of the $Y_\pm$ functions implies that 2PCFs with the integration boundaries as arguments dominate the circle statistics' estimator. In particular, the noise of these boundary 2PCFs strongly affects the estimator, making the circle statistics inapplicable as a cosmic shear measure.

The ring statistics improves on this issue; when expressing it through the 2PCF, its filter functions do not contain any singularities. This new cosmic shear measure can be expressed as an integral over the circle statistics, which means that instead of measuring the shear on a circle we now consider the shear inside an annulus with $\zeta_1 \leq \theta \leq \zeta_2$.

$$\mathcal{R} = \mathcal{R}_t + i\mathcal{R}_\times = \int_{\zeta_1}^{\zeta_2} \mathrm{d}\theta\, W(\theta)\, C(\theta), \tag{6.10}$$

where $W(\theta)$ is a weight function which fulfills

$$\int_{\zeta_1}^{\zeta_2} \mathrm{d}\theta\, W(\theta) = 1. \tag{6.11}$$

We want to calculate the correlator of two rings with the annuli $\zeta_1 \leq \theta_1 \leq \zeta_2$ for the first ring and $\zeta_3 \leq \theta_2 \leq \zeta_4$ for the second (see Fig. 6.3). The rings are non-overlapping, more precisely we require $\zeta_i < \zeta_j$ if $i < j$. Outer and inner ring are separated by $\vartheta_{\min} = \zeta_3 - \zeta_2$ which guarantees that the ring statistics correlator can be calculated properly from 2PCFs covering the range $[\vartheta_{\min}; \Psi = \zeta_2 + \zeta_4]$. The correlator is defined as

$$\begin{aligned} \langle \mathcal{R}\mathcal{R} \rangle &= \int_{\zeta_1}^{\zeta_2} \mathrm{d}\theta_1\, W_1(\theta_1) \int_{\zeta_3}^{\zeta_4} \mathrm{d}\theta_2\, W_2(\theta_2)\, \langle C(\theta_1)C(\theta_2)\rangle \\ &= \int_{\zeta_1}^{\zeta_2} \mathrm{d}\theta_1\, W_1(\theta_1) \int_{\zeta_3}^{\zeta_4} \mathrm{d}\theta_2\, W_2(\theta_2) \times \int_{\theta_2-\theta_1}^{\theta_2+\theta_1} \frac{\mathrm{d}\vartheta}{\vartheta}\, \xi_-(\vartheta)\, Y_-\left(\frac{\vartheta}{\theta_2},\frac{\theta_1}{\theta_2}\right). \end{aligned} \tag{6.12}$$
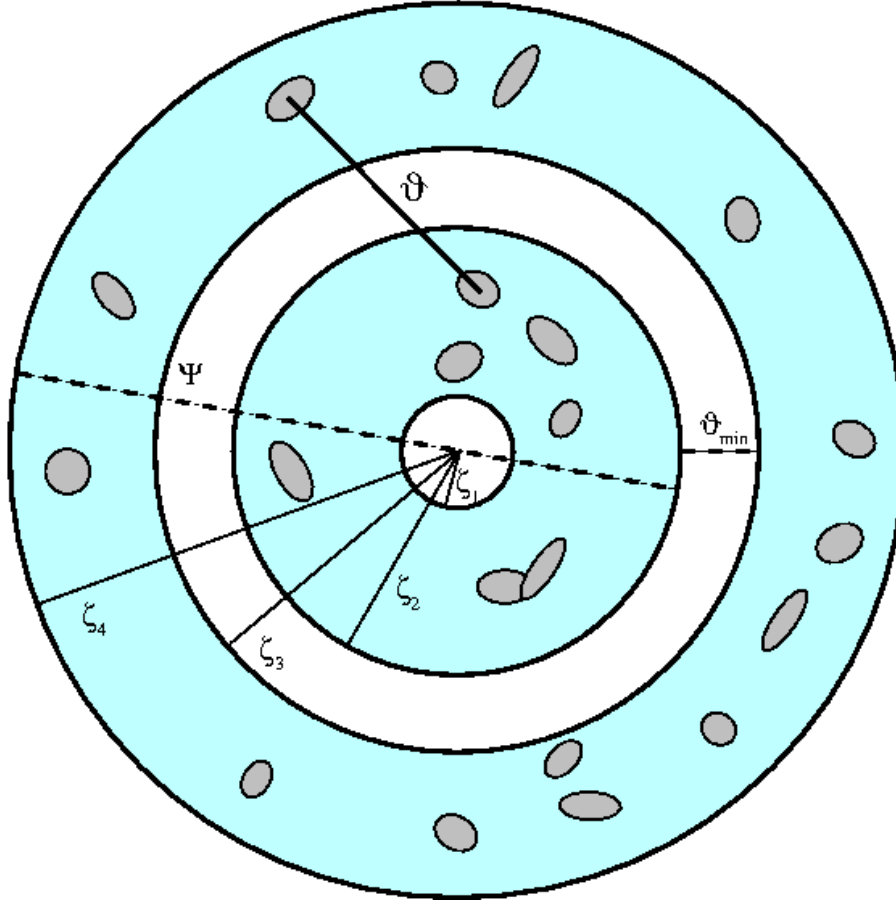
Figure 6.3: *This figure illustrates the principle of the ring statistics. We calculate the 2PCF for each galaxy in the inner ring with all galaxies in the outer ring, which implies that the angular separation of the 2PCF extends over a range of $\vartheta \in [\vartheta_{\min}; \Psi]$. The ring statistics is then calculated as an integral over the 2PCF with the filter functions $Z_\pm$.*

Changing the order of integration twice, first between θ_2 and ϑ, then between θ_1 and ϑ, one obtains

$$\langle \mathcal{RR} \rangle = \int_{\zeta_3 - \zeta_2}^{\zeta_2 + \zeta_4} \frac{d\vartheta}{\vartheta} \xi_-(\vartheta) \underbrace{\int_{\max(\zeta_1, \vartheta - \zeta_4, \zeta_3 - \vartheta)}^{\zeta_2} d\theta_1 \, W_1(\theta_1) \int_{\max(\zeta_3, \vartheta - \theta_1)}^{\min(\zeta_4, \vartheta + \theta_1)} d\theta_2 \, W_2(\theta_2) \, Y_- \left(\frac{\vartheta}{\theta_2}, \frac{\theta_1}{\theta_2} \right)}_{Z_-}. \quad (6.13)$$

The last two integrals define the Z_--function, i.e. the equivalent for the T_- function of the aperture mass dispersion (see 3.65). Similar to Schneider & Kilbinger (2007) we consider the special case

$$\zeta_1 = \frac{1 - \eta}{8} \Psi, \quad (6.14)$$

$$\zeta_2 = \frac{3(1 - \eta)}{8} \Psi, \quad (6.15)$$

$$\zeta_3 = \frac{5\eta + 3}{8} \Psi, \quad (6.16)$$

$$\zeta_4 = \frac{3\eta + 5}{8} \Psi, \quad (6.17)$$

where η is the ratio of minimum to maximum separation $\eta = \vartheta_{\min}/\Psi$. In addition one has to specify the form for the weight functions W in (6.13). We follow the choice of Schneider &
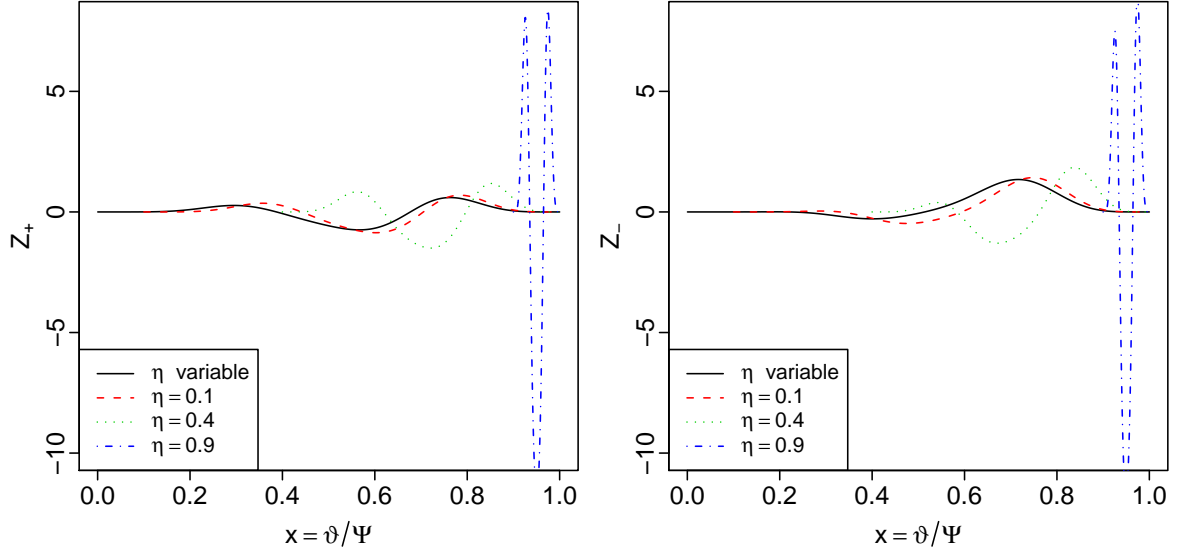
Figure 6.4: *This plot shows the filter functions Z_+ (left panel) and Z_- (right panel) depending on $x = \vartheta/\Psi$ for 4 different choices of η.*

Kilbinger (2007)

$$W_1(\theta_1) = \frac{30(\theta_1 - \zeta_1)^2 (\zeta_2 - \theta_1)^2}{(\zeta_2 - \zeta_1)^5}, \qquad (6.18)$$

$$W_2(\theta_2) = \frac{30(\theta_2 - \zeta_3)^2 (\zeta_4 - \theta_2)^2}{(\zeta_4 - \zeta_3)^5}. \qquad (6.19)$$

We examine the ring statistics as a function of Ψ; with the above specifications (and for a given Ψ) the Z_--function only depend on ϑ and η

$$\langle \mathcal{R}\mathcal{R} \rangle (\Psi) = \int_{\vartheta_{\min}}^{\Psi} \frac{\mathrm{d}\vartheta}{\vartheta} \, \xi_-(\vartheta) \, Z_-(\vartheta, \eta), \qquad (6.20)$$

and correspondingly

$$\langle \mathcal{R}\mathcal{R}^* \rangle (\Psi) = \int_{\vartheta_{\min}}^{\Psi} \frac{\mathrm{d}\vartheta}{\vartheta} \, \xi_+(\vartheta) \, Z_+(\vartheta, \eta), \qquad (6.21)$$

where the filter function Z_+ is defined similar to Z_- (6.13) but replacing Y_- by Y_+. The functions are illustrated in Fig. 6.4. For η we examine two different cases, it can be either fixed or kept variable. The first case means that $\vartheta_{\min}$ in the integral in (6.20) and (6.21) changes with Ψ, whereas a variable η allows for $\vartheta_{\min}$ to remain fixed. The latter approach includes more and smaller ϑ-bins into $\langle \mathcal{R}\mathcal{R} \rangle$. We outline advantages and disadvantages of these methods in more detail in Sect. 6.3.

Using (6.10) we see that $\langle \mathcal{R}\mathcal{R} \rangle$ and $\langle \mathcal{R}\mathcal{R}^* \rangle$ read

$$\langle \mathcal{R}\mathcal{R} \rangle = \langle (\mathcal{R}_t + i\mathcal{R}_\times)(\mathcal{R}_t + i\mathcal{R}_\times) \rangle = \langle \mathcal{R}_t\mathcal{R}_t - \mathcal{R}_\times\mathcal{R}_\times \rangle \tag{6.22}$$

$$\langle \mathcal{R}\mathcal{R}^* \rangle = \langle (\mathcal{R}_t + i\mathcal{R}_\times)(\mathcal{R}_t - i\mathcal{R}_\times) \rangle = \langle \mathcal{R}_t\mathcal{R}_t + \mathcal{R}_\times\mathcal{R}_\times \rangle \,, \tag{6.23}$$

$$\tag{6.24}$$

where we used the fact that the imaginary terms vanish for a parity invariant shear field. Recall that $\mathcal{R}_t$ is due to E-modes and $\mathcal{R}_\times$ due to B-modes only. Then the E and B-mode decomposition for the ring statistics reads

$$\langle \mathcal{R}\mathcal{R}_E \rangle (\Psi) = \frac{1}{2} \left(\langle \mathcal{R}\mathcal{R}^* \rangle + \langle \mathcal{R}\mathcal{R} \rangle \right) \tag{6.25}$$

$$= \frac{1}{2} \int_{\vartheta_{\min}}^{\Psi} \frac{\mathrm{d}\vartheta}{\vartheta} \left[\xi_+(\vartheta) \, Z_+(\vartheta, \eta) + \xi_-(\vartheta) \, Z_-(\vartheta, \eta) \right] \tag{6.26}$$

$$\langle \mathcal{R}\mathcal{R}_B \rangle (\Psi) = \frac{1}{2} \left(\langle \mathcal{R}\mathcal{R}^* \rangle - \langle \mathcal{R}\mathcal{R} \rangle \right) \tag{6.27}$$

$$= \frac{1}{2} \int_{\vartheta_{\min}}^{\Psi} \frac{\mathrm{d}\vartheta}{\vartheta} \left[\xi_+(\vartheta) \, Z_+(\vartheta, \eta) - \xi_-(\vartheta) \, Z_-(\vartheta, \eta) \right] \,. \tag{6.28}$$

Similar to the case of the aperture mass dispersion, $\langle \mathcal{R}\mathcal{R}_E \rangle$ and $\langle \mathcal{R}\mathcal{R}_B \rangle$ can be related to the power spectra in case it only consists of E-modes. We therefore insert the relation of $\xi_\pm$ to P_κ (3.51), (3.52) respectively, into (6.26) and obtain

$$\langle \mathcal{R}\mathcal{R}_E \rangle (\Psi) = \int_0^\infty \frac{\mathrm{d}\ell \, \ell}{2\pi} \, P_E \, \mathcal{W}_E(\ell\Psi, \eta) \,, \tag{6.29}$$

with

$$\mathcal{W}_E(\ell\Psi, \eta) = \int_{\vartheta_{\min}}^{\Psi} \frac{\mathrm{d}\vartheta}{2\vartheta} \left[J_0(\ell\vartheta) \, Z_+(\vartheta, \eta) + J_4(\ell\vartheta) \, Z_-(\vartheta, \eta) \right] \,. \tag{6.30}$$

6.3 Optimization of the ring statistics' signal

The signal strength of the ring statistics increases with the angular range covered by the 2PCFs which enter in (6.26). From Fig. 6.3 it becomes clear that the ring statistics' signal for given Ψ will be large if the area between the rings is small, hence $\vartheta_{\min}$ is small. In Schneider & Kilbinger (2007) the authors hold $\eta = \vartheta_{\min}/\Psi$ fixed; in order to obtain a high signal this implies that η must be chosen as small as possible.

The minimum value of $\vartheta_{\min}$ is restricted to the smallest possible separation in the 2PCF which we denote as $\vartheta_{\xi_{\min}}$. A fixed η therefore limits a data vector of the ring statistics to $\Psi \geq \vartheta_{\xi_{\min}}/\eta$. Choosing a small η (in order to increase the signal strength) implies that the ring statistics is restricted to larger scales compared to the case of a large η. This tradeoff between signal strength and small-scale sensitivity can be overcome when relaxing the condition of a fixed η. Instead we suggest a variable η and fix $\vartheta_{\min} = \vartheta_{\xi_{\min}}$ for all Ψ. We assume that this choice increases the signal
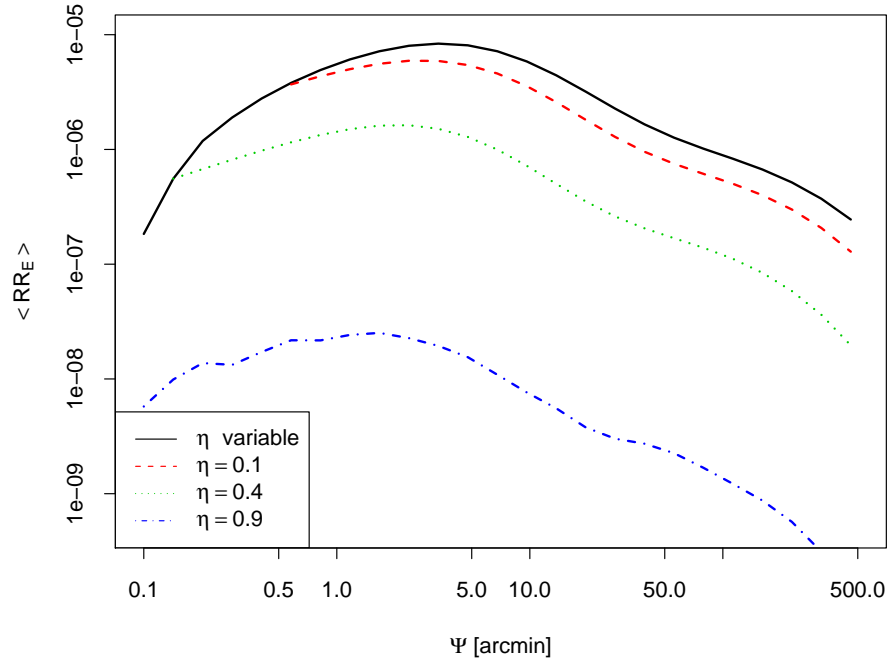
Figure 6.5: *This figure shows the signal of the ring statistics calculated from a set of theoretical 2PCFs with $\vartheta \in [0.05; 470]$ arcmin. The individual lines correspond to 4 different cases of η; from top to bottom we see $\eta = 0\overset{''}{.}05/\Psi$, $\eta = 0.1$, $\eta = 0.4$, $\eta = 0.9$.*

strength of the ring statistics; in addition, it does not restrict the range of Ψ. The only deficit of this method is that the filter function $Z_{\pm}$ has to be recalculated every time the argument of the ring statistics (Ψ) changes. Figure 6.5 compares the signal of the ring statistics for several choices of η to the case of a variable η, which from now on we denote as η_{var}. Here, the ring statistics is calculated from a set of theoretical 2PCFs with the same angular range as the 2PCFs from the CFHTLS which we use in Sect. 6.6, i.e. $\vartheta \in [0\overset{''}{.}05; 470\overset{''}{.}0]$. We clearly see the anticipated behavior, namely that the ring statistics with η_{var} gives a stronger signal, which becomes most obvious on large scales. In addition, it can be measured down to arbitrary small values of Ψ (above $\vartheta_{\xi_{\mathrm{min}}}$), which is not possible when choosing a fixed η.

When comparing the ring statistics to other 2-point statistics of cosmic shear (Fig. 6.6), such as the shear dispersion, $\xi_{\pm}$ and the aperture mass dispersion, we find that the ring statistics' signal is lower. Even the improved version of the ring statistics is on average by a factor of ≈ 2 smaller than the aperture mass dispersion. This low signal can be explained when comparing the filter functions of $\langle \mathcal{RR}_{\mathrm{E}} \rangle$ and $\langle M_{\mathrm{ap}}^2 \rangle$, $Z_{\pm}$ (Fig. 6.4) and $T_{\pm}$ (Fig. 6.2), respectively. The Z-functions have two roots at their boundaries whereas the T_{+}-function becomes particularly large for small x. However, more important than the signal strength is the signal-to-noise ratio; the latter quantifies the ability of a measure to constrain cosmological parameters. In the next section we derive an expression for the covariance of the ring statistics and employ this expression to compare the information content of ring statistics and aperture mass dispersion.
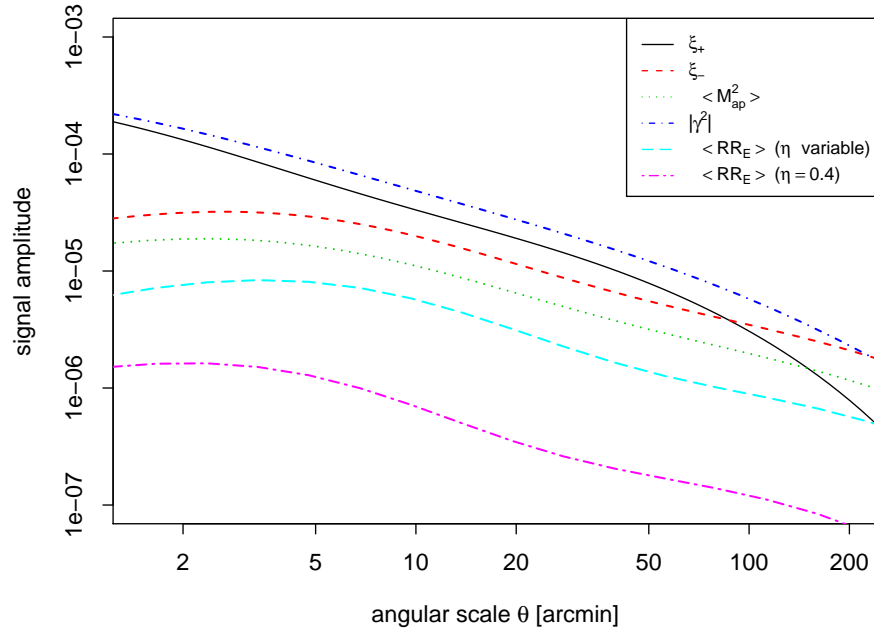
Figure 6.6: *This figure shows the ring statistics compared to other second-order cosmic shear measures, namely the shear dispersion, $\xi_{\pm}$ and the aperture mass dispersion. We see that even in the case of the improved ring statistics the signal is low compared to the other measures.*

6.4   Covariance of the ring statistics

In order to compare $\langle \mathcal{R}\mathcal{R}_{\mathrm{E}} \rangle$ and $\langle M_{\mathrm{ap}}^2 \rangle$ in a likelihood analysis we have to derive the corresponding covariances. Both are calculated from the 2PCF covariance; a corresponding expression for $\langle M_{\mathrm{ap}}^2 \rangle$ is given in (5.2). The covariance of the ring statistics is defined as

$$C_{\mathcal{R}}(\Psi_k, \Psi_l) = \left\langle \hat{R}_{\mathrm{E}}^2(\Psi_k)\, \hat{R}_{\mathrm{E}}^2(\Psi_l) \right\rangle - \langle \mathcal{R}\mathcal{R}_{\mathrm{E}} \rangle (\Psi_k)\, \langle \mathcal{R}\mathcal{R}_{\mathrm{E}} \rangle (\Psi_l), \tag{6.31}$$

where $\hat{R}_{\mathrm{E}}^2$ denotes the estimator of the ring statistics. To calculate this estimator from a binned 2PCF data vector with bin width $\Delta\vartheta_i$ we replace the integrals (6.26) by a sum over the bins.

$$\hat{R}_{\mathrm{E}}^2(\Psi) = \frac{1}{2} \sum_{i=1}^{I} \frac{\Delta\vartheta_i}{\vartheta_i} \left[\hat{\xi}_+(\vartheta_i)\, Z_+(\vartheta_i, \eta) + \hat{\xi}_-(\vartheta_i)\, Z_-(\vartheta_i, \eta) \right], \tag{6.32}$$

with $\hat{\xi}_{\pm}(\vartheta_i)$ denoting the estimator of the i-th 2PCF bin. The upper limit I in (6.32) denotes the bin up to which $\vartheta_i < \Psi$. Inserting (6.32) into (6.31) we derive

$$
\begin{aligned}
C_{\mathcal{R}}(\Psi_k, \Psi_l) \;=\; & \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{\Delta\vartheta_i \Delta\vartheta_j}{4\,\vartheta_i \vartheta_j} \\[2mm]
& \times \left[\sum_{m,n=+,-} Z_m(\vartheta_i, \Psi_k)\, Z_n\!\left(\vartheta_j, \Psi_l\right) C_{mn}(\vartheta_i, \vartheta_j) \right].
\end{aligned}
\tag{6.33}
$$

where I and J denote the bins up to which $\vartheta_i < \Psi_k$ ($\vartheta_j < \Psi_l$) holds and C_{mn} is the 2PCF covariance.

In order to illustrate the correlation between the individual data points we calculate the correlation matrix $\mathbf{R}$ for $\langle M_{\mathrm{ap}}^2 \rangle$ and $\langle \mathcal{RR}_{\mathrm{E}} \rangle$ from the corresponding covariance matrix. For $\mathbf{C}$ being the covariance of either $\langle M_{\mathrm{ap}}^2 \rangle$ or $\langle \mathcal{RR}_{\mathrm{E}} \rangle$ the correlation coefficient is defined as

$$ \mathrm{R}_{ij} = \frac{\mathrm{C}_{ij}}{\sqrt{\mathrm{C}_{ii} \mathrm{C}_{jj}}}, \tag{6.34} $$

which implies that $\mathbf{R}$ is one on the diagonal.

The upper row in Fig. 6.7 shows the correlation matrix of ring statistics (left panel) and aperture mass dispersion (right panel). Here, we calculate the covariance of $\langle \mathcal{RR}_{\mathrm{E}} \rangle$ and $\langle M_{\mathrm{ap}}^2 \rangle$ from a Gaussian 2PCF covariance matrix and then calculate the correlation matrix via (6.34). Starting from the diagonal, which contains the maximum values, the correlation drops quickly in both cases. The n-th contour line corresponds to values of 0.11^n (hence, "red"$\equiv$0.11, "orange" $\equiv$ 0.0121, etc.). We clearly see that the data points of the ring statistics are much less correlated than those of the aperture mass dispersion. The contour level depends mainly on the ratio θ_1/θ_2 and Ψ_1/Ψ_2, respectively. We compare the correlation coefficients for $\theta_1/\theta_2 = \Psi_1/\Psi_2 = 0.5$ and find that $\mathrm{R}_{ij}(\langle M_{\mathrm{ap}}^2 \rangle) = 0.35$, whereas $\mathrm{R}_{ij}(\langle \mathcal{RR}_{\mathrm{E}} \rangle) = 0.20$.

A qualitatively similar result is obtained when performing the same analysis with the 2PCF covariance matrix used in the latest CFHTLS analysis (Fu et al. 2008). There, the covariance was calculated from a theoretical model ($\Omega_{\mathrm{m}} = 0.27$, $\Omega_{\Lambda} = 0.73$, $h = 0.73$, $\Omega_{\mathrm{b}} = 0.044$, $\sigma_8 = 0.8$ and $n_{\mathrm{s}} = 1.0$) using the formula given in Schneider et al. (2002a). To account for non-Gaussianity Fu et al. (2008) applied the calibration factor introduced by Semboloni et al. (2007). From this 2PCF covariance we calculate covariances of the ring statistics and aperture mass dispersion and derive the corresponding correlation matrices. The result is illustrated in the lower row of Fig. 6.7. Here, the n-th contour line corresponds to values of 0.5^n (hence, "red" $\equiv$ 0.5, "orange" $\equiv$ 0.25, etc.). Compared to the Gaussian case the contours broaden, which is not directly seen in the figure due to the different contour levels. Comparing the correlation between the ring statistics and the aperture mass dispersion we obtain a similar result as in the Gaussian case. The data points of the ring statistics are significantly less correlated compared to those of the aperture mass dispersion. However, in the non-Gaussian case the correlation coefficients of $\langle M_{\mathrm{ap}}^2 \rangle$ and $\langle \mathcal{RR}_{\mathrm{E}} \rangle$ do not only depend on the ratio θ_1/θ_2 and Ψ_1/Ψ_2, respectively; they also depend on the absolute values θ_1, θ_2 and Ψ_1, Ψ_2, respectively. A quantitative comparison similar to the Gaussian case is more difficult here; we will analyze this in more detail in a future paper.

Note that the angular range of $\langle \mathcal{RR}_{\mathrm{E}} \rangle$ and $\langle M_{\mathrm{ap}}^2 \rangle$ is different. The first extends from $\Psi = 0.1'$ to $470'$, whereas the second is limited to an angular scale of $\theta \in [1; 230]$ arcmin. The lower limit is a consequence of the E/B-mode mixing of $\langle M_{\mathrm{ap}}^2 \rangle$ on small scales (Sect. 6.1). In addition, it should be mentioned that the scales of the ring statistics and aperture mass dispersion cannot be compared easily. For example, $\langle \mathcal{RR}_{\mathrm{E}} \rangle (\Psi)$ contains information from the 2PCF with $\vartheta \leq \Psi$, whereas $\langle M_{\mathrm{ap}}^2 \rangle (\theta)$ contains information from the 2PCF up to 2ϑ.
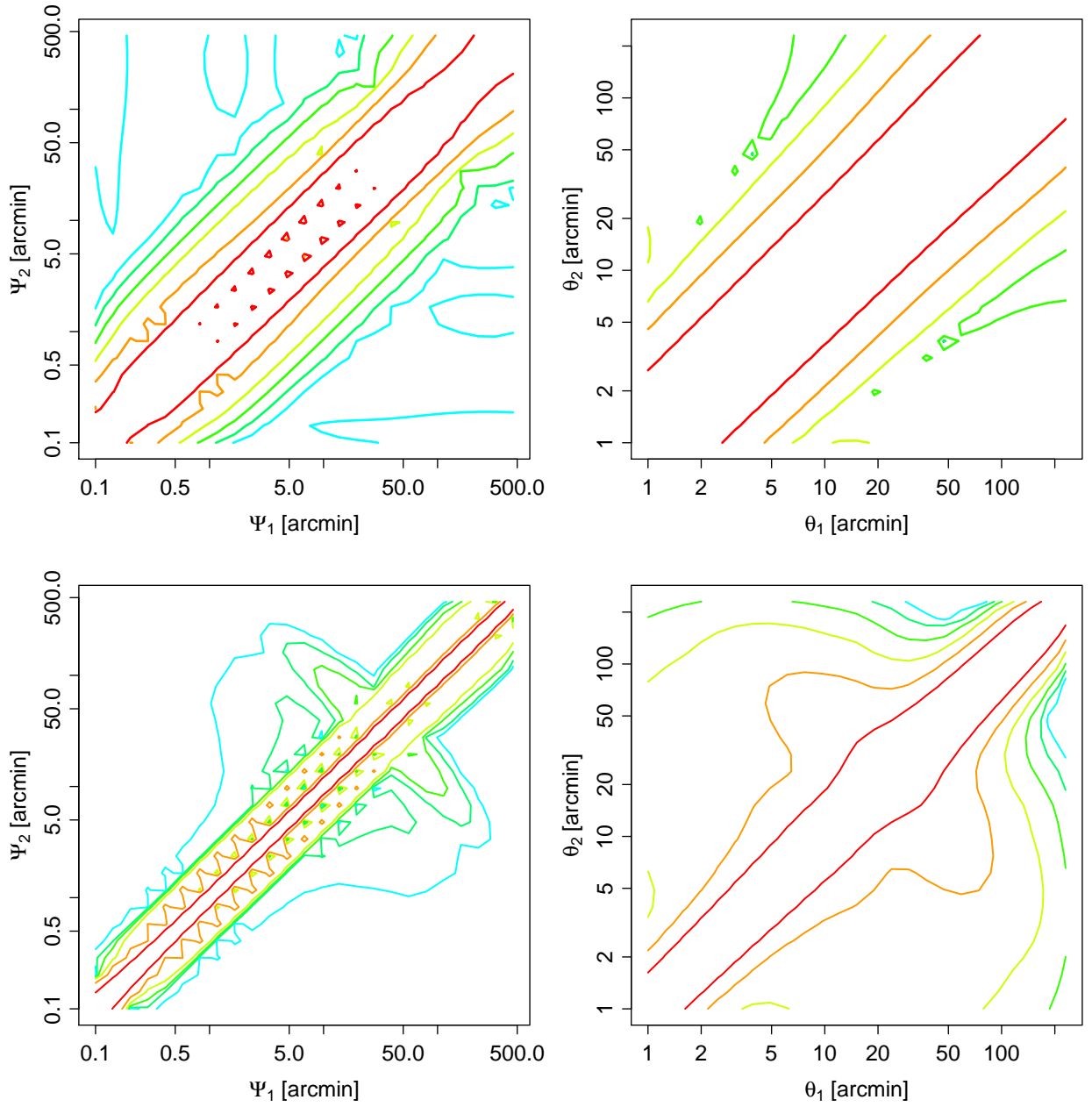
Figure 6.7: *This figure shows the correlation matrices of $\langle \mathcal{R}\mathcal{R}_{\mathrm{E}} \rangle$ (left panels) and $\langle M_{\mathrm{ap}}^2 \rangle$ (right panels) under the assumption of a Gaussian shear field (upper row) and when taking non-Gaussianity into account (lower row). The n-th contour line marks values of $(0.11)^n$ in the upper panels and $(0.5)^n$ in the lower ones. More precisely, for the Gaussian case: "red" contours corresponds to 0.11, "orange" to 0.0121, etc. For the non-Gaussian case: "red" corresponds to 0.5, "orange" to 0.25, etc.*

Figure 6.8: *This figure shows the result of the likelihood analysis for $\langle \mathcal{RR}_E \rangle$ (left panel) and $\langle M_{ap}^2 \rangle$ (right panel) with simulated data but employing the original covariance matrix of the latest CFHTLS analysis (Fu et al. 2008).*

6.5 Comparison of the information content of $\langle \mathcal{RR}_E \rangle$ and $\langle M_{ap}^2 \rangle$

We now compare the information content of ring statistics and aperture mass dispersion. We simulate a set of 2PCFs on an angular range which is similar to that of the CFHTLS 2PCFs, i.e. $\vartheta \in [0.05; 470]$ arcmin. The 2PCFs are calculated from P_E as described in Sect. 5.2, but with the transfer function from Efstathiou et al. (1992). From this set of theoretical 2PCFs we calculate the data vectors of $\langle \mathcal{RR}_E \rangle$ (with η_{var}) and $\langle M_{ap}^2 \rangle$

$$\mathcal{R} = \begin{pmatrix} \langle \mathcal{RR}_E \rangle (\Psi_1) \\ \vdots \\ \langle \mathcal{RR}_E \rangle (\Psi_n) \end{pmatrix} \quad \text{and} \quad \langle \boldsymbol{M_{ap}^2} \rangle = \begin{pmatrix} \langle M_{ap}^2 \rangle (\theta_1) \\ \vdots \\ \langle M_{ap}^2 \rangle (\theta_n) \end{pmatrix}, \tag{6.35}$$

with $\Psi \in [0.1; 460]$ and $\theta \in [1; 230]$ in 20 logarithmic bins each. Our fiducial model is chosen to be similar to that of the Millennium simulation, i.e. $\Omega_m = 0.25$, $\sigma_8 = 0.9$, $H_0 = 0.73$ and $\Omega_b = 0.04$; the source galaxies are assumed to be at a constant redshift $z = 1$. In the following likelihood analyses we consider a two-dimensional parameter space, namely Ω_m and σ_8. The corresponding covariances of $\langle \mathcal{RR}_E \rangle$ and $\langle M_{ap}^2 \rangle$ are calculated from th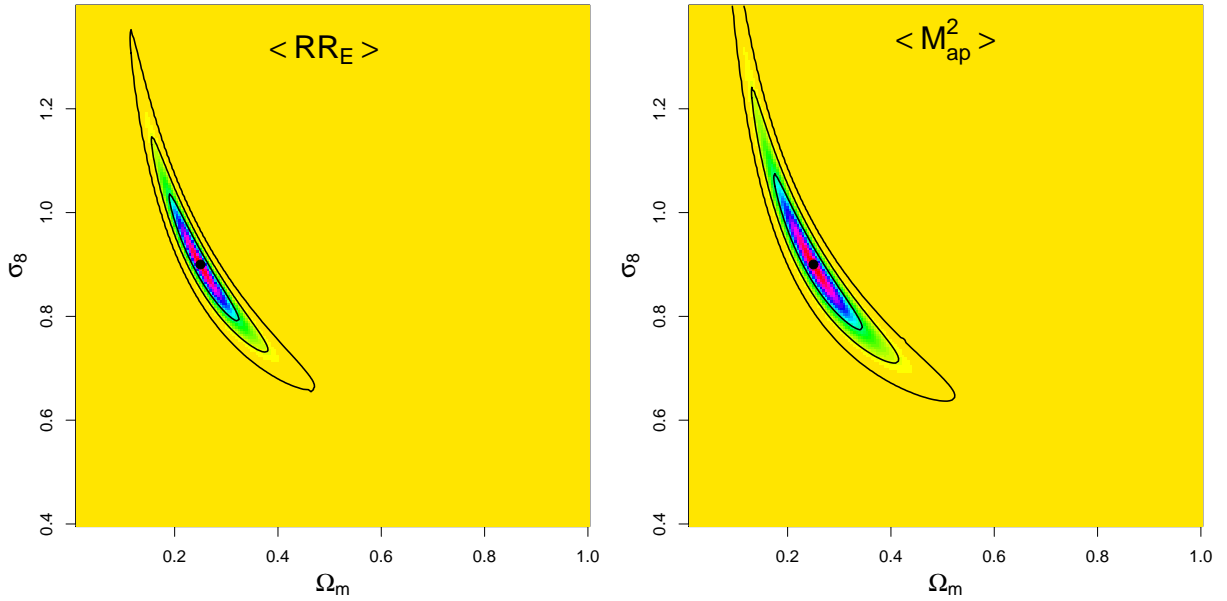e CFHTLS 2PCF covariance via (6.33) and (5.2), respectively. Figure 6.8 shows the result of the likelihood analysis for $\langle \mathcal{RR}_E \rangle$ (left panel) and $\langle M_{ap}^2 \rangle$ (right panel). We see that the ring statistics data vector is a clear improvement over the aperture mass dispersion. This can be explained by the fact that the ring statistics can probe much smaller scales than $\langle M_{ap}^2 \rangle$ and, more important, these scales are much less correlated.

6.6 Ring statistics with the CFHTLS

In this section we measure the ring statistics signal with CFHTLS data. We perform this analysis for the optimized ring statistics (with η_{var}) as well as for seven fixed η, namely 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, and 0.7. The results are illustrated in Fig. 6.9, the black data points correspond to the E-mode signal calculated via (6.26), the red data points correspond to the B-mode signal calculated via (6.28). The error bars are calculated from the diagonal elements of the ring statistics covariance matrix which was calculated from the CFHTLS covariance (see Sect. 6.4 and Fu et al. 2008). From $\mathbf{C}_{\mathcal{R}}$ we calculate the error for the i-th E-mode data point as $\sqrt{C_{ii}}$. This method accounts for statistical noise and non-Gaussian cosmic variance. The error bars on the B-modes contain only statistical noise. As pointed out by Fu et al. (2008) this error analysis does not include systematic errors, which might lead to an underestimation in the error bars .
There are several things to mention regarding Fig. 6.9. First, the improved ring statistics gives a significantly stronger signal compared to the ring statistics with fixed η. When choosing $\eta > 0.2$ hardly any signal is noticeable. Second, we see that the data points for η_{var} and $\eta = 0.1$ are almost identical on scales below 2′. This can be explained when looking at Fig. 6.5, which shows that the difference between both expected signals vanishes on small scales. Third, below 5′ we measure an E- and a B-mode signal which have approximately the same amplitude. This feature will be subject of future research; in particular, we will sample these small scales with more data points in order to verify our preliminary finding.
In Fig. 6.10 (left column) we compare the signal of the ring statistics above 5′ for two cases, namely η_{var} and $\eta = 0.1$. We see that the shear signal improves significantly when using the η_{var}; this becomes evident in particular on large scales. This behavior is expected from Fig. 6.5 which shows that the difference between both signals increases at large Ψ. Comparing the signal obtained from the ring statistics to the aperture mass signal obtained in the analysis of Fu et al. (2008) (right column of Fig. 6.10) we see a good qualitative agreement between both results. Similar to the $\langle M_{\mathrm{ap}}^2 \rangle$ signal, the ring statistics measures a significant B-mode on larger scales, more precisely at $\Psi \approx 120'$. In the $\langle M_{\mathrm{ap}}^2 \rangle$ signal this B-mode occurs at $\theta \approx 60'$. This is not a contradiction; it can be explained by the fact that $\langle \mathcal{R}\mathcal{R}_{\mathrm{E}} \rangle (\Psi)$ contains information from the 2PCF with $\vartheta \leq \Psi$, whereas $\langle M_{\mathrm{ap}}^2 \rangle(\theta)$ contains information from the 2PCF up to 2ϑ.

6.6.1 Parameter constraints with the ring statistics

In this section we employ the ring statistics to constrain cosmological parameters; more precisely, we constrain σ_8 depending on Ω_{m} under the assumption of a flat universe with $h = 0.73$, $\Omega_{\mathrm{b}} = 0.044$, and $n_{\mathrm{s}} = 1.0$. The model data vectors for the ring statistics are calculated from the 2PCF as described in Sect. 6.5, however we choose a redshift distribution of source galaxies similar to that of Benjamin et al. (2007)

$$n(z) = \frac{\beta}{z_0 \Gamma\left((1 + \alpha)/\beta\right)} \left(\frac{z}{z_0}\right)^{\alpha} \exp\left[-\left(\frac{z}{z_0}\right)^{\beta}\right], \qquad (6.36)$$

with $\alpha = 0.836$, $\beta = 3.425$, $z_0 = 1.171$.
In the likelihood analysis we choose a flat prior probability with cutoffs, which means $p(\boldsymbol{\pi})$ is
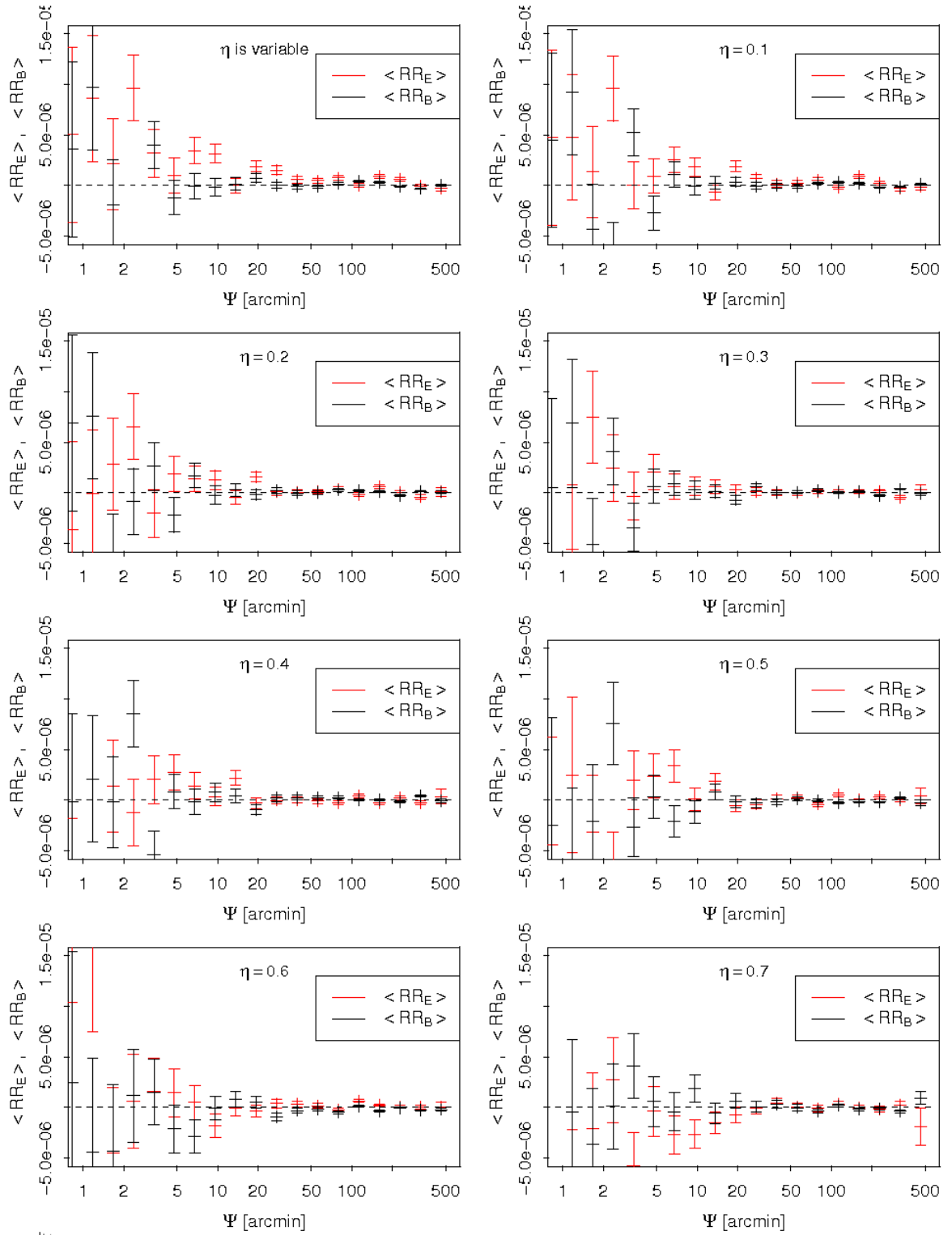
Figure 6.9: *This figure shows the ring statistics signal measured from the CFHTLS for various cases of η. The red data points correspond to the E-mode signal, the black data points to the B-mode signal.*
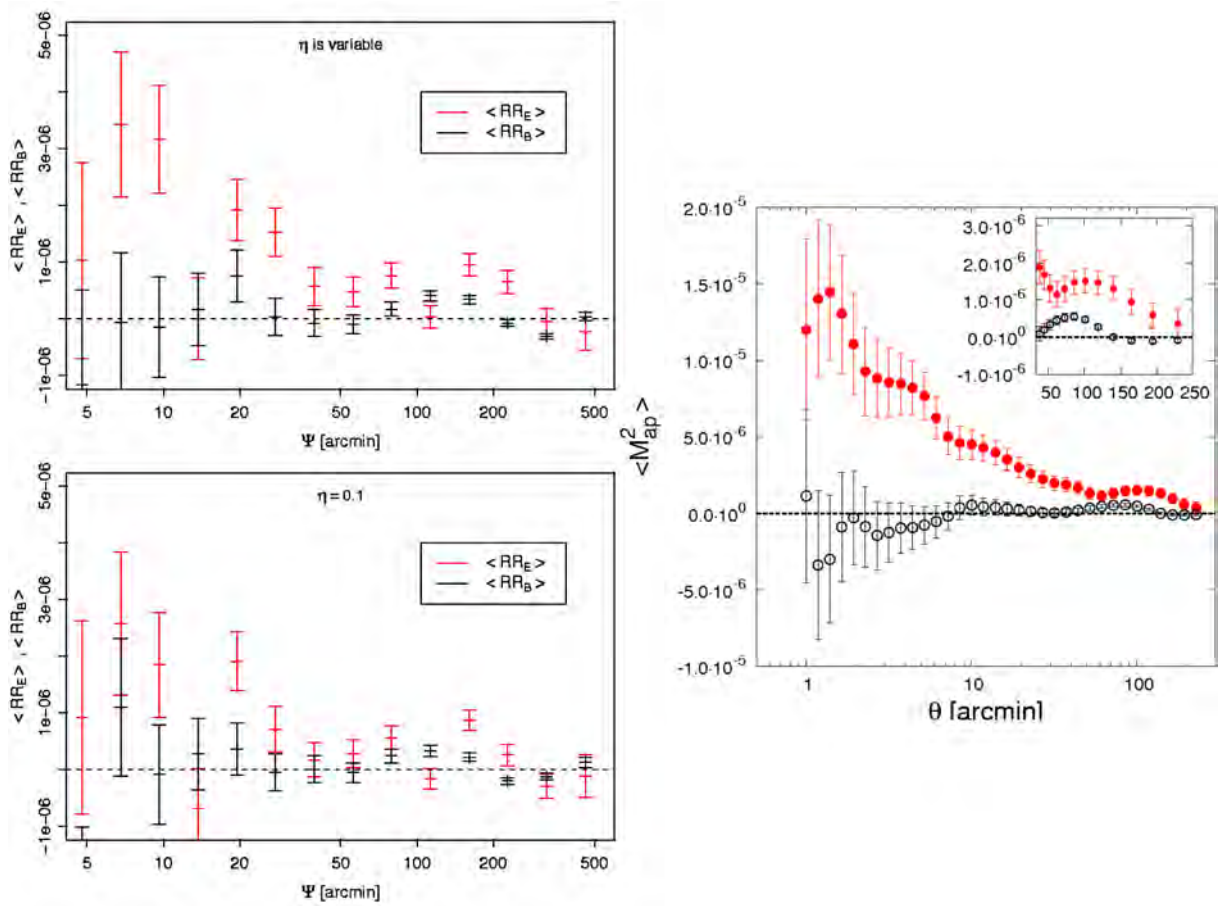
Figure 6.10: *The ring statistics for the case of a variable η (top left panel) and η = 0.1 (bottom left panel) compared to the aperture mass dispersion signal from the latest CFHTLS analysis (Fu et al. 2008).*

constant for all parameters inside a fixed interval (i.e. $\Omega_m \in [0.01; 1.0]$, $\sigma_8 \in [0.4; 1.4]$) and $p(\pi) = 0$ else. The ring statistics' covariance is derived from the 2PCF covariance of the Fu et al. (2008) analysis as described in Sect. 6.4. The result of the ring statistics likelihood analysis is shown in Fig. 6.11. For given Ω_m we indicate the best-fit value of σ_8 by the dotted red curve. This curve can be parameterized as $\sigma_8(\Omega_m/0.25)^{0.49} = 0.82^{+0.02}_{-0.04}$. The error bars, illustrated by the horizontal dashed lines in Fig. 6.11, correspond to the 1σ limits of σ_8 when assuming the prior $\Omega_m = 0.25$ (vertical dashed line). A comparison of our results using the ring statistics with the aperture mass dispersion analysis of Fu et al. (2008) and Benjamin et al. (2007) can be found in Table 6.1. In both cases the results agree well within the error bars. In addition, our estimates agree with those of the WMAP5 analysis (see Tab. 2.1). As expected from our analysis in Sect. 6.5 the likelihood contours of the ring statistics are tighter compared to those of the aperture mass dispersion (compare Fig. 6.11 and Fig. 6.12). This can also be seen when comparing the error bars of Table 6.1. However, the error analysis is ongoing work and must be improved in the future, e.g. systematic errors should be taken into account. Furthermore, it will be important to

Table 6.1: *Comparison of the Ω_m vs. σ_8 parameter constraints from the ring statistics to the $\langle M_{ap}^2 \rangle$ analysis of Fu et al. (2008) and Benjamin et al. (2007).*

Prior	Analysis	Result
$\Omega_m = 0.25$	Fu et al. (2008)	$\sigma_8(\Omega_m/0.25)^{0.64} = 0.785 \pm 0.043$
$\Omega_m = 0.25$	Ring statistics	$\sigma_8(\Omega_m/0.25)^{0.48} = 0.820^{+0.021}_{-0.037}$
$\Omega_m = 0.24$	Benjamin et al. (2007)	$\sigma_8(\Omega_m/0.24)^{0.56} = 0.86 \pm 0.06$
$\Omega_m = 0.24$	Ring statistics	$\sigma_8(\Omega_m/0.24)^{0.48} = 0.837^{+0.018}_{-0.035}$

Figure 6.11: *This figure shows the 68%, 95%, 99.9% credible intervals in the Ω_m-σ_8-plane when measuring the shear from CFHTLS data with the ring statistics. For given Ω_m the red dots indicate the value of σ_8 with the highest probability. The dotted vertical line corresponds to $\Omega_m = 0.25$, the dotted vertical lines represent the 1σ error of σ_8 for a given $\Omega_m = 0.25$.*
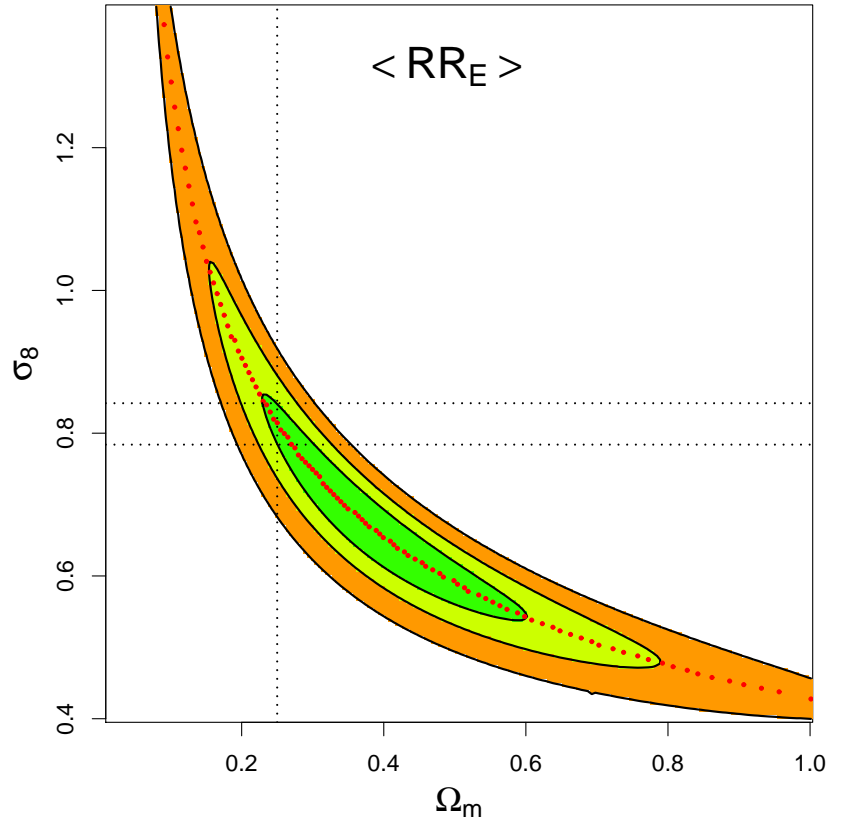
work on theoretical predictions; the model data vectors derived from the fit-formula of of Smith et al. (2003) is not sufficient for precision cosmology.

6.7 Concluding remarks on the ring statistics

In this chapter we outlined that the ring statistics improves on deficits of commonly used methods to decompose E- and B-modes. These methods require knowledge on the 2PCF on scales which are not measurable; usually this missing information is added in form of 2PCFs calculated from a theoretical model. In contrast, the ring statistics separates E- and B-modes properly using
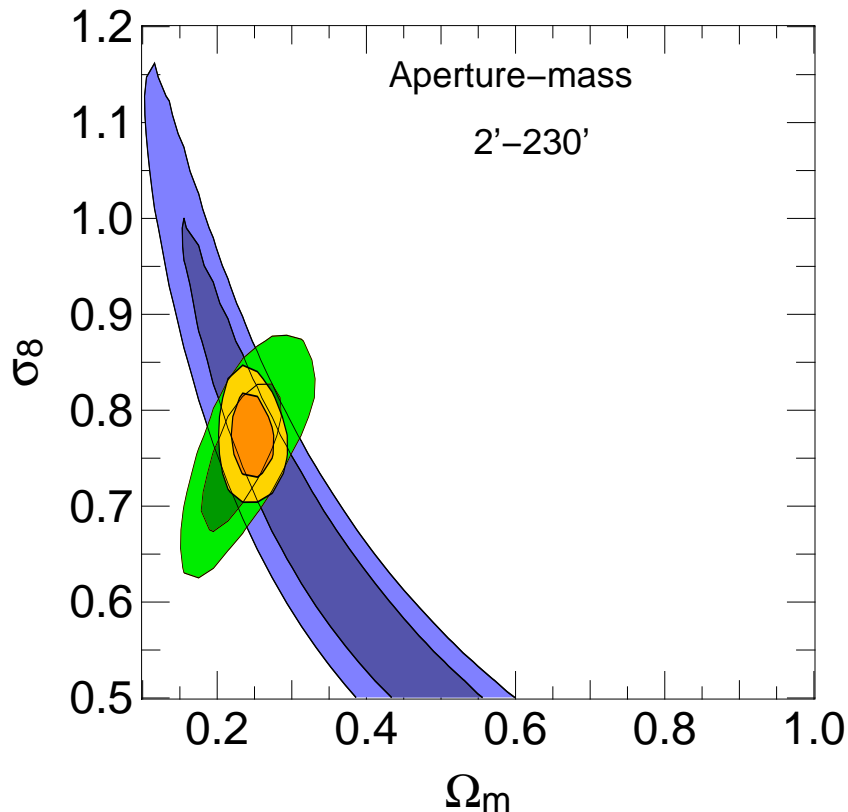
Figure 6.12: *This figure shows the 1σ and 2σ likelihood contours in the Ω_m-σ_8-plane of the $\langle M_{ap}^2 \rangle$ analysis of Fu et al. (2008) (purple). The green contours correspond to the result of WMAP3 and the yellow contours represent the combined constraints.*

2PCFs measured on a finite interval of ϑ-values. As outlined in Schneider & Kilbinger (2007) the filter functions of the ring statistics, i.e. $Z_\pm$, are in general complicated to calculate; the authors restrict the free parameters of the ring statistics to one parameter, namely η. This parameter is held fixed, independent of the angular scale Ψ at which the ring statistics is evaluated. In this chapter, we improve on the condition of a fixed η by choosing a variable η which improves the ring statistics' signal, particularly on large scales.

Furthermore, we present a formula to calculate the ring statistics' covariance from the 2PCF covariance. This covariance is employed to compare the correlation of data points and the information content of ring statistics and aperture mass dispersion. We find that the ring statistics is less correlated and therefore has a higher information content than the aperture mass dispersion. A second reason for the high information content of $\langle \mathcal{R}\mathcal{R}_E \rangle$ is that it can be calculated on smaller angular scales than $\langle M_{ap}^2 \rangle$ which suffers from E/B-mode mixing on these scales.

We apply the ring statistics to CFHTLS data and measure a shear signal. This signal vanishes when choosing a fixed $\eta > 0.2$; hence, we conclude that the ring statistics with η_{var} significantly improves on the signal-to-noise ratio. Above $5'$ we measure a clear E-mode shear signal, whereas the B-mode is mostly consistent with zero. These results agree with the aperture mass dispersion analysis of Fu et al. (2008); similar to them we only measure a B-mode signal on large scales. More precisely, we measure a B-mode signal at $\Psi \approx 120'$ with $\langle \mathcal{R}\mathcal{R}_E \rangle$, which is comparable to the B-mode signal at $\theta \approx 60'$ measured with $\langle M_{ap}^2 \rangle$. In the last section of this chapter we employ the ring statistics to measure cosmological parameters, in particular σ_8 in combination with Ω_m.

Our result $\sigma_8(\Omega_m/0.25)^{0.49} = 0.82^{+0.02}_{-0.04}$ is comparable to the results of foregoing cosmic shear analyses which are based on the aperture mass dispersion or the E/B-mode correlation function, e.g. Benjamin et al. (2007) or Fu et al. (2008).

Below $5'$ we measure an E- as well as a B-mode signal. This region of small Ψ will be examined closely in the near future, in particular we will sample this range with more data points. The fact, that data points of the ring statistics have so small correlations enables us to determine the contaminated scales very accurately.

The noise-level of the ring statistics on small scales can be reduced by increasing the number of galaxy pairs within the contributing 2PCF-bins. The number of galaxy pairs inside a 2PCF-bin increases quadratically with $\bar{n}$, therefore it would be interesting to test the ring statistics on a data set with higher $\bar{n}$. Similarly, an increased survey volume will significantly enhance the constraints, for the reason that the cosmic variance scales with $1/A$. For example, the CFHTLS data we consider here covers an area of $34.2\,\text{deg}^2$ with $\bar{n} = 13.3$. Testing the ring statistics on the full CFHTLS sample ($172\,\text{deg}^2$) would be an interesting project in the future.

A second possibility to increase the signal-to-noise ratio, is to optimize the filter function of the ring statistics. As outlined in Schneider & Kilbinger (2007) the ring statistics is a special case of a general E- and B-mode decomposition; there exists a large set of filter functions which separate E-and B-modes using a 2PCF measured over a finite angular range. It will be future work to find the filter function for an E- and B-mode decomposition on a finite interval with optimal signal strength.

Chapter 7

Improved Likelihood Analysis for Cosmic Shear Data

In cosmic shear likelihood analyses the covariance is most commonly assumed to be constant in parameter space. Therefore, when calculating the covariance matrix analytically or from simulations, its underlying cosmology should not influence the likelihood contours. In this chapter we examine whether the aforementioned assumptions hold and quantify how strong cosmic shear covariances vary within a reasonable parameter range. We calculate Gaussian covariances analytically for 2500 different cosmologies; in order to quantify the impact on the parameter constraints we perform a likelihood analysis for each covariance matrix and compare the likelihood contours. To improve on the assumption of a constant covariance, we use an adaptive covariance matrix, which is continuously updated according to the point in parameter space where the likelihood is evaluated. As a side-effect, this cosmology-dependent covariance improves the parameter constraints. We examine this fact more closely using the Fisher-matrix formalism. In addition, we quantify the impact of non-Gaussian covariances on the likelihood contours using a ray-tracing covariance derived from the Millennium simulation. In this ansatz we return to the approximation of a cosmology-independent covariance matrix; in order to minimize the error due to this approximation, we develop the concept of an iterative likelihood analysis.

7.1 Introduction to covariances

Obtaining appropriate covariances is a crucial issue in this context of a precision cosmology likelihood analysis. Several methods are suggested in the literature and have been applied to cosmic shear data. An analytic expression for covariances assuming a Gaussian shear field is derived in Schneider et al. (2002a) and confirmed in Joachimi et al. (2008) who use a power spectrum approach which significantly reduces the computational effort in the calculation. This analytic expression has been used for parameter estimation in many surveys (e.g. van Waerbeke et al. 2005; Semboloni et al. 2006; Hoekstra et al. 2006). However, the assumption of a Gaussian shear field breaks down on small scales; according to Kilbinger & Schneider (2005) and Semboloni et al. (2007) non-linear effects already become important at angular scales $\lesssim 10$ arcmin.

To account for non-Gaussianity, Semboloni et al. (2007) obtain a calibration factor which is derived from a comparison of Gaussian to ray-tracing covariances. An application of this method to real data can be found in Fu et al. (2008). A second approach is the derivation of the covariance matrix from the data (e.g. Hetterscheidt et al. 2007; Massey et al. 2007b). Here, the covariance is calculated via field-to-field variation which involves a separation of the data set into many independent subsamples. This might lead to a loss of information on large scales if the survey is not sufficiently large. Third, one can estimate the covariance matrix from ray-tracing simulations, a method which circumvents the aforementioned loss in information. Although, in this method the covariance is again derived via field-to-field variation, we can choose a sufficiently large numerical simulation to create many independent subsamples of adequate size.

Note, that the last two methods involve an estimation process in the determination of the covariance matrix, which means that the inverse is biased and one has to correct for this effect (Anderson 2003; Hartlap et al. 2007). Nevertheless, deriving covariance matrices from ray-tracing simulations seems to be a promising method as it preserves all the information in the data and additionally takes the non-Gaussianity of the shear field into account.

The analytic expression and the ray-tracing covariance assume a specific cosmological model in their derivation. So far, cosmic shear likelihood analyses treat the covariance matrix as constant in parameter space, hence its underlying cosmology is assumed not to influence the parameter constraints. In the following we check for this assumption and in case it does not hold, to present an improved likelihood formalism for future surveys.

7.2 Variation of covariances in parameter space

We select a two-dimensional parameter grid with 50×50 gridpoints of $\Omega_m \in [0.2; 0.4]$ and $\sigma_8 \in [0.6; 1.0]$. For each grid point we calculate a covariance analytically using (4.3) - (4.8). The shear power spectra P_E are obtained from the density power spectra P_δ employing Limber's equation. To derive P_δ we assume an initial Harrison-Zeldovich power spectrum ($P_\delta(k) \propto k^{n_s}$ where $n_s = 1$) with the transfer function from Efstathiou et al. (1992). For the calculation of the non-linear evolution we use the fitting formula of Smith et al. (2003). Throughout this chapter, we assume a flat universe and fix all cosmological parameters except Ω_m and σ_8, more precisely $H_0 = 0.73$ and $\Omega_b = 0.04$. These values for H_0 and Ω_b together with $\Omega_m = 0.25$ and $\sigma_8 = 0.9$ define our fiducial cosmological model, which we have chosen similar to the cosmology of the Millennium Simulation (Springel et al. 2005) for a later comparison of Gaussian and ray-tracing covariances. We assume all source galaxies to be at redshift $z_0 = 1.0$. Using a redshift distribution instead would not change our results qualitatively. In addition to cosmology, the covariance depends on survey parameters. The scaling relations given in Sect. 7.2 are generally valid and independent of survey parameters. In case of the likelihood analyses in Sects. 7.3 and 7.4 we choose, unless stated otherwise, an intrinsic ellipticity noise of $\sigma_\epsilon = 0.4$, a number density of source galaxies of $\bar{n} = 10/\text{arcmin}^2$ (similar to the values of the Dark Energy survey), and a survey which covers $A = 900 \text{ deg}^2$. The angular scale of the 2PCF data vector for which we calculate the covariances covers a range from 0.1 arcmin to 180 arcmin, which is divided into
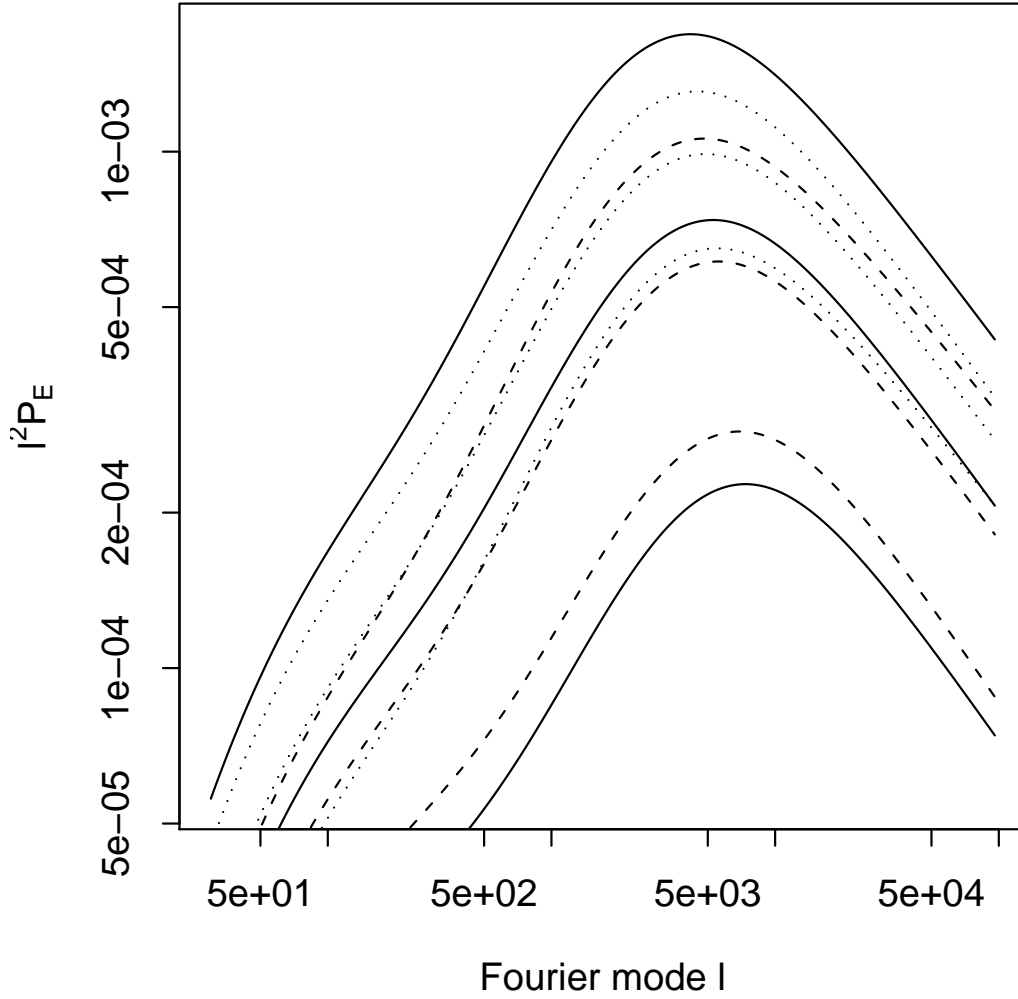
Figure 7.1: *The dimensionless shear power spectrum $\ell^2 P_E$. The solid curves correspond to variation in Ω_m and σ_8: $\Omega_m = 0.2$, $\sigma_8 = 0.6$ (lower), $\Omega_m = 0.3$, $\sigma_8 = 0.8$ (middle), $\Omega_m = 0.4$, $\sigma_8 = 1.0$ (top). The dashed curves show variation in σ_8 with $\Omega_m = 0.25$: $\sigma_8 = 0.6$ (lower), $\sigma_8 = 0.8$ (middle), $\sigma_8 = 1.0$ (top). The dotted curves show variation in Ω_m with a constant $\sigma_8 = 0.9$: $\Omega_m = 0.2$ (lower), $\Omega_m = 0.3$ (middle), $\Omega_m = 0.4$ (top).*

50 logarithmic bins.

7.2.1 A fast method to calculate covariances for arbitrary Ω_m and σ_8

From (4.7) and (4.8) one directly sees that the covariance matrix depends on the cosmological model, which enters with the power spectrum P_E. Figure 7.1 illustrates the change in P_E when varying only Ω_m, or σ_8, and both parameters simultaneously; we see that it increases with Ω_m as

well as with σ_8.

For a given cosmological model we can calculate the covariance directly from (4.3) - (4.8). Performing this calculation for many sets of parameters is time-consuming; hence we seek a scaling relation, which relates the covariances of an arbitrary cosmology $\mathbf{C}_\pi$ to a reference model $\mathbf{C}_{\pi_0}$. A basic theorem in statistics states (e.g. Anderson 2003), that if there is a relation between two data vectors x and y which reads $y = \mathbf{A}x$ ($\mathbf{A}$ being a matrix), the relation of the covariances of x and y can be written as

$$
\begin{aligned}
\mathbf{C}_y &= \left\langle (y - \langle y \rangle)(y - \langle y \rangle)^{\mathrm{t}} \right\rangle \\
&= \left\langle (\mathbf{A}x - \langle \mathbf{A}x \rangle)(\mathbf{A}x - \langle \mathbf{A}x \rangle)^{\mathrm{t}} \right\rangle \\
&= \mathbf{A}\mathbf{C}_x\mathbf{A}^t .
\end{aligned}
\tag{7.1}
$$

In this derivation $\mathbf{A}$ must be independent of the ensemble average. If we apply the above ansatz to the 2PCF, it seems reasonable to define a scaling relation for parameter dependent covariances as

$$
\mathbf{C}_{\xi_\pi} = \mathbf{A}\mathbf{C}_{\xi_{\pi_0}}\mathbf{A}^{\mathrm{t}} ,
\tag{7.2}
$$

where we can calculate the scaling matrices $\mathbf{A}$ using the 2PCF

$$
diag(\mathbf{A}) = \boldsymbol{\xi}_\pi/\boldsymbol{\xi}_{\pi_0} .
\tag{7.3}
$$

In contrast to a covariance matrix, the 2PCF can be calculated extremely fast for many different cosmologies via (3.51) and (3.52). Hence, it would be a fast and convenient method to calculate the covariance for a reference cosmology and then apply (7.2) to obtain covariances for arbitrary cosmological parameters. Unfortunately, we cannot transfer this method directly to the cosmic shear case. Recall, that the 2PCF is derived from the measured ellipticities of galaxies. Schneider et al. (2002b) have shown that the intrinsic ellipticity terms cancel out in the derivation of the 2PCF estimator, hence the 2PCF is defined only in terms of the shear. In contrast, the 2PCF covariance does not only consist of terms coming from the shear, but has additional noise terms which arise from the intrinsic ellipticity of galaxies. The pure shot noise term (4.6) is independent of cosmology and, as explained below, the mixed term cannot be scaled with relation (7.2), which is quadratic in the 2PCF.

However, in the limit of a noise-free covariance, i.e. considering only the cosmic variance term, a scaling relation similar to (7.2) exists. We explicitly prove this below; in particular, we show that the scaling matrices are independent of the ensemble average. The cosmic variance term can be calculated via (4.7). Cosmology only enters with the power spectrum, hence the relation of π to π_0 can be described as $P_{\mathrm{E}}(\ell, \pi) = a(\ell, \pi)P_{\mathrm{E}}(\ell, \pi_0)$. Using this relation we transform the cosmic variance term (4.7) for given bins ϑ_i, ϑ_j as follows

$$
\begin{aligned}
V_{\pm\pm}(\boldsymbol{\pi}) &= \frac{1}{\pi A} \int_0^\infty \mathrm{d}\ell\, \ell\, \mathrm{J}_{0/4}(\ell\vartheta_i)\, \mathrm{J}_{0/4}(\ell\vartheta_j)\, P_{\mathrm{E}}^2(\ell, \boldsymbol{\pi}) \\
&= \frac{1}{\pi A} \sum_{\bar{\ell}} \Delta\bar{\ell}\, \bar{\ell}\, \mathrm{J}_{0/4}(\bar{\ell}\vartheta_i)\, \mathrm{J}_{0/4}(\bar{\ell}\vartheta_j)\, a^2(\bar{\ell}, \boldsymbol{\pi})P_{\mathrm{E}}^2(\bar{\ell}, \boldsymbol{\pi}_0),
\end{aligned}
\tag{7.4}
$$

where we discretize the integral into a sum of $\bar{\ell}$-bins. Next we insert equation (26) of Joachimi et al. (2008) (see also Kaiser 1998) but with $\sigma_\epsilon = 0$

$$\left\langle \Delta P_{\rm E}(\bar{\ell})\Delta P_{\rm E}(\bar{\ell}') \right\rangle = \frac{4\pi}{A\bar{\ell}\Delta\bar{\ell}}\, P_{\rm E}^2(\bar{\ell})\, \delta_{\bar{\ell}\bar{\ell}'}\,, \tag{7.5}$$

to rewrite (7.4) as

$$\begin{aligned}
V_{\pm\pm}(\boldsymbol{\pi}) &= \frac{1}{4\pi^2}\sum_{\bar{\ell},\bar{\ell}'}\Delta\bar{\ell}^2\bar{\ell}'\bar{\ell}\, {\rm J}_{0/4}(\bar{\ell}\vartheta_i)\, {\rm J}_{0/4}(\bar{\ell}'\vartheta_j)\, a(\bar{\ell},\boldsymbol{\pi})\, a(\bar{\ell}',\boldsymbol{\pi}) \\
&\quad \times \left\langle \Delta P_{\rm E}(\bar{\ell},\boldsymbol{\pi}_0)\Delta P_{\rm E}(\bar{\ell}',\boldsymbol{\pi}_0) \right\rangle\,.
\end{aligned} \tag{7.6}$$

The mean value theorem guarantees that there exist values $\bar{a}(\vartheta_i,\boldsymbol{\pi})$, $\bar{a}(\vartheta_j,\boldsymbol{\pi})$ such that (7.6) becomes

$$\begin{aligned}
V_{\pm\pm}(\boldsymbol{\pi}) &= \frac{\bar{a}(\vartheta_i,\boldsymbol{\pi})\, \bar{a}(\vartheta_j,\boldsymbol{\pi})}{4\pi^2}\left\langle \int_0^\infty {\rm d}\ell\, \ell\, {\rm J}_{0/4}(\ell\vartheta_i)\, \Delta P_{\rm E}(\ell,\boldsymbol{\pi}_0) \right. \\
&\quad \left. \times \int_0^\infty {\rm d}\ell'\, \ell'\, {\rm J}_{0/4}(\ell'\vartheta_j)\, \Delta P_{\rm E}(\ell',\boldsymbol{\pi}_0) \right\rangle \\
&= \bar{a}(\vartheta_i,\boldsymbol{\pi})\, \bar{a}(\vartheta_j,\boldsymbol{\pi})\, V_{\pm\pm}(\boldsymbol{\pi}_0)\,,
\end{aligned} \tag{7.7}$$

where we consider the limit $\Delta\bar{\ell} \longrightarrow 0$ in the first step. Comparing the expressions of $V_{\pm\pm}(\boldsymbol{\pi})$ and $V_{\pm\pm}(\boldsymbol{\pi}_0)$ we can calculate the scaling factors as

$$\bar{a}_{\pm\pm}(\vartheta_i,\boldsymbol{\pi}) = \frac{\int_0^\infty {\rm d}\ell\, \ell\, {\rm J}_{0/4}(\ell\vartheta_i)P_{\rm E}(\ell,\boldsymbol{\pi})}{\int_0^\infty {\rm d}\ell\, \ell\, {\rm J}_{0/4}(\ell\vartheta_i)P_{\rm E}(\ell,\boldsymbol{\pi}_0)} = \frac{\xi(\vartheta_i,\boldsymbol{\pi})}{\xi(\vartheta_i,\boldsymbol{\pi}_0)}\,, \tag{7.8}$$

where we inserted (3.51), (3.51) respectively, in the last step. This provides a fast and convenient method to scale the cosmic variance term in parameter space, due to the fact we can use a computationally efficient Hankel transformation for the calculation of the 2PCF.

In order to examine whether a scaling relation can be found for the mixed term, we refer to the corresponding expression derived in Schneider et al. (2002a).

$$M_{++} = \frac{2\sigma_\epsilon^2}{\pi A n}\int_0^\pi {\rm d}\varphi\, \xi_+(|\boldsymbol{\phi}|)\,, \tag{7.9}$$

$$M_{--} = \frac{2\sigma_\epsilon^2}{\pi A n}\int_0^\pi {\rm d}\varphi\, \xi_+(|\boldsymbol{\phi}|)\, \cos(4\varphi)\,, \tag{7.10}$$

$$\begin{aligned}
M_{+-} &= \frac{2\sigma_\epsilon^2}{\pi A n}\int_0^\pi {\rm d}\varphi\left[\sum_{k=0}^4 \binom{4}{k}(-1)^k\vartheta_i^k\,\vartheta_j^{4-k}\cos(k\varphi)\right] \\
&\quad \times (|\boldsymbol{\phi}|)^{-4}\,\xi_-(|\phi|)\,\cos(4\varphi)\,,
\end{aligned} \tag{7.11}$$

where we denote $|\boldsymbol{\phi}| = \sqrt{\vartheta_i^2 + \vartheta_j^2 - 2\vartheta_i\vartheta_j\cos\varphi}$. From (7.9) - (7.11) we see that the mixed term $M_{\pm\pm}$ scales linearly with the 2PCF which prevents a scaling relation similar to (7.2). Fortunately,

the direct calculation of the mixed term via (4.8) is comparatively fast, therefore, the scaling relation for the cosmic variance term already reduces the computational costs significantly. Nonetheless, we numerically derive a fit-formula for the linear term based on the following expression

$$M_{\pm\pm}(\vartheta_i, \vartheta_j, \boldsymbol{\pi}) = M_{\pm\pm}(\vartheta_i, \vartheta_j, \boldsymbol{\pi}_0) \left(\frac{\Omega_{\mathrm{m}}}{0.25}\right)^{\alpha} \left(\frac{\sigma_8}{0.9}\right)^{\beta}, \qquad (7.12)$$

with $\boldsymbol{\pi}_0$ being the fiducial model explained at the beginning of Sect. 7.2. The structure of this fit-formula is motivated by the intention to use as few fit-parameters as possible; additionally we require that in the limit of the fiducial model, $M_{\pm\pm}(\vartheta_i, \vartheta_j, \boldsymbol{\pi}) = M_{\pm\pm}(\vartheta_i, \vartheta_j, \boldsymbol{\pi}_0)$ must hold. The fit-parameters α and β vary depending on the scale ϑ_i, ϑ_j and are different for the different parts of the covariance matrix, $\mathbf{C}_{++}$, $\mathbf{C}_{--}$, and $\mathbf{C}_{+-}$. The tables with α and β can be found in the Appendix B.

7.2.2   Variation of the inverse covariance with Ω_{m} and σ_8

From the variation of the power spectrum with Ω_{m} and σ_8 (Sect. 7.2.1) it is clear that covariances vary with respect to comological parameters. For simplicity and in order to increase the readability of the following sections we refer to this variation as CDC-effect (CDC ≡ Cosmology-Dependent Covariances). In order to examine the CDC-effect more closely, recall that the structure of the covariance is given by

$$\mathbf{C} = \left(\begin{array}{c|c} \mathbf{C}_{++} & \mathbf{C}_{+-} \\ \hline \mathbf{C}_{+-}^{\mathrm{t}} & \mathbf{C}_{--} \end{array} \right)$$

and the individual parts are calculated from (4.3) - (4.8). From these equations we see that the covariances are filtered versions of the power spectrum, either filtered by a product of J_0's (in case of $\mathbf{C}_{++}$), J_4's ($\mathbf{C}_{--}$), or a combination of both ($\mathbf{C}_{+-}$). The strength of the CDC-effect depends on these filter functions, as they determine which parts of the power spectrum are sampled. A change in Ω_{m} and σ_8 affects all scales of the power spectrum almost similarly (see Fig. 7.1); therefore, the CDC-effect for the individual parts of $\mathbf{C}$ is also similar. However, this might change when considering different cosmological parameters, e.g. the shape parameter Γ. I contrast to a change in Ω_{m} and σ_8, which increases (decreases) the power spectrum on all scales, a change in Γ increases (decreases) small scales while decreasing (increasing) large scales. Therefore, a change in Γ causes a "rotation" in the power spectrum. The covariances are integrals over P_{E}, and depending on the filter function, such a rotational change in P_{E} can average out. A second argument why the individual covariance parts have different sensitivity to the CDC-effect is that $\mathbf{C}_{+-}$ is not affected from shot noise, hence a change in cosmology has a stronger impact on $\mathbf{C}_{+-}$ compared to $\mathbf{C}_{++}$ and $\mathbf{C}_{--}$.

In order to quantify the CDC-effect we examine the trace of the inverse covariance matrix $\mathbf{C}^{-1}$. The trace of the covariance itself is an improper measure for this effect, as it depends on the binning, which can be seen from (4.6). The trace of $\mathbf{C}$ becomes arbitrarily large when decreasing the bin width. In contrast, we checked numerically that for the trace of $\mathbf{C}^{-1}$ binning effects are negligible, once one has exceeded a minimum bin number (see Sect. 4.4.1). More precisely, once

the bin width of the 2PCF data vector is small enough that discretization effects are unimportant, the trace of $\mathbf{C}^{-1}$ hardly changes for different binning.

Figure 7.2 shows the trace of the inverse covariance matrix depending on Ω_m for various constant σ_8 (top) and vice versa (bottom). Here, we normalize the survey size to $A = 1\,\mathrm{deg}^2$; the other survey parameters are $\sigma_\epsilon = 0.4$ and $\bar{n} = 10/\mathrm{arcmin}^2$. We postpone a detailed analysis of how survey parameters influence the CDC-effect to Sect. 7.3. Qualitatively the result does not change for different survey parameters; the trace of $\mathbf{C}^{-1}$ decreases with increasing Ω_m or σ_8.
In addition, we perform a singular value decomposition (SVD) for each inverse covariance matrix. For the case of a symmetric and positive definite matrix, such as the inverse covariance matrix, an SVD yields the eigenvalues in decreasing order. For arbitrary i, we find that the i-th eigenvalue decreases when increasing Ω_m or σ_8. The strength of the CDC-effect, i.e. the gradient of the traces, depends on the considered point in parameter space.

7.3 Impact of the CDC-effect on parameter estimation

7.3.1 Results of the likelihood analysis

In Sect. 7.2 we calculate 2500 covariances covering a parameter range of $\Omega_m \in [0.2; 0.4]$ and $\sigma_8 \in [0.6; 1.0]$. Here, we want to examine how the CDC-effect influences the likelihood contours, hence for each of the 2500 covariance matrices we perform a likelihood analysis. Throughout the whole likelihood analysis we assume the ΛCDM model. Here, we assume flat priors with cut-offs, which means $p(\boldsymbol{\pi})$ is constant for all parameters inside a fixed interval (i.e. $\Omega_m \in [0.2; 0.4]$, $\sigma_8 \in [0.6; 1.0]$) and $p(\boldsymbol{\pi}) = 0$ else. In our case we calculate $\hat{\boldsymbol{\xi}}$ from P_E via (3.51) and (3.52) assuming our fiducial cosmology; $\boldsymbol{\xi}_\pi$ is calculated similarly but its cosmological model varies according to the considered point in parameter space. The result of a likelihood analysis is usually summarized in contour plots. In a Bayesian approach, these likelihood contours represent so-called credible regions, i.e. a region in parameter space, where the true parameter is located with a probability of 68%, 95%, 99,9%, respectively. In addition, we quantify the size of these credible regions through the values of q, which we explained in Sect. 4.2. In these analyses, we consider the same parameter space, similar priors, similar $\hat{\boldsymbol{\xi}}$ and $\boldsymbol{\xi}_\pi$, only the covariance in (4.11) is changed. The left panel of Fig. 7.3 shows the 95%-credible intervals when choosing $\Omega_m = 0.2$ and $\sigma_8 = 0.6$ (solid), $\Omega_m = 0.4$, and $\sigma_8 = 1.0$ (dotted) as a model for the covariance matrix. We compare these to the (dashed) case when the covariance is calculated from the fiducial model ($\Omega_m = 0.25$, $\sigma_8 = 0.9$). These examples illustrate that assuming different cosmologies in the covariance can significantly broaden or narrow the likelihood contours. As expected from the foregoing analysis of the inverse covariance traces (Sect. 7.2) the contours broaden for increasing Ω_m and σ_8.

Without any information which cosmology to choose in our covariance matrix, it is reasonable to include prior information coming from other cosmological probes into our covariance cosmology. The middle panel of Fig. 7.3 shows the 95% credible intervals when calculating the covariance from the minimum, mean, and maximum values of the 68% confidence region of the recent WMAP 5-years analysis (Komatsu et al. 2008). Compared to the left panel the deviation
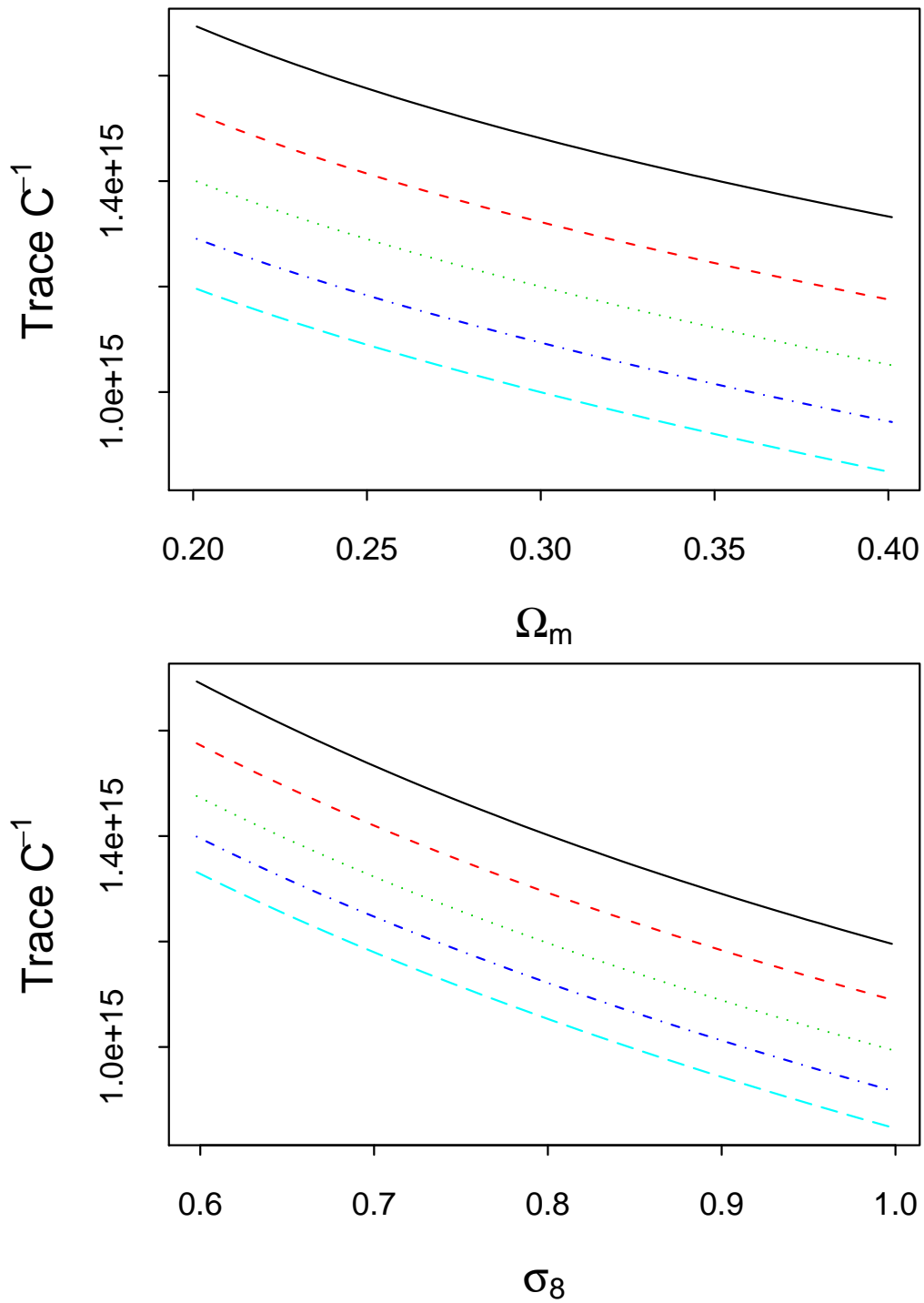
Figure 7.2: *The trace of the inverse covariance matrix* $\mathbf{C}^{-1}$ *depending on* Ω_m *(top), the individual lines in each figure correspond to (from top to bottom)* $\sigma_8 = [0.6, 0.7, 0.8, 0.9, 1.0]$. *The lower panel shows the dependence on* σ_8, *the individual lines corresponding to (from top to bottom)* $\Omega_m = [0.2, 0.25, 0.3, 0.35, 0.4]$.
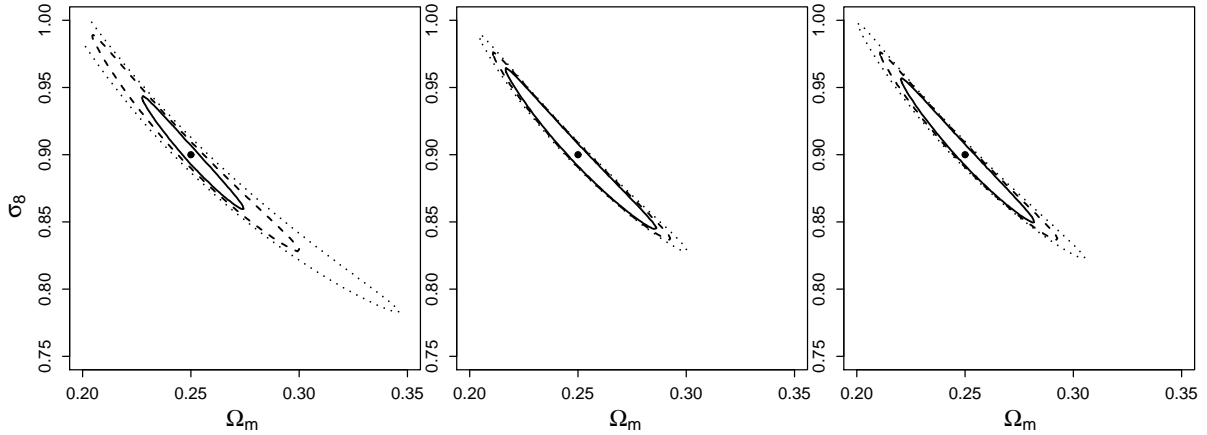
Figure 7.3: *The 95%-credible intervals obtained from likelihood analyses with different cosmological models assumed in their covariance matrix. The left panel corresponds to the following covariance parameters: $\Omega_m = 0.2$, $\sigma_8 = 0.6$ (solid), $\Omega_m = 0.25$, $\sigma_8 = 0.9$ (dashed), and $\Omega_m = 0.4$, $\sigma_8 = 1.0$ (dotted). The middle panel shows the deviation which occurs when restricting the range of possible covariance models to the 68% confidence interval of the WMAP5 analysis, i.e. $\Omega_m = 0.237$, $\sigma_8 = 0.74$ (solid), $\Omega_m = 0.259$, $\sigma_8 = 0.796$ (dashed), and $\Omega_m = 0.274$, $\sigma_8 = 0.85$ (dotted). The right panel shows the same analysis but for the 95% confidence interval of the WMAP5 analysis, i.e. $\Omega_m = 0.226$, $\sigma_8 = 0.70$ (solid), $\Omega_m = 0.237$, $\sigma_8 = 0.74$ (dashed), and $\Omega_m = 0.288$, $\sigma_8 = 0.885$ (dotted).*

of the contours reduces significantly, nevertheless it is still noticeable and cannot be neglected in a precision cosmology analysis. Similarly, the right panel shows the impact of the CDC-effect when calculating the covariance from parameters within the 95% confidence region of the recent WMAP5 analysis. For a better comparison we calculate the values of q (Sect. 4.2) for all contour plots and summarize them in Table 7.1. Restricting the possible cosmologies for the covariance to the 68% contour region of the WMAP5 analysis, the values of q deviate by a factor of ≈ 1.84. This factor increases to ≈ 2.76 when considering the minimum and maximum values of the 95% confidence region of the WMAP5 constraints. In Fig. 7.4 we show the values of q for all 2500 likelihood analyses depending on Ω_m (top) and σ_8 (bottom). Similar to the parameter dependence of the inverse covariances in Sect. 7.2, the strength of the CDC-effect, i.e. the gradient of the curves in Fig. 7.4, depends on the considered point in parameter space. At the fiducial model we calculate $(\partial q/\partial \Omega_m)_{\text{fid}} = 7.5$, whereas in case of σ_8 we find $(\partial q/\partial \sigma_8)_{\text{fid}} = 3.5$.

7.3.2 Impact of survey parameters on the CDC-effect

In the last section we have shown that the CDC-effect non-negligibly affects the likelihood contours. However, we only quantify this for one specific set of survey parameters. In this section we examine how the impact of the CDC-effect on likelihood contours depends on survey parameters, namely survey size A, ellipticity dispersion σ_ϵ, and number density of source galaxies $\bar{n}$, where in case of the latter two only the combination $\sigma_\epsilon^2/\bar{n}$ is of interest. We perform likelihood analyses
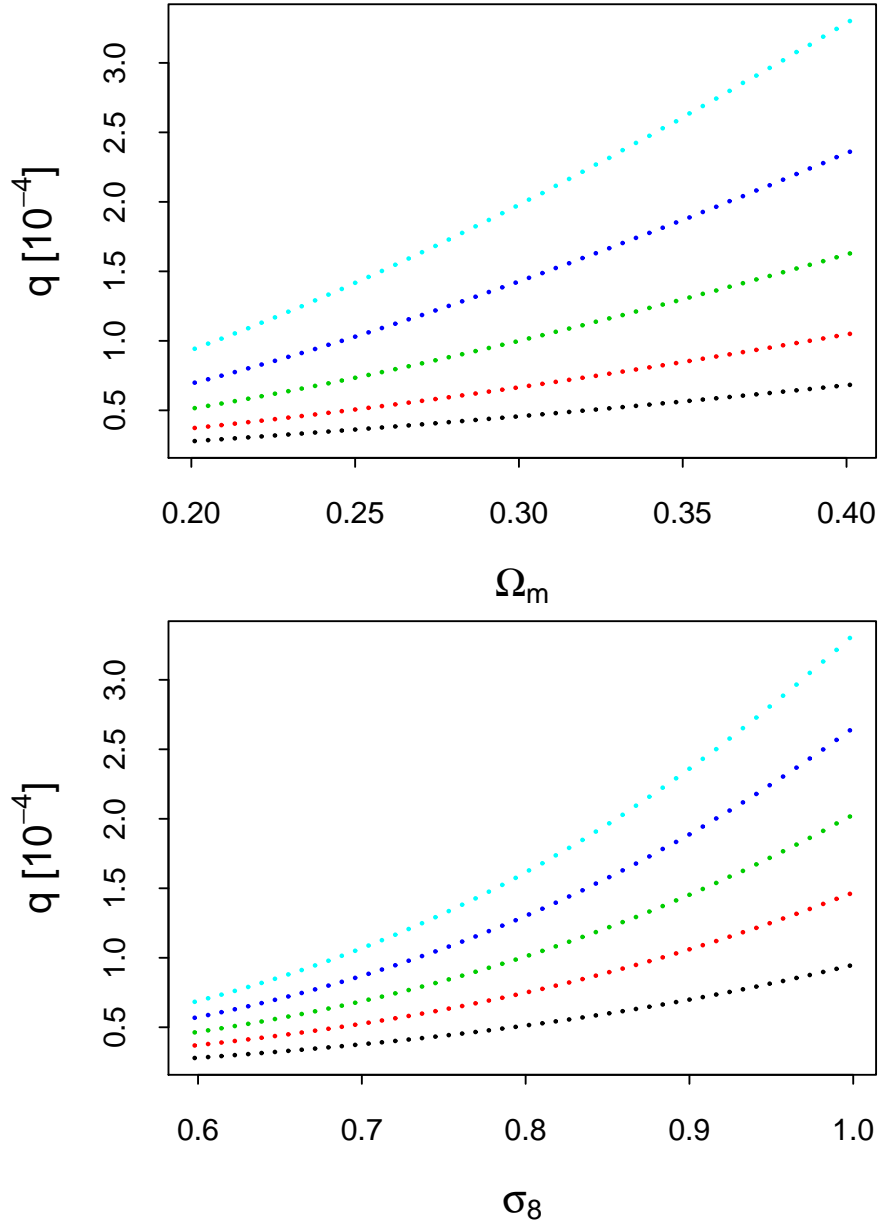
Figure 7.4: *The values of q depending on* Ω_m *(top), the individual lines in each figure correspond to (from top to bottom)* $\sigma_8 = [0.6, 0.7, 0.8, 0.9, 1.0]$. *The lower panel shows the dependence on* σ_8, *the individual lines corresponding to (from top to bottom)* $\Omega_m = [0.2, 0.25, 0.3, 0.35, 0.4]$.
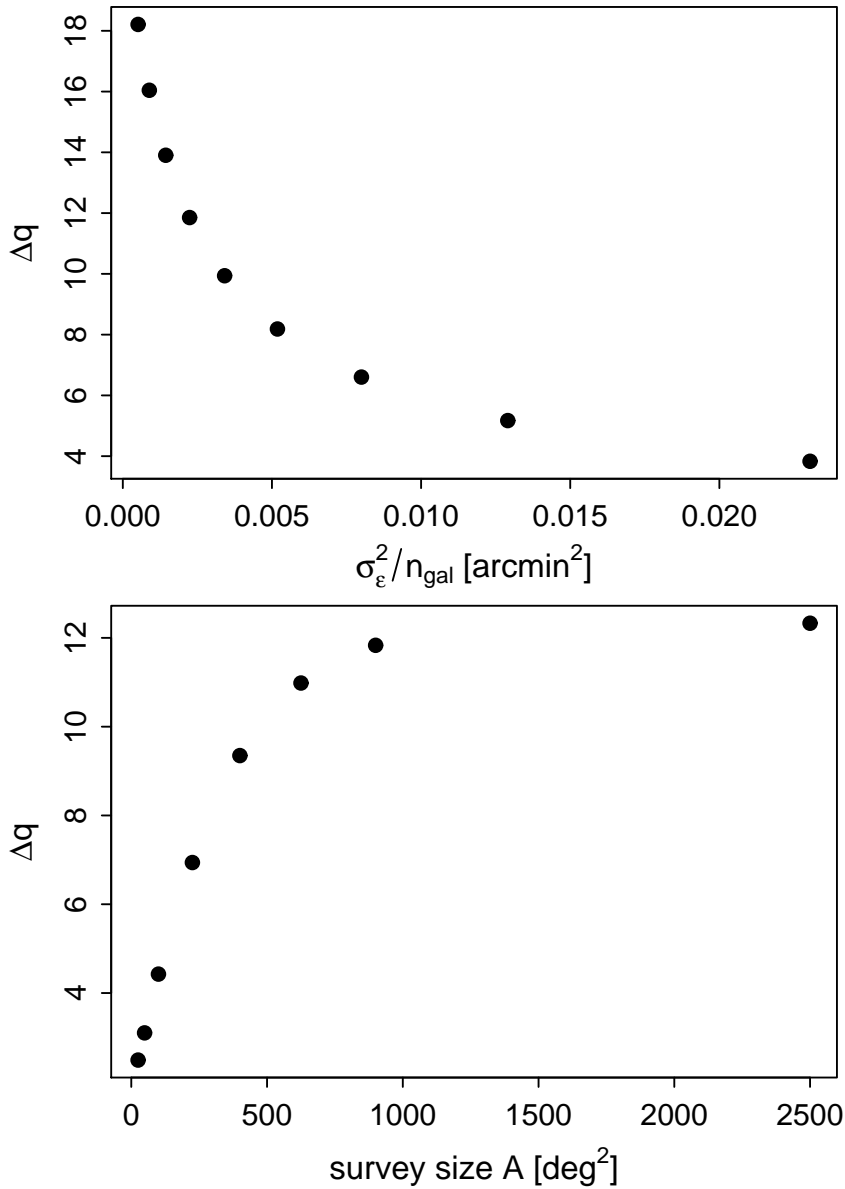
Figure 7.5: *This figure illustrates the strength of the CDC-effect depending on the ratio $\sigma_\epsilon^2/\bar{n}$ (upper panel) and depending on the survey size A (lower panel). We quantify this strength as the ratio $\Delta q = q(\boldsymbol{\pi}_{\max})/q(\boldsymbol{\pi}_{\min})$, where $q(\boldsymbol{\pi}_{\max})$ is obtained when assuming the maximum parameter values of Ω_m and σ_8 in the calculation of the covariance (i.e. $\boldsymbol{\pi}_{\max} = (\Omega_m = 0.4, \sigma_8 = 1.0)$). Correspondingly, $q(\boldsymbol{\pi}_{\min})$ is obtained when assuming $\boldsymbol{\pi}_{\min} = (\Omega_m = 0.2, \sigma_8 = 0.6)$.*

Table 7.1: *Values of q for different covariance models*

parameters used for the covariance	$q \, [10^{-4}]$
$\Omega_{\rm m} = 0.25, \sigma_8 = 0.9$	1.03
$\Omega_{\rm m} = 0.2, \sigma_8 = 0.6$	0.28
$\Omega_{\rm m} = 0.4, \sigma_8 = 1.0$	3.30
$\Omega_{\rm m} = 0.259, \sigma_8 = 0.796$ (WMAP5 68 % CL mean)	0.75
$\Omega_{\rm m} = 0.237, \sigma_8 = 0.740$ (WMAP5 68 % CL min)	0.56
$\Omega_{\rm m} = 0.274, \sigma_8 = 0.850$ (WMAP5 68 % CL max)	1.02
$\Omega_{\rm m} = 0.226, \sigma_8 = 0.700$ (WMAP5 95 % CL min)	0.45
$\Omega_{\rm m} = 0.288, \sigma_8 = 0.885$ (WMAP5 95 % CL max)	1.24

for 9 different combinations of $\sigma_\epsilon^2/\bar{n}$ and 8 different survey sizes. The strength of the CDC-effect is quantified by the ratio of maximum to minimum value of q, which occur within the considered range of $\Omega_{\rm m}$ and σ_8, we define $\Delta q = q(\boldsymbol{\pi}_{\rm max})/q(\boldsymbol{\pi}_{\rm min})$. The minimum q is obtained when choosing the minimum parameter set in the calculation of the covariance, i.e. $\boldsymbol{\pi}_{\rm min} = (\Omega_{\rm m} = 0.2, \sigma_8 = 0.6)$. Correspondingly, choosing the maximum parameter set $\boldsymbol{\pi}_{\rm max} = (\Omega_{\rm m} = 0.4, \sigma_8 = 1.0)$ results in the maximal q. The values of q represent the size of credible intervals, hence Δq can be interpreted as their ratio.

Unfortunately, it is not possible to derive an analytical expression for the relation between Δq and the survey parameters. From (4.6) - (4.8) we see that the individual covariance terms scale differently with $\sigma_\epsilon^2/\bar{n}$. This already prohibits an analytically derived relation between Δq and $\sigma_\epsilon^2/\bar{n}$. Considering the survey size A, (4.6) - (4.8) imply that the total covariance scales with $1/A$. When comparing two (inverse) covariances with different cosmologies by taking their ratio, the survey size cancels, suggesting the strength of CDC-effect to be independent of A. However, when considering the likelihood, the inverse covariance enters in the exponent, furthermore the values of q are an integral over the posterior likelihood. This non-linearity in the inverse covariance causes that the strength of the CDC-effect varies with the survey size. An analytic expression of this dependence cannot be derived, for similar reasons as for the case of $\sigma_\epsilon^2/\bar{n}$. We therefore calculate Δq depending on the survey parameters numerically.

The upper panel of Fig. 7.5 shows $\Delta q = q(\boldsymbol{\pi}_{\rm max})/q(\boldsymbol{\pi}_{\rm min})$ as a function of $\sigma_\epsilon^2/\bar{n}$. The ratio Δq changes from 4 to 18 over the considered interval of $\sigma_\epsilon^2/\bar{n}$. When increasing the survey size A (Fig. 7.5 , lower panel), we find that the impact of the CDC-effect increases from $\Delta q = 2.5$ (for a 25 deg^2 survey) up to $\Delta q = 12.3$ (for a 2500 deg^2 survey). Note that the size of the likelihood contours, hence the values of q themselves, decrease with decreasing $\sigma_\epsilon^2/\bar{n}$ and increasing A. In contrast, Δq increases with decreasing $\sigma_\epsilon^2/\bar{n}$ and increasing A. Hence relatively, the CDC-effect becomes more important when increasing the survey size or when decreasing the ratio $\sigma_\epsilon^2/\bar{n}$.

7.4 Likelihood analysis with a model dependent covariance

7.4.1 Adaptive covariance matrix

For a given cosmological model we can calculate the covariance directly from (4.3) - (4.8). This enables us to perform a likelihood analysis, where the covariance is calculated individually for every point in parameter space. We denote this parameter dependent covariance as $\mathbf{C}_\pi$ and rewrite the likelihood (4.11) as

$$p(\boldsymbol{\xi}|\boldsymbol{\pi}) = \frac{\exp\left[-\frac{1}{2}\left((\boldsymbol{\xi}_\pi - \hat{\boldsymbol{\xi}})^{\mathrm{t}}\, \mathbf{C}_\pi^{-1}\,(\boldsymbol{\xi}_\pi - \hat{\boldsymbol{\xi}})\right)\right]}{(2\pi)^{d/2}\,|\mathbf{C}_\pi|^{\frac{1}{2}}}\,. \tag{7.13}$$

Compared to the case of a constant covariance, there are two main differences. First, the covariance in the exponential term of (7.13) changes according to the considered point in parameter space. Second, $|\mathbf{C}_\pi|^{\frac{1}{2}}$ is now parameter dependent, therefore the determinant no longer cancels with a similar term in the evidence. As a consequence, the posterior likelihood does not only depend on the exponential terms, which basically compare $\boldsymbol{\xi}_\pi$ and $\hat{\boldsymbol{\xi}}$, but it is also affected by the determinants of the covariance matrices, more precisely by their behavior in parameter space. In the following we quantify the impact of the determinant term.

The upper left panel in Fig. 7.6 shows the likelihood contours for a 84 deg^2 survey, where the posterior probability is calculated via the new likelihood (7.13). For comparison, the right panel shows the likelihood contours when neglecting the parameter dependence in the determinant terms, hence considering a parameter dependent covariance only in the exponential terms. One clearly sees that the determinant terms shift the likelihood contours and cause a difference between the best-fit value and the fiducial model. In order to explain this shift we overlay the right panels of Fig. 7.6 with the contours of constant $|\mathbf{C}_\pi|^{-1/2}$ (for numerical reasons we plot $\ln|\mathbf{C}_\pi|^{-1/2}$). We see, that the covariance determinant is a monotonic function of Ω_{m} and σ_8; it decreases with increasing Ω_{m} or σ_8. Hence, $|\mathbf{C}_\pi|^{-\frac{1}{2}}$ induces a parameter-dependent weighting, which increases the likelihood at small Ω_{m} and σ_8 and vice versa suppresses large Ω_{m} and σ_8.

In general, the exponential term dominates the likelihood, $|\mathbf{C}_\pi|^{-1/2}$ only has significant impact on parameter regions where the exponential hardly changes. For the highly degenerate case of Ω_{m} and σ_8, this applies to curves where $\sigma_8 \approx const \times \Omega_{\mathrm{m}}^{-0.6}$. Compared to these curves, the contours of constant $|\mathbf{C}_\pi|^{-1/2}$ are slightly rotated, which allows for different values of the latter in regions where the exponential term is constant. As a result, the likelihood contours in the left panel are shifted and stretched towards regions of larger $|\mathbf{C}_\pi|^{-1/2}$ compared to the right panel. Note that for a different parameter combination this bias might not cause such a large shift of the best-fit value.

The second row of Fig. 7.6 shows the same analysis but for a 900 deg^2 survey. Comparing the left and right panel, we see that the likelihood contours are, similar to the 84 deg^2 survey, shifted and stretched towards regions of larger $|\mathbf{C}_\pi|^{-1/2}$. However, the effect is hardly noticeable and the bias of the best-fit value has basically vanished. This can be explained when looking at the

expression of the posterior likelihood

$$p(\boldsymbol{\pi}|\boldsymbol{\xi}) = \frac{\exp\left[-\frac{1}{2}\left((\boldsymbol{\xi}_\pi - \hat{\boldsymbol{\xi}})^{\mathrm{t}}\,\mathbf{C}_\pi^{-1}\,(\boldsymbol{\xi}_\pi - \hat{\boldsymbol{\xi}})\right)\right]}{\int \mathrm{d}\boldsymbol{\pi}'|\mathbf{C}_{\pi'}^{-1}\,\mathbf{C}_\pi|^{\frac{1}{2}}\exp\left[-\frac{1}{2}\left((\boldsymbol{\xi}_{\pi'} - \hat{\boldsymbol{\xi}})^{\mathrm{t}}\,\mathbf{C}_{\pi'}^{-1}\,(\boldsymbol{\xi}_{\pi'} - \hat{\boldsymbol{\xi}})\right)\right]} \; . \tag{7.14}$$

Compared to the case of a constant covariance the above expression has an additional factor in the denominator, i.e. $|\mathbf{C}_\pi \mathbf{C}_{\pi'}^{-1}|^{1/2}$. Note, that this factor is independent of the survey size A, whereas the importance of the exponential term increases with increasing A. As a result, the cosmology dependence of the covariance determinant becomes negligible for sufficiently large surveys.

7.4.2 Fisher matrix analysis

We expect tighter constraints on cosmological parameters if the cosmology dependence of both, mean data vector and covariance matrix, is incorporated into the likelihood analysis, instead of only using the mean data vector (Tegmark et al. 1997). The Fisher information matrix can be used to illustrate this effect. In Sect. 4.3 we explained the concept of the Fisher matrix in detail; recall that for the case that $p(\boldsymbol{\xi}|\boldsymbol{\pi})$ is Gaussian one can directly express the Fisher matrix in terms of the mean data vector and the data covariance matrix (e.g. Tegmark et al. 1997)

$$\mathbf{F}_{ij} = \frac{1}{2}\mathrm{tr}\left[\mathbf{C}^{-1}\mathbf{C}_{,i}\mathbf{C}^{-1}\mathbf{C}_{,j} + \mathbf{C}^{-1}\mathbf{M}_{ij}\right] , \tag{7.15}$$

where $\mathbf{C}_{,i} \equiv \partial\mathbf{C}/\partial\pi_i$ denotes the derivative of the covariance matrix with respect to the i-th component of the parameter vector and $\mathbf{M}_{ij} \equiv \boldsymbol{\xi}_{,i}\boldsymbol{\xi}_{,j}^{\mathrm{t}} + \boldsymbol{\xi}_{,j}\boldsymbol{\xi}_{,i}^{\mathrm{t}}$. The first term of (7.15) vanishes in case the covariance matrix is constant in parameter space, the second term vanishes in case of a constant mean. For cosmic shear we have seen that neither the mean data vector, nor the covariance matrix are independent of cosmological parameters; hence, when calculating the Fisher matrix both terms must be taken into account. Recall that $\mathbf{C} \propto 1/A$, which also holds for the derivatives $\mathbf{C}_{,i}$, hence the first term is independent of the survey size. The second term increases proportional to the survey volume, therefore the information gain on cosmological parameters, through incorporating the cosmology dependence of covariances, becomes less important for large surveys (see also Kilbinger & Munshi 2006, for a similar analysis).

Figure 7.7 shows the results of the Fisher matrix analysis for two different survey sizes (84 deg^2 on the left and 900 deg^2 on the right). As expected, the left panel (smaller survey) shows a small improvement, which vanishes completely in case of the larger survey (right panel). Nevertheless, one should keep in mind that we only consider Gaussian covariances. The cosmology dependence of the covariance becomes larger for the case of non-Gaussian covariances for the following reason. Non-Gaussianity increases the cosmic variance term, in particular it becomes important on small scales, which are still dominated by shot noise in the pure Gaussian case. As the CDC-effect mainly results from the cosmic variance term, its strength also increases in the non-Gaussian case. A stronger dependence of the covariance on parameters enlarges the first term in (7.15), which implies that for the case of truly non-Gaussian covariances the improvement on parameter constraints is more significant than shown in Fig. 7.7.
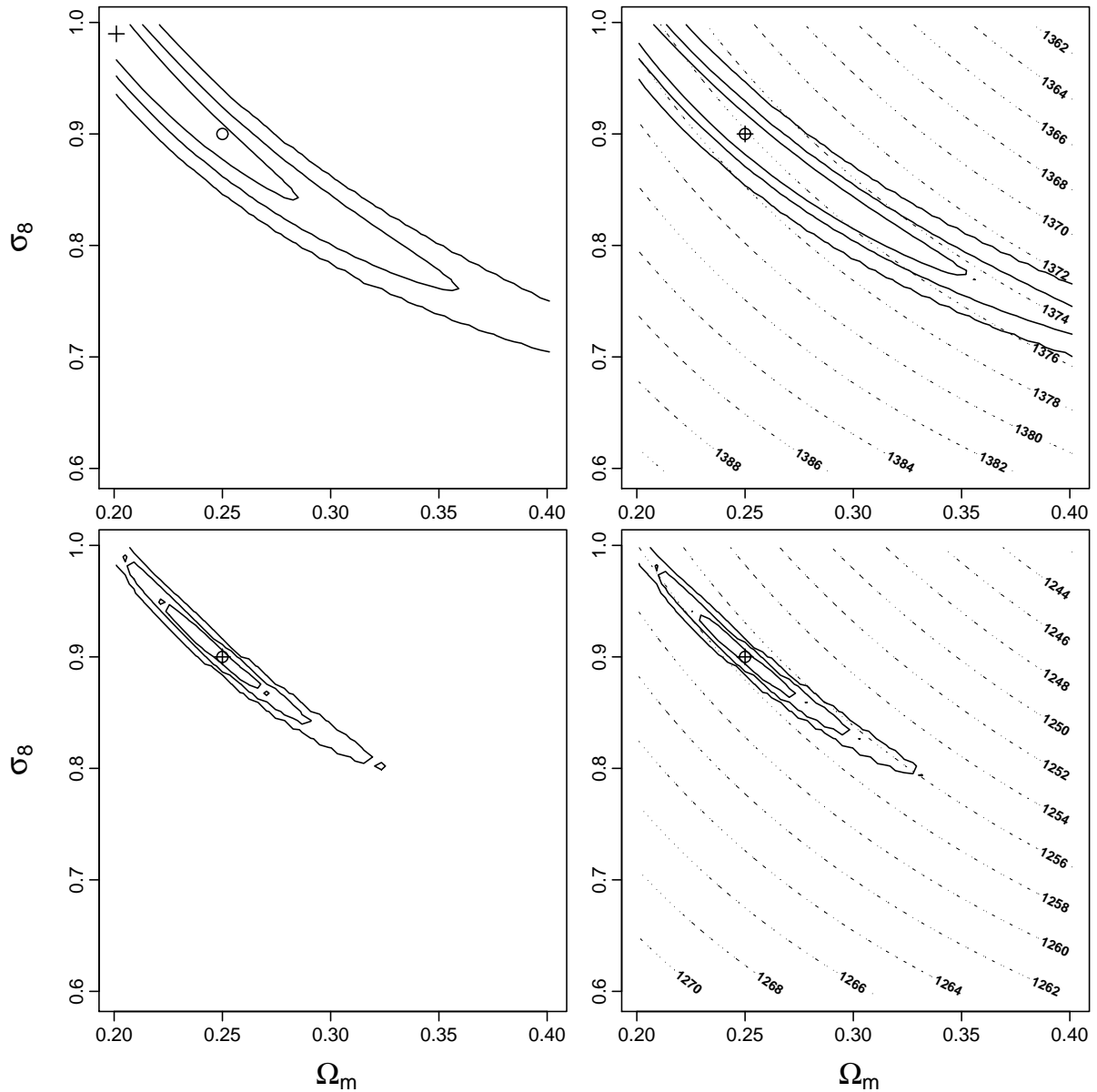
Figure 7.6: *The left plots shows the likelihood contours when using a model- dependent co-variance, more explicitly, when calculating the posterior from (7.13). The cross illustrates the best-fit value, whereas the circle indicates our fiducial model. The panels on the r.h.s. show the likelihood contours obtained when neglecting the determinant-terms (7.13). The dotted contours visualize regions of constant $\ln |\mathbf{C}_{\pi'}|^{-1/2}$. The likelihood contours in the upper row correspond to a survey size of 84 deg^2, whereas the lower panels correspond to $A = 900$ deg^2.*
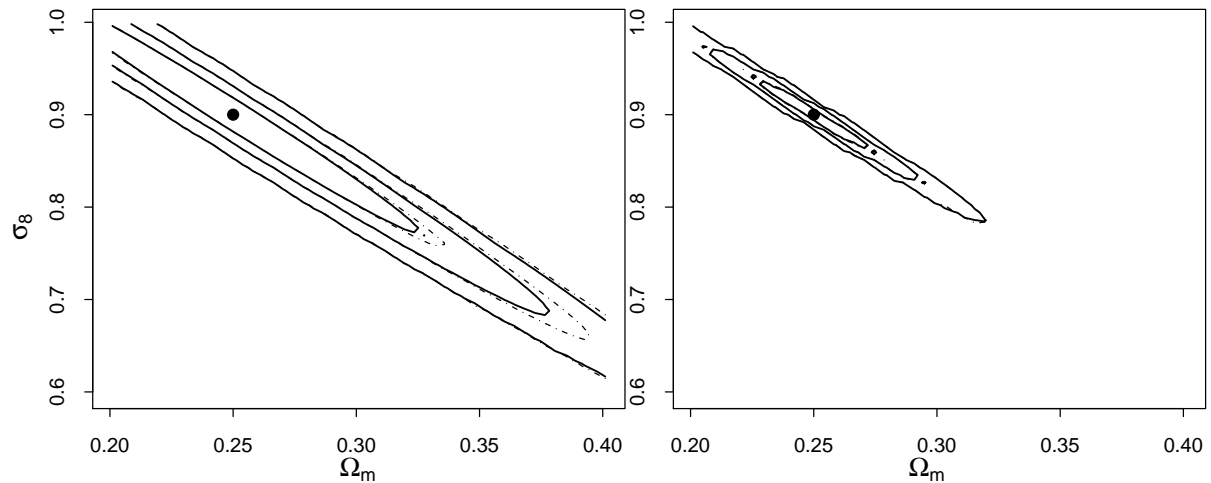
Figure 7.7: *Likelihood contours from a Fisher matrix analysis for a 84 deg^2 survey (left), and for a 900 deg^2 (right). The solid lines correspond to the case when the Fisher matrix is calculated from both terms in (7.15); the dashed lines correspond to the case when the parameter dependence of the covariances, i.e. the first term in (7.15), is neglected. The filled circle indicates the fiducial model at which the Fisher matrix was calculated. Note, that in the right panel dashed and solid contours are identical*

Table 7.2: *The ML-parameter sets which occur when choosing different starting cosmologies in the iterative likelihood analysis.*

| step	run1		run2		run3		run4		run5	
	Ω_m	σ_8	Ω_m	σ_8	Ω_m	σ_8	Ω_m	σ_8	Ω_m	σ_8
π_{start}	0.20	0.60	0.237	0.740	0.250	0.90	0.274	0.850	0.40	1.0
π_{ML} 1	0.254	0.892	0.254	0.892	0.245	0.914	0.259	0.884	0.277	0.858
π_{ML} 2	0.260	0.882	0.260	0.882	0.245	0.914	0.260	0.882	0.259	0.884
π_{ML} 3	0.260	0.882	0.260	0.882	converged		0.260	0.882	0.260	0.882
π_{ML} 4	converged		converged		converged		converged		0.260	0.882

7.4.3   Iterative likelihood analysis

In Sect. 7.4.1 we have introduced the adaptive covariance, which is a proper way to incorporate cosmology-dependent covariances into a likelihood analysis. Its disadvantage is the large computational effort, which is high already for Gaussian covariances. In order to account for non-Gaussianity, one must employ ray-tracing covariances derived from many numerical simulations with different underlying cosmologies. In a multi-dimensional parameter space, this is clearly unfeasible with today's computer power.

In this section we quantify the impact on likelihood contours when using non-Gaussian instead of Gaussian covariances. We use a ray-tracing covariance taken from the Millennium simulation (Hilbert et al. 2008), neglect the CDC-effect and approximate the covariance to be constant in

parameter space. The error in the posterior likelihood caused by this approximation increases with increasing distance to the cosmology of the ray-tracing simulation. As we are mainly interested in regions around the maximum likelihood parameter set, $\boldsymbol{\pi}_{\mathrm{ML}}$, this suggests the following strategy for a likelihood analysis. First, perform an iterative likelihood analysis using Gaussian covariances in order to derive $\boldsymbol{\pi}_{\mathrm{ML}}$. Then, start a numerical simulation with this cosmology, derive a ray-tracing covariance, and perform the final likelihood analysis. This ansatz minimizes the errors due to the CDC-effect in the region of interest and additionally incorporates non-Gaussianity.

In order to derive $\boldsymbol{\pi}_{\mathrm{ML}}$ iteratively, we start from an arbitrary cosmology, calculate a Gaussian covariance matrix therefrom using (4.3) - (4.8), and perform a likelihood analysis. Throughout this first iteration step the covariance matrix is kept constant. In the second step we choose the ML-parameter set of the first analysis as the underlying cosmology for the new covariance matrix, and again perform a likelihood analysis. We continue this iteration process until the ML-parameter set converges.

The main difficulty of this ansatz is that the choice of the starting cosmology might influence the final ML-parameter estimate and therefore also the final covariance. In order to check for this, we take the noise of a ray-tracing data vector, add it to our fiducial data vector and thereby simulate measurement uncertainties in the latter. When performing the analysis without noise the iteration converges after one step, as the model data vector (at the fiducial model) exactly fits the fiducial data vector, $\boldsymbol{\xi}_{\boldsymbol{\pi}_{\mathrm{fid}}} = \hat{\boldsymbol{\xi}}$. Table 7.2 shows the results for 5 iterative likelihood analyses, each starting from a different cosmology in the covariance. We see that all 5 runs converge quickly, 4 of them to the same cosmology. Only the run which started from the fiducial model deviates from the others. Although the suggested $\boldsymbol{\pi}_{\mathrm{ML}}$ are close to $\boldsymbol{\pi}_{\mathrm{fid}}$, we note that none of the runs converges to the fiducial model. This implies that the starting cosmology can bias the final outcome of the iterative likelihood analysis and can shift the ML-estimate. In general, such a bias occurs if the function $\boldsymbol{\xi}_{\boldsymbol{\pi}} - \hat{\boldsymbol{\xi}}$ does not fall off steeply enough around the ML-parameter set, which especially applies to higher-dimensional likelihood analyses.

Our iterative pre-analysis has converged to $\Omega_{\mathrm{m}} = 0.26$, $\sigma_8 = 0.882$, however we "only" have a ray-tracing simulation with $\Omega_{\mathrm{m}} = 0.25$, $\sigma_8 = 0.9$ available. Figure 7.8 shows the result of our likelihood analysis, when using the ray-tracing covariance of the Millennium simulation (left panel). Compared to a likelihood analysis using a Gaussian covariance (right panel), the contours broaden significantly; q increases from 0.44×10^{-4} in the Gaussian to 0.78×10^{-4} in the non-Gaussian case. Note that the value of q in the Gaussian case does not correspond to that in Table 7.1, because we use different survey parameters (here, $\sigma_\epsilon = 0.3$, $\bar{n} = 15/\mathrm{arcmin}^2$) and a different data vector (here, 30 logarithmic bins from 0.2-130 arcmin) in order to exactly match the corresponding parameters of the ray-tracing covariance.

The impact of non-Gaussianity depends on the scales probed by the data vector. In our case 20 bins are below 10 arcmin, therefore the impact is relatively high. Choosing linear bins or probing higher ϑ reduces the difference to the Gaussian case. For the data vector considered here, this difference is of the same order as the impact of the CDC-effect we described in Sect. 7.3.1. However, the strength of the latter will most likely increase for non-Gaussian covariances, as we explained at the end of the last section.
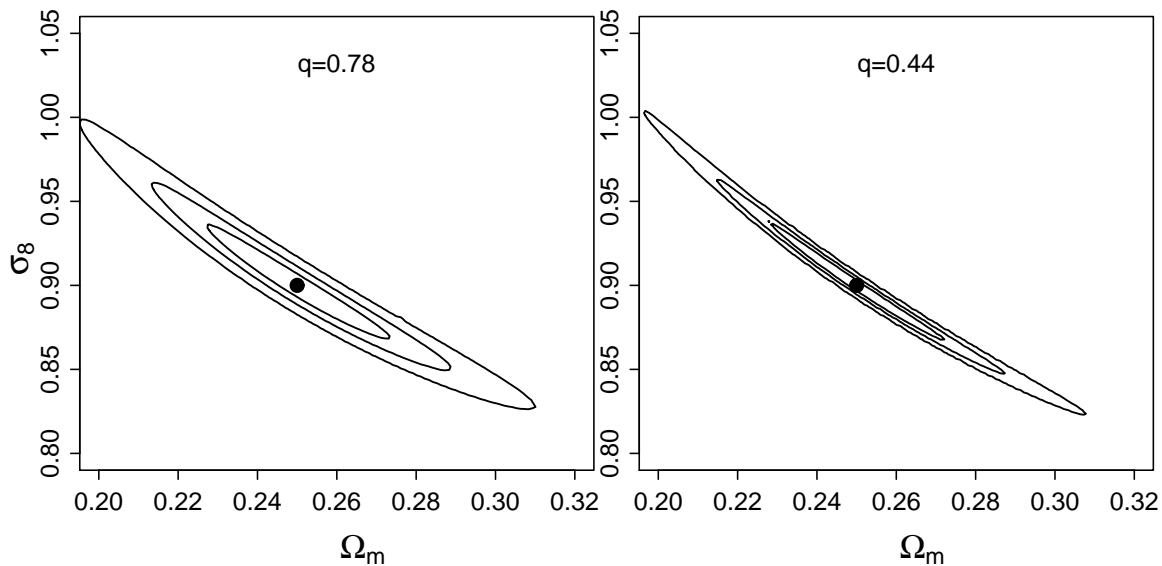
Figure 7.8: *The likelihood contours when using a ray-tracing covariance derived from the Millennium Simulation via field-to-field variation (left panel), compared to the case of a Gaussian covariance (right panel). Although the original size of each field is only 16* $\deg^2$*, we extrapolated the covariance to a 900* $\deg^2$ *survey. The values of q are given in units of* 10^{-4}*.*

7.5　Concluding remarks on the cosmology dependence of covariances

An accurate likelihood analysis plays an essential role in future precision cosmology. We can only exploit the full potential of upcoming high quality data, if we use appropriate statistical methods. In this context the derivation of covariances is an important issue in order not to bias the parameter constraints.

In cosmic shear, there are several methods to derive covariances. First, one can calculate $\mathbf{C}$ analytically assuming a Gaussian shear field. This assumption breaks down on small angular scales (< 10 arcmin), where non-linearities of the matter density field start to become important. Second, covariances can be estimated from ray-tracing simulations. Although computationally more expensive, this method automatically accounts for the non-Gaussianity of the shear field. In both methods the covariance is calculated assuming a specific cosmology. In the first case, this cosmology enters in the power spectrum from which $\mathbf{C}$ is calculated, in the second case we estimate $\mathbf{C}$ from numerical simulations, which are also based on a given cosmology. Past cosmic shear data analyses approximate the covariance to be constant in parameter space, therefore assume that its underlying cosmology does not influence the result of a likelihood analysis significantly. In this chapter we have shown that the covariance matrix depends non-negligibly on its underlying cosmology and that this CDC-effect significantly influences the likelihood contours of parameter constraints. To prove this, we calculate 2500 Gaussian covariance matrices for var-

ious parameters of $\Omega_m \in [0.2; 0.4]$ and $\sigma_8 \in [0.6; 1.0]$; all other cosmological parameter are held fixed. Even a change of Ω_m and σ_8 within the WMAP5 68% confidence levels has a non-negligible impact on the likelihood contours. Here, the value of q deviates by a factor of 1.84 and this deviation increases to 2.76 if one considers the WMAP5 95% confidence levels. Furthermore, we show that the impact of the CDC-effect depends on survey parameters. Although the likelihood contours become smaller, relatively the CDC-effect becomes more important when increasing the survey size or when decreasing the ratio $\sigma_\epsilon^2/\bar{n}$. Therefore, a proper treatment becomes more important in the future, for large and deep surveys.

To take cosmology-dependent covariances into account we present two methods. First, we perform a likelihood analysis with an adaptive covariance matrix. Here, $\mathbf{C}$ is calculated individually for every point in parameter space, assuming the corresponding parameters as the underlying cosmology. For small surveys this method introduces a bias to the best-fit parameter set, which vanishes when going to larger survey sizes. A disadvantage of this approach is its computational costs. Using the analytic expression for Gaussian covariances is already time-consuming; using ray-tracing covariances to include the non-Gaussianity is not feasible with today's computer power. For the Gaussian case we present a fast and convenient scaling relation to derive covariances on a parameter grid. As a side-effect this approach enhances the constraints on cosmology, for the reason that we now incorporate two cosmology-dependent quantities into the likelihood analysis instead of only the mean data vector.

In a strict sense the second method does not account properly for the CDC-effect, however it minimizes the error around the maximum likelihood parameter set ($\boldsymbol{\pi}_{\mathrm{ML}}$). The method consists of two steps, first derive $\boldsymbol{\pi}_{\mathrm{ML}}$ through an iterative process, then derive a ray-tracing covariance with $\boldsymbol{\pi}_{\mathrm{ML}}$ as underlying cosmology and incorporate this in the final likelihood analysis. Here, the approximation of a constant covariance is made, however the error in the posterior probability is minimized in the region of interest; in addition, this ansatz incorporates non-Gaussianity which is non-negligible for future surveys. A drawback is the fact the the starting point of the iteration might bias $\boldsymbol{\pi}_{\mathrm{ML}}$. This must be checked carefully before employing this method, otherwise the approximation of a constant covariance fails around $\boldsymbol{\pi}_{\mathrm{ML}}$.

Chapter 8

Impact of non-Gaussianity and shot noise on parameter estimation

One of the most exciting discoveries in cosmology was made in 1998, namely that the expansion of the Universe accelerates. This acceleration suggests the existence of an energy density component which acts repulsively, so-called dark energy (see Sect. 2.1.6). As pointed out by the *Dark Energy Task Force* (Albrecht et al. 2006) and the *ESA-ESO Working Group on Fundamental Cosmology* (Peacock et al. 2006), cosmic shear is an ideal subject for future dark energy missions such as DES, JDEM, or Euclid. Therefore, I could not resist to carry out one project related to dark energy parameter estimation with cosmic shear.

At the end of the last chapter we illustrated the impact of non-Gaussianity on parameter constraints for the case of Ω_m and σ_8. In this chapter we extend the analysis to the dark energy parameters Ω_Λ, w_0, and w_a (see Sect. 2.1.6 for more details). In addition to non-Gaussianity we examine the impact of shot noise on the parameter constraints. The analysis presented in this chapter is an ongoing project with preliminary results which require more work in the future. We will indicate open questions and plans how to address them.

8.1 Comparison of P_κ derived from the Smith et al. (2003) fit-formula to P_κ from the Millennium simulation

In the last chapter we calculated the Gaussian covariances from the shear power spectrum. Therefore, we obtain P_δ via the fit-formula of Smith et al. (2003) and calculate P_κ via (3.44). We want to stress, that the accuracy of this fit-formula is an inherent limitation to the accuracy of covariances and data vectors derived therefrom. Figure 8.1 compares P_κ derived from the Millennium simulation to P_κ derived from the fit-formula of Smith et al. (2003) using the Millennium simulation cosmology. In the latter case we show the results when using the transfer function of Efstathiou et al. (1992) (solid) and a dewiggeled version of the transfer function of Eisenstein & Hu (1999) (dashed). Whereas the impact of the transfer function is comparatively small, the two P_κ calculated from the fit-formula significantly deviate from P_κ of the Millennium simulation when going above $\ell = 1000$ (see right panel of Fig. 8.1). For more details on this topic the

reader is referred to Hilbert et al. (2008). This deviation also occurs in the corresponding 2PCF
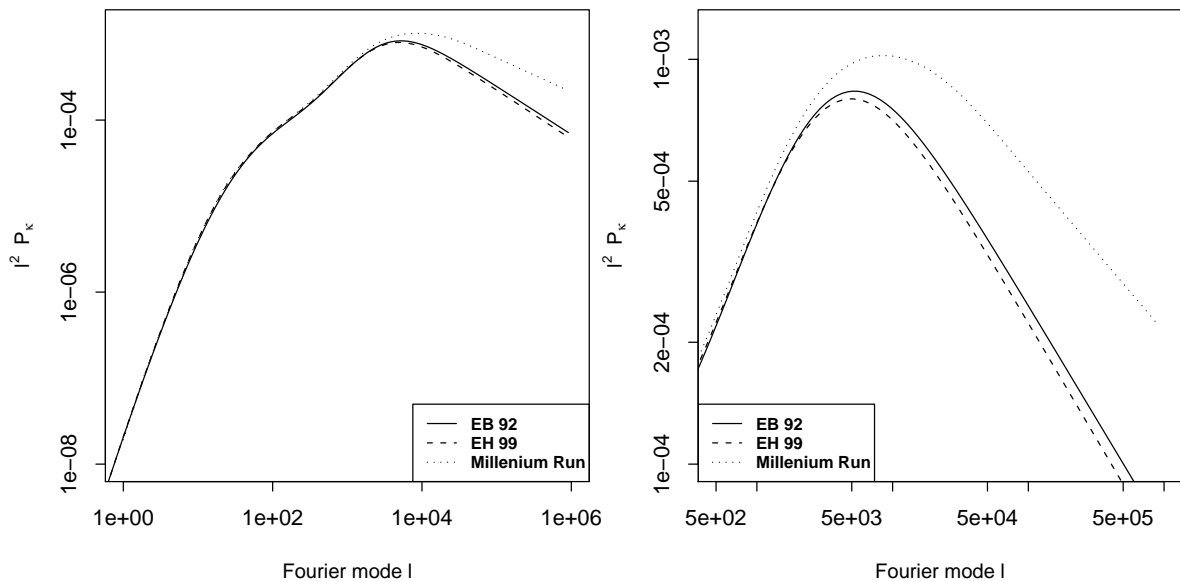

Figure 8.1: *This figure shows three shear power spectra; the two panels only differ in the considered range of Fourier modes. The dotted line corresponds to P_κ obtained from the Millennium Simulation directly, whereas the other two are calculated from the fitting formula of Smith et al. (2003), employing the transfer function of Efstathiou et al. (1992) (solid) and a dewiggeled version of the transfer function of Eisenstein & Hu (1999) (dashed).*

covariances calculated via (4.3) - (4.8). In Fig. 8.2 we compare the contour levels of these covariances; the left panel corresponds to the covariance obtained from the fitting formula, whereas the right was calculated from the Millennium simulation shear power spectrum. For given angular scales the values of the latter covariance are larger compared to those of the first. This effect is clearly noticeable in the upper panel ($\mathbf{C}_{++}$); it also exists for $\mathbf{C}_{--}$, $\mathbf{C}_{+-}$, but is less significant. For the rest of this chapter, I refer to Gaussian covariances as covariances calculated via (4.3) - (4.8) using P_κ of the Millennium simulation. Non-Gaussian covariances in this chapter are obtained via field-to-field variation of 128 independent ray-tracing realizations through the Millennium simulation.

8.2   Gaussian and non-Gaussian covariances

In Sect. 7.1 we outline three possibilities how to obtain non-Gaussian covariances. First, they can be estimated from the data directly via field-to-field variation, however this method is inappropriate for small surveys. Second, one can employ ray-tracing simulations, a method we used in the context of the iterative likelihood analysis in Sect. 7.4.3. Third, one can calculate a Gaussian covariance analytically using the power spectrum (Joachimi et al. 2008) or the 2PCF (Schneider et al. 2002a). To account for non-Gaussianity one then applies the calibration factor
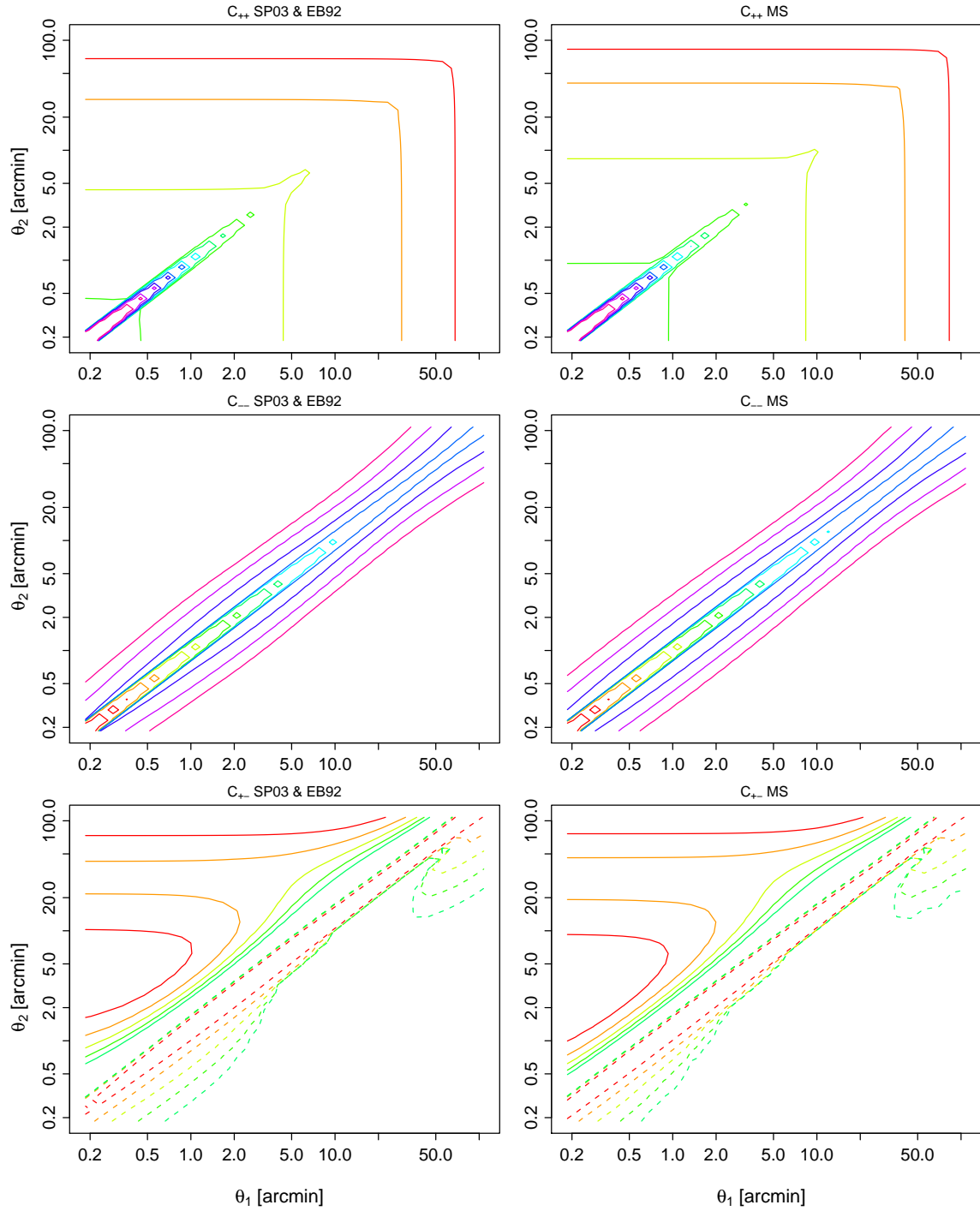
Figure 8.2: *This figure compares the contour lines of the 2PCF covariance matrices ($\mathbf{C}_{++}$ (top), $\mathbf{C}_{--}$ (middle), $\mathbf{C}_{+-}$ (bottom)) when using the fitting formula of Smith et al. (2003) (left panel) in the calculation (4.3) - (4.8) and when using the shear power spectrum of the Millennium Run (right panel). The contours levels are monotonically decreasing, with red marking the lowest absolute value in the upper row, and the highest absolute value for the lower two rows, respectively. The x-th contour level of left and right panels are similar, i.e. ($5.58 \times 10^{-11} \times 1.5^x$) for $\mathbf{C}_{++}$, $4.6 \times 10^{-8}/3.5^x$ for $\mathbf{C}_{--}$. In case of $\mathbf{C}_{+-}$ the contours correspond to $3.6 \times 10^{-8}/(1.2^x)$ for the positive values (solid) and $-4.7 \times 10^{-8}/5.0^x$ for the negative values (dashed).*

Table 8.1: *Comparison of my results for $F_{++}(\theta_1, \theta_2)$ to the analysis of Semboloni et al. (2007).*

Analysis	Functional form	α	β
This thesis	$\alpha(\theta_1\theta_2)^{-\beta}$	8.100	0.514
Semboloni et al. (2007)	$\alpha(\theta_1\theta_2)^{-\beta}$	16.48	0.56

of Semboloni et al. (2007). This factor was calibrated using ray-tracing simulations, but only for the case of $\mathbf{C}_{++}$. In the following, we repeat their analysis using the Millennium simulation. More precisely, we derive a non-Gaussian covariance matrix from the via field-to-field variation and compare it to a Gaussian covariance obtained from the shear power spectrum of the Millennium simulation. Besides the fact that we use the power spectrum approach to calculate the Gaussian 2PCF covariances whereas Semboloni et al. (2007) employ the 2PCF directly, there are two main differences between both analyses. First, we do not include a redshift dependence into the calibration factor, and second the cosmologies of the ray-tracing simulations are different. The calibration factor is defined as

$$F_{\pm\pm}(\theta_1, \theta_2) = \frac{C_{\pm\pm}^{\text{non-Gaussian}}(\theta_1, \theta_2)}{C_{\pm\pm}^{\text{Gaussian}}(\theta_1, \theta_2)} \; . \tag{8.1}$$

Note that in our derivation we consider a redshift distribution of source galaxies which has the functional form of (6.36) with parameters $\alpha = 0$, $\beta = 1.5$, $z_0 = 0.9$. The different survey sizes A (5.49 deg^2 of Semboloni et al. (2007) compared to 16 deg^2 here) should not influence the results on $F_{\pm\pm}(\theta_1, \theta_2)$ significantly, for the reason that $C^{\text{non-Gaussian}}$ and C^{Gaussian} should scale similarly with A.

Figure 8.3 shows the calibration factor for the diagonal terms of $\mathbf{C}_{++}$. The result we obtain for $F_{++}(\theta, \theta)$ is qualitatively similar to the results in Semboloni et al. (2007) (see Fig. 1 in their paper). We therefore choose the same parameterization as Semboloni et al. (2007); the results of the fit are given in Table 8.1. Although our results ($\alpha = 8.1$, $\beta = 0.51$) deviate from the corresponding parameters of Semboloni et al. (2007) ($\alpha = 16.48$, $\beta = 0.56$) this can be explained by the different cosmologies of the Millennium simulation and the numerical simulations used in Semboloni et al. (2007). In particular, their values for $\Omega_{\text{m}} = 0.3$ and $\sigma_8 = 1.0$ are higher compared to those of the Millennium simulation ($\Omega_{\text{m}} = 0.25$ and $\sigma_8 = 0.9$). This leads to a shear field which is more non-Gaussian compared to that of the Millennium simulation.

We also tried to extend the calibration factor to the $\mathbf{C}_{--}$ and $\mathbf{C}_{+-}$ sub-matrices of the 2PCF covariance, however this is more complicated. On small angular scales $F_{\pm-}(\theta, \theta)$ becomes very steep which prevents an easy functional form as used for $\mathbf{C}_{++}$. We already started to analyze this issue, however we did not succeed finding an appropriate functional form for the calibration factor. We plan to examine this in more detail in the future. As an improvement over the analytic calculation of the covariance via (4.3) - (4.8) one can simulate Gaussian realizations from the Millennium power spectrum and obtain a covariance via field-to-field variation of these realizations. This would exclude uncertainties in the analytic calculation.

As a next step, after deriving calibration factors of $\mathbf{C}_{--}$ and $\mathbf{C}_{+-}$, we plan to apply the correction factor to the adaptive likelihood formalism introduced in chapter 7. From Fig. 8.2 we see that the
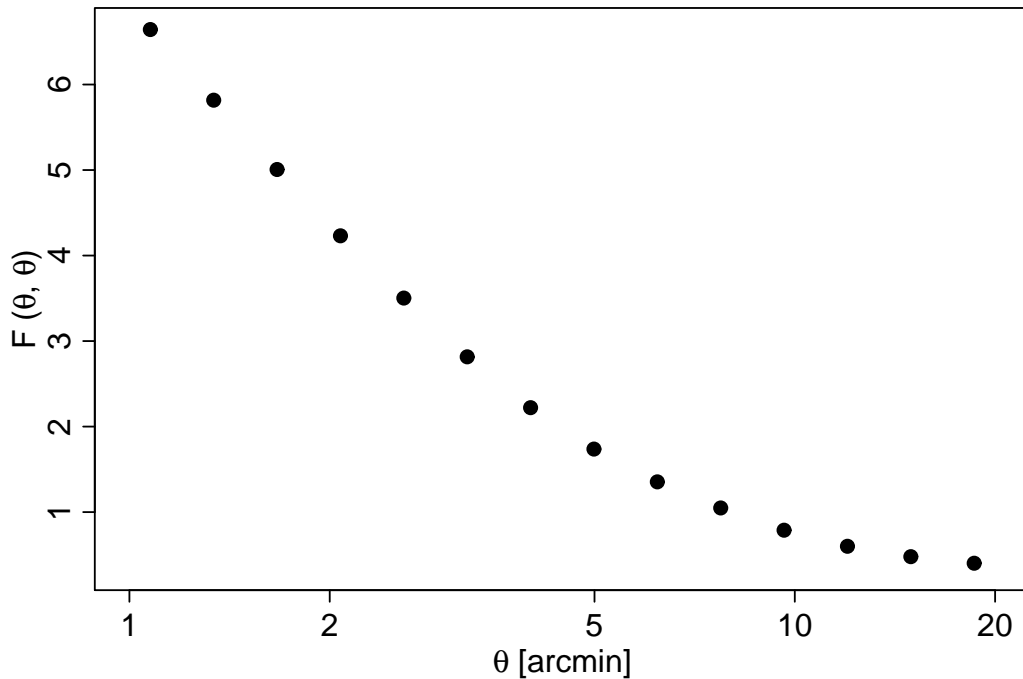
Figure 8.3: *This figure shows the non-Gaussian calibration factor for the diagonal terms of the 2PCF covariances* $\mathbf{C}_{++}$.

structure of a Gaussian covariance obtained through a power spectrum derived from Smith et al. (2003) is very similar to that derived from P_κ of the Millennium simulation. This suggests that applying the correction factor to Gaussian covariances calculated from an analytic model, is in principle a valid ansatz. However, one must keep in mind that the calibration factor is very likely to be cosmology-dependent. A comparison to the results of Semboloni et al. (2007) already showed significant deviation, which could be explained by the different underlying cosmologies of the ray-tracing simulations. Applying a constant calibration factor in the adaptive likelihood formalism is therefore probably not the best approach. In collaboration with AstroGrid[1] we have access to ray-tracing simulations for 200 cosmologies, more precisely 200 different combinations of Ω_m and σ_8. We intend to compare the non-Gaussian covariances obtained by applying a constant $F(\theta_1, \theta_2)$ to the case of deriving non-Gaussian covariances from these ray-tracing simulations.

---

[1]http://www.astrogrid.org/
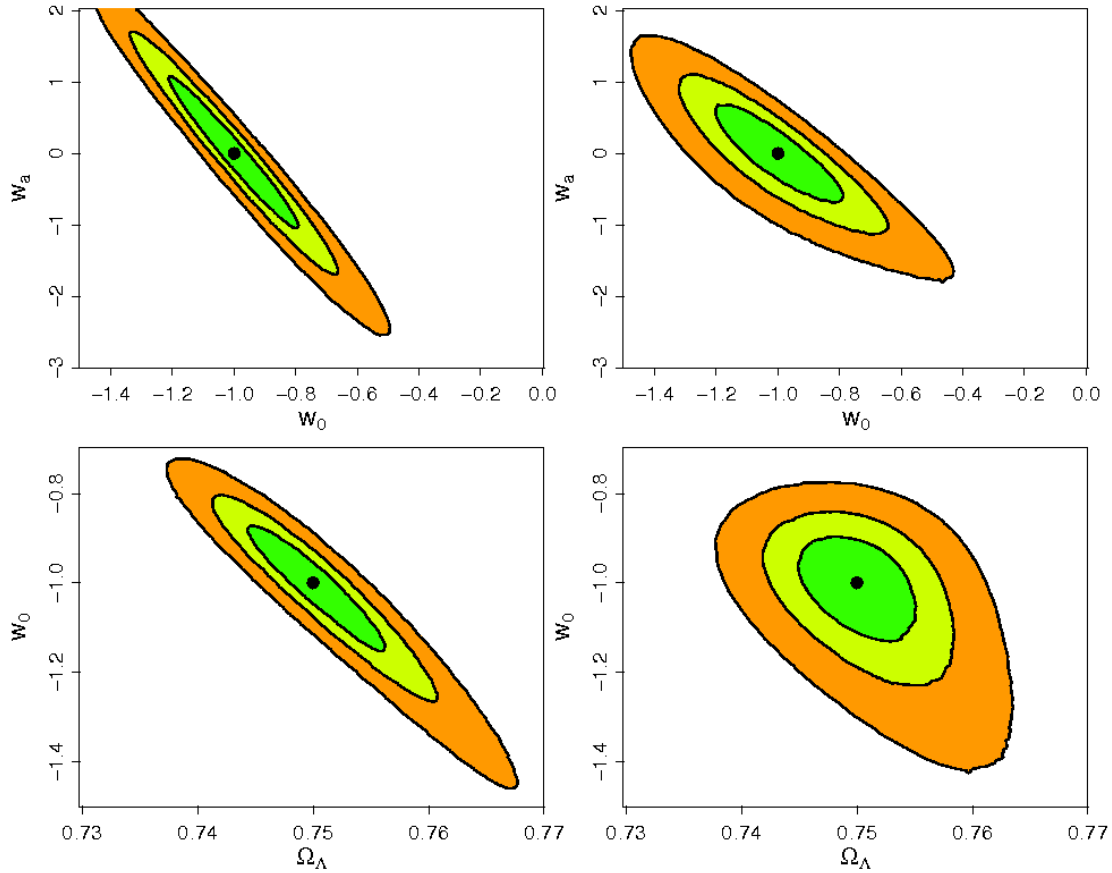
Figure 8.4: *This figure shows the impact of non-Gaussianity on the constraints of dark energy parameters, namely Ω_Λ vs. w_0 (bottom panels) and w_0 vs. w_a (top panels). The left panels correspond to the result when using Gaussian covariances and the right corresponds to non-Gaussian covariances.*

8.3   Impact of non-Gaussian covariances on parameter estimation

In this section we examine how non-Gaussianity affects the likelihood contours. For the case of Ω_m vs. σ_8 we already illustrated the impact in Fig. 7.8. In the following analysis we also examine dark energy parameters, namely Ω_Λ vs. w_0 and w_0 vs. w_a (see Fig. 8.4). All likelihood analyses in this chapter are performed assuming a cosmology-independent covariance in parameter space; the data vectors are again calculated from an analytical model (as explained e.g. in Sect. 5.2). In Fig. 8.4 we show the likelihood contours for a 900 deg^2 survey ($\sigma_\epsilon = 0.3$, $n_{gal} = 15$) when using Gaussian (left panels) and non-Gaussian covariances (right panel). It is interesting to note, that the contours corresponding to $\mathbf{C}^{Gaussian}$ are more elongated than the likelihood contours of $\mathbf{C}^{non-Gaussian}$. However, in the latter case the contours are significantly broadened; the parameter space within the individual credible intervals is larger. The goal of this section is to relate the

importance of non-Gaussianity to the angular range of the considered 2PCF data vector. In Fig. 8.4 the angular range extends from $\theta \in [0\rlap{.}'2 - 108\rlap{.}'0]$. In the following analysis we constantly decrease this interval (by cutting off the small scale data points) and compare the size of the credible intervals again. From now on we change the survey size to $A = 3200 \deg^2$, in order to assure that the major part of the credible intervals is inside the considered parameter spaces, i.e. $\Omega_{\mathrm{m}} \in [0.15; 0.35]$ vs. $\sigma_8 \in [0.8; 1.05]$, $\Omega_{\Lambda} \in [0.73; 0.77]$ vs. $w_0 \in [-1.5; -0.7]$, and $w_0 \in [-3; 2]$ vs. $w_a \in [-1.5; 0.0]$. As a measure for the size of the credible intervals we again use the values of q (see Sect. 4.2). The results are summarized in Fig. 8.5. The upper row shows the increase of q depending on the angular scale at which the data vector is cut-off (θ_{cut}) when using non-Gaussian covariances, the middle row illustrates the same effect but for non-Gaussian covariances. The lower panel shows the ratio

$$\Delta q = \frac{q^{\mathrm{Gaussian}}}{q^{\mathrm{non-Gaussian}}}, \qquad (8.2)$$

also depending on θ_{cut}. For the case of w_0 vs. w_a one clearly sees a break in the growth of the values of q at $\theta_{\mathrm{cut}} \approx 8'$. This is due to the fact that the likelihood contours exceed the considered parameter space when going above this threshold (see Fig. 8.8 and 8.9). In general one would expect that the impact of non-Gaussianity becomes less important when going to larger scales. Therefore, we expected Δq to approach unity as θ_{cut} increases. This effect is only noticeable for the Ω_{m} vs. σ_8 case; even there it is weak. We assume that the effect becomes more evident when increasing θ_{cut} even further. The decrease of Δq at $\theta_{\mathrm{cut}} \approx 8'$ is interesting, as it occurs in all considered parameter spaces. For the case of w_0 vs. w_a, one could explain this behavior by the fact that a significant part of the posterior likelihood is outside the considered parameter space, however this explanation fails for the other two parameter combinations. The effect might be an inherent feature of the covariance matrix related to the choice of the survey parameters, however the effect also occurred when changing the survey size. We postpone a more detailed analysis of this behavior to the future.

8.4 Impact of shot noise on parameter estimation

In this section we perform a similar analysis as in Sect. 8.3 but now comparing non-Gaussian covariances with and without shot noise ($\sigma_\epsilon = 0$). This implies that in the latter case only the cosmic variance term of the covariance contributes. Corresponding to (8.2) we define

$$\Delta q = \frac{q^{\mathrm{nonoise}}}{q^{\mathrm{noise}}} \qquad (8.3)$$

and examine the dependence of this ratio on the angular scale of the considered data vector. The results are summarized in Fig. 8.6. The left column shows the trend of q when increasing θ_{cut} for the case of a noise-free covariance matrix. Again there is a break at $\theta_{\mathrm{cut}} \approx 8'$ for the case of w_0 vs. w_a, for the same reasons as mentioned before. The corresponding plots for noisy covariances are shown in the middle row of Fig. 8.5. As shot noise is only dominant on small angular scales we again expected Δq to approach unity when going to higher θ_{cut}. This effect indeed occurs in the
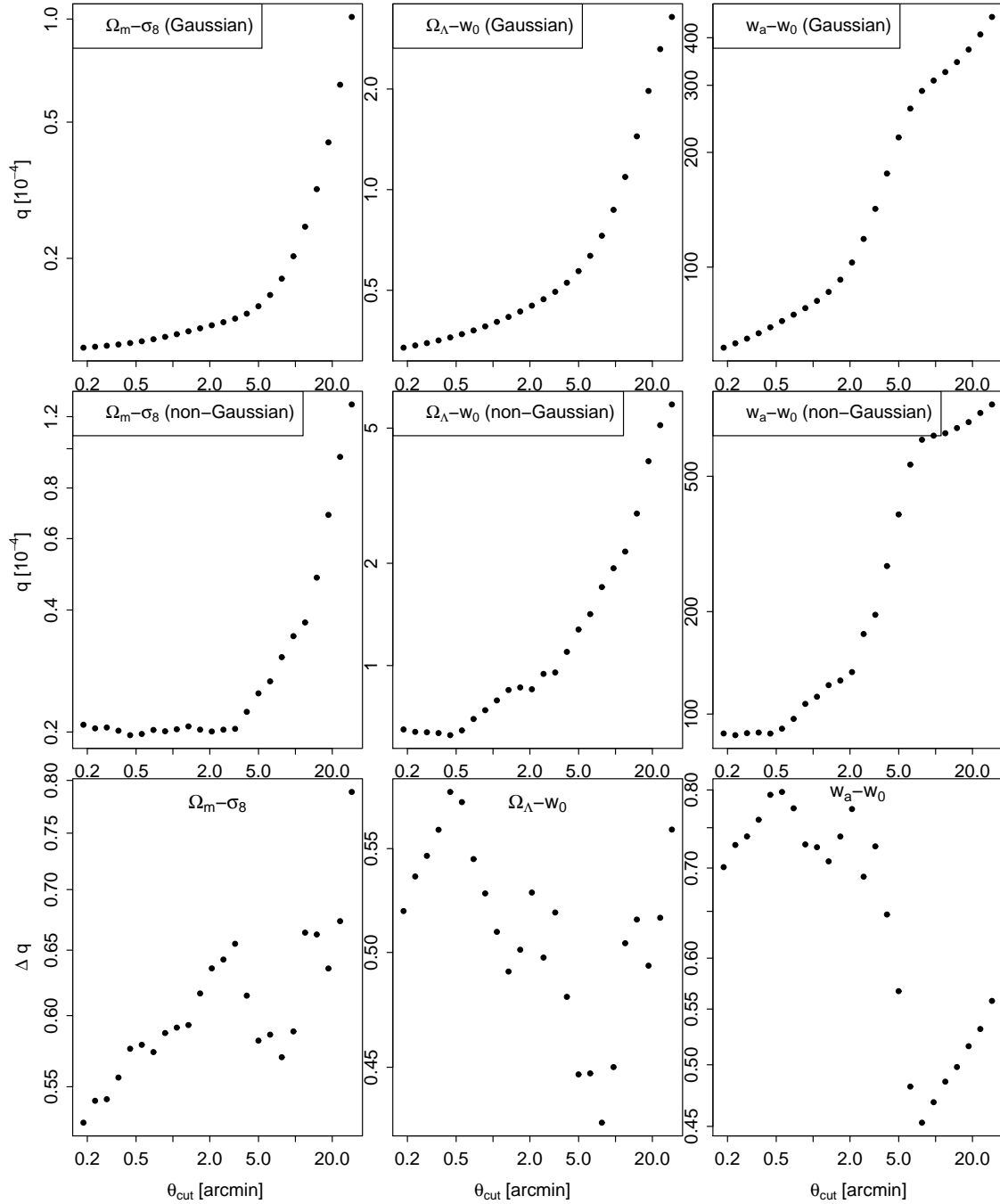
Figure 8.5: *This figure illustrates the growth of the credible regions (represented through qs) depending on the angular range of the considered data vector. This range is determined by θ_{cut}, more precisely the data vector extends from $[\theta_{cut} - 108'.0]$. We consider likelihood analyses employing Gaussian and non-Gaussian covariances (top and middle row, respectively) and show the dependence of the ratio $\Delta q = q^{Gaussian}/q^{non-Gaussian}$ on θ_{cut} in the last row.*

right panels of Fig. 8.6 where we see an overall increase in Δq throughout all examined parameter combinations. Still, the ratio does not become unity but stagnates at ≈ 0.6. These results imply that shot noise has a major impact on parameter constraints on small scales, and even on scales above $10'$ it still affects the parameter constraints significantly. For a better illustration of these effects we included a selection of contour plots in the next section.

8.5 Selected contour plots

In this section we show a selection of contour plots which further illustrate the results presented in Sects. 8.3 and 8.4 . The three figures in this section correspond to the three considered parameter spaces, namely Ω_m vs. σ_8 (Fig. 8.7) , Ω_Λ vs. w_0 (Fig. 8.8), and w_0 vs. w_a (Fig. 8.9). The left column always corresponds to the non-Gaussian noise-free covariance, the middle column is the noisy non-Gaussian covariance, and the right column shows the results when using a (noisy) Gaussian covariance.
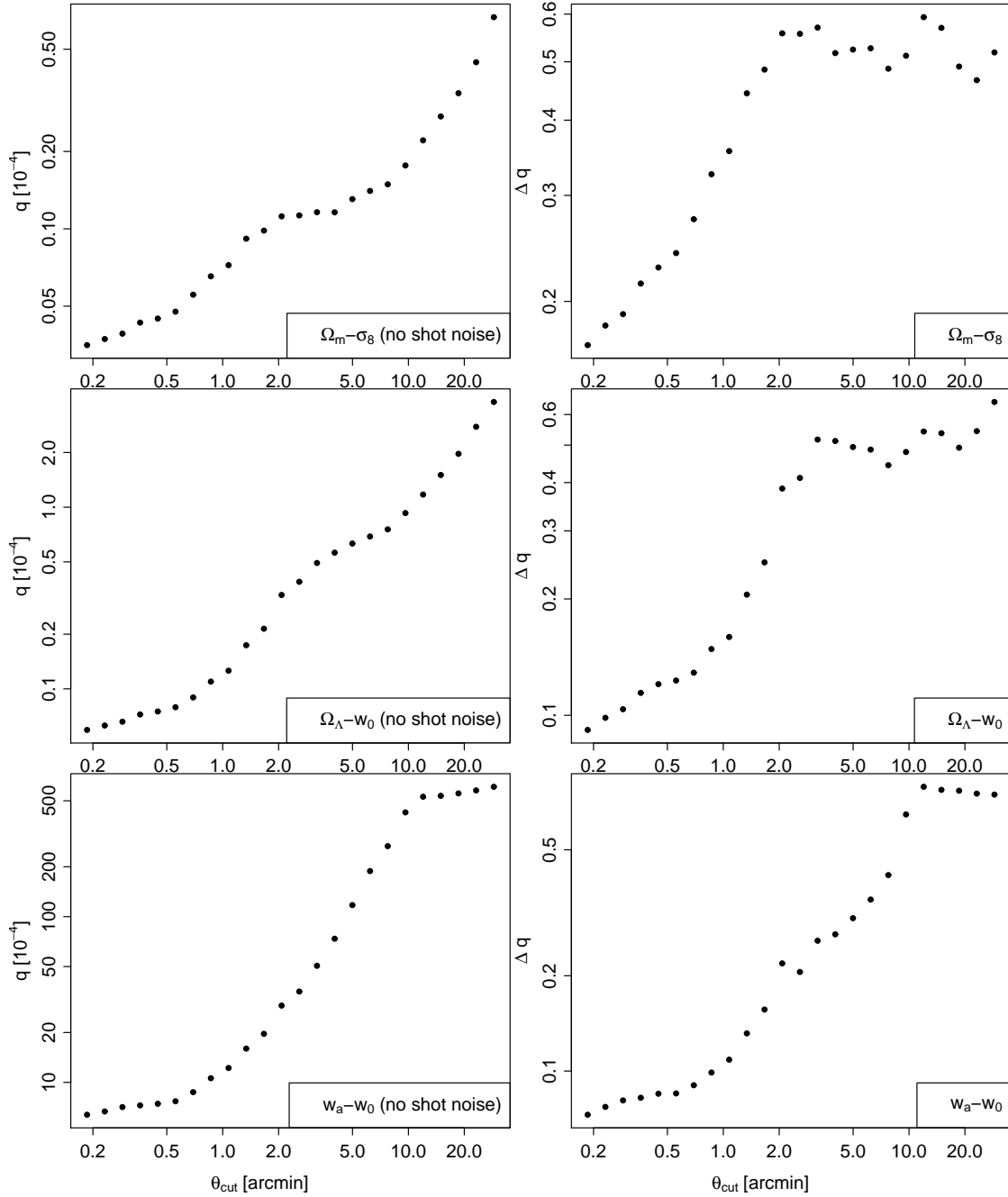
Figure 8.6: *This figure illustrates the impact of shot noise on the likelihood contours (q) depending on the considered angular range of the 2PCF data vector ($\vartheta \in [\theta_{\mathrm{cut}} - 108.'0]$). The left column shows the q when using a noise-free non-Gaussian covariance matrix; the right plot shows the ratio $\Delta q = q^{\mathrm{nonoise}}/q^{\mathrm{noise}}$, i.e. the ratio of the q when using noisy or noise-free covariances, respectively.*
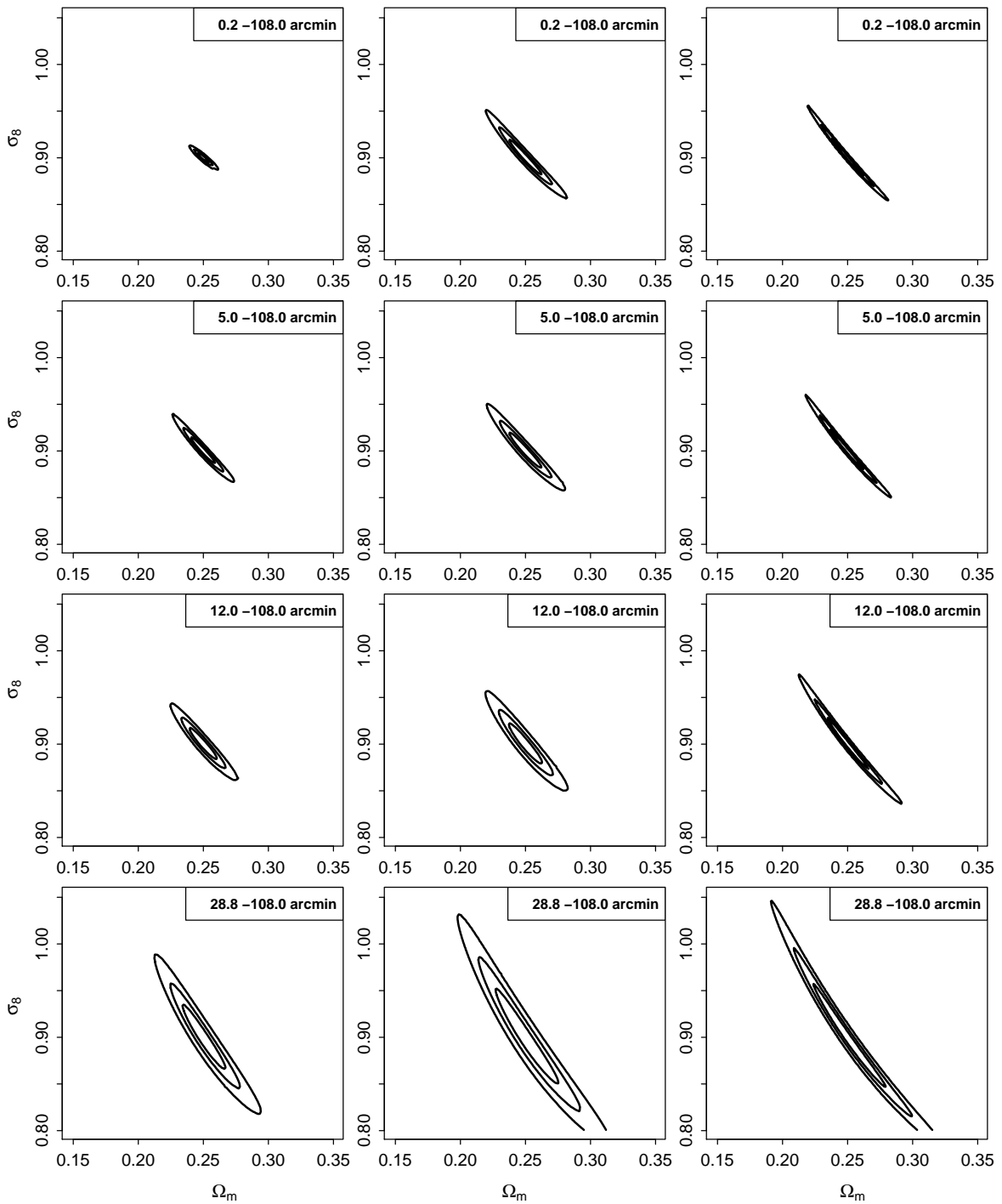
Figure 8.7: *Likelihood contours of* Ω_{m} *vs.* σ_8 *when using non-Gaussian noise-free covariances (left column), non-Gaussian noisy covariances (middle column), and noisy Gaussian covariances (right column). From top to bottom the considered angular range constantly decreases; it extends from* $[\theta_{\mathrm{cut}} - 108.'0]$.
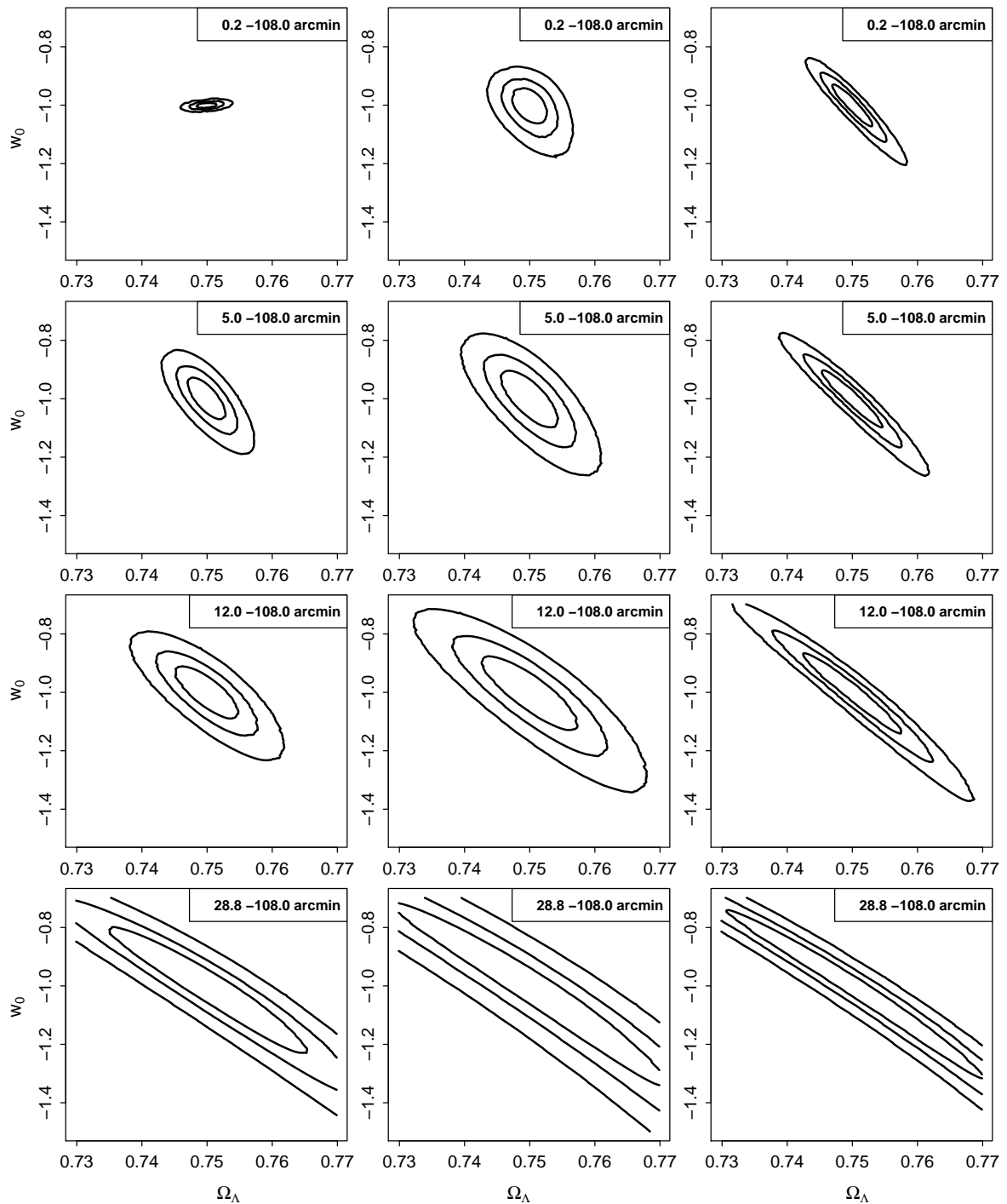
Figure 8.8: *Likelihood contours of Ω_Λ vs. w_0 when using non-Gaussian noise-free covariances (left column), non-Gaussian noisy covariances (middle column), and noisy Gaussian covariances (right column). From top to bottom the considered angular range constantly decreases; it extends from $[\theta_{\rm cut} - 108.'0]$.*
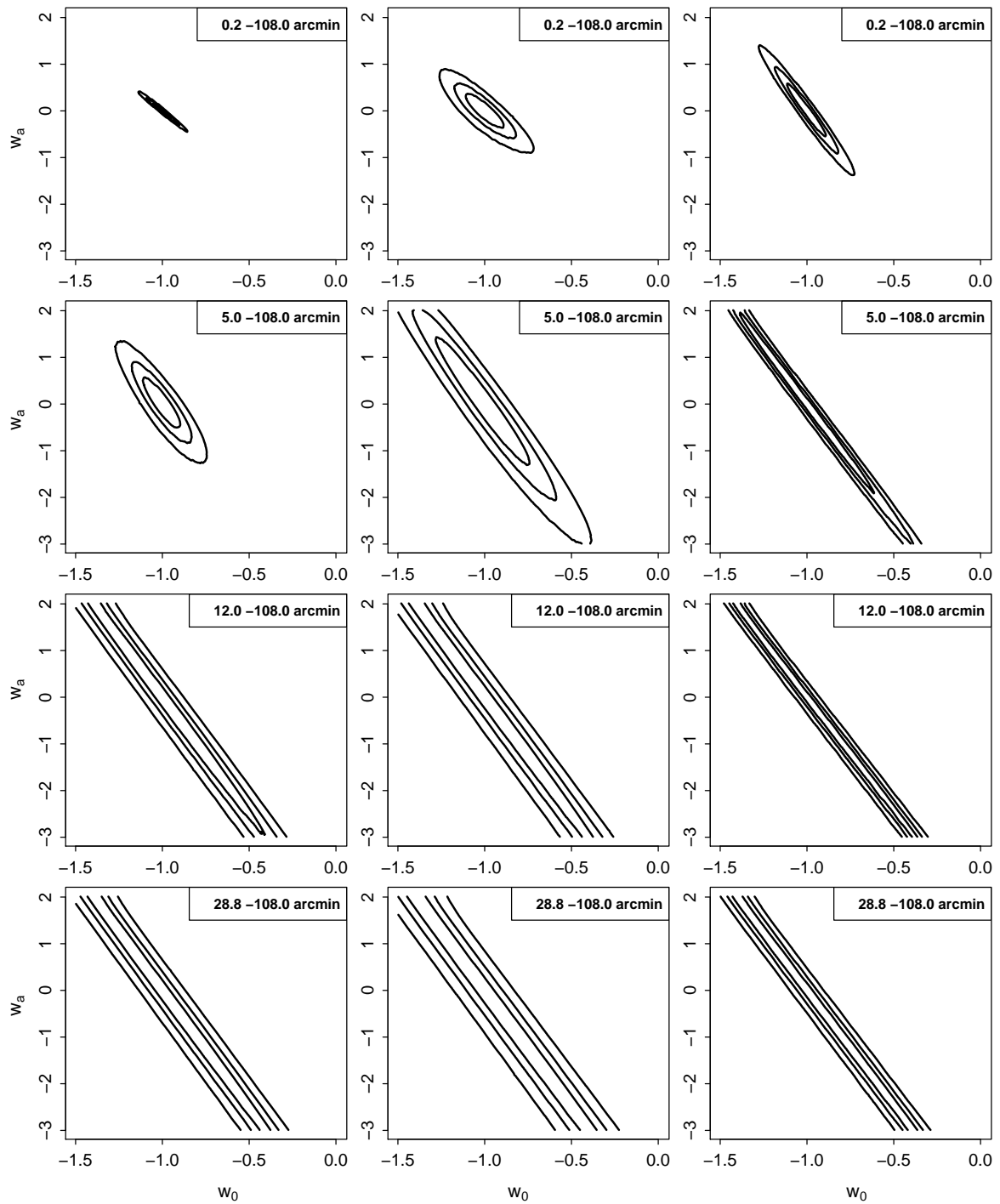
Figure 8.9: *Likelihood contours of w_0 vs. w_a when using non-Gaussian noise-free covariances (left column), non-Gaussian noisy covariances (middle column), and noisy Gaussian covariances (right column). From top to bottom the considered angular range constantly decreases; it extends from $[\theta_{\mathrm{cut}} - 108'\!.0]$.*

Chapter 9

Summary

In my PhD thesis I improve on theoretical aspects related to the estimation of cosmological parameters with cosmic shear. This involves the optimization of cosmic shear data vectors, accurate descriptions of the corresponding covariances, and the statistical methods used in the likelihood analysis. Furthermore, I applied my theoretical work to cosmic shear data. More precisely, I improve on the most recently developed cosmic shear measure, namely the ring statistics, and employ this statistics to constrain cosmological parameters using data from the Canada-France-Hawaii Telescope Legacy Survey (CFHTLS).

Chapter 2: Cosmology
In the first chapter I give an introduction to the basics of cosmology, explaining the theoretical framework which describes the evolution of our Universe, the details of the ΛCDM model, and the observational pillars on which this model is based. This chapter also contains a brief summary of the main concepts of structure formation.

Chapter 3: Gravitational lensing and cosmic shear
This chapter contains a summary of the concepts used in gravitational lensing with the focus on cosmic shear. In the following I summarize the most important information which is required in this summary.
All second-order statistical information on the shear field is contained in the power spectrum of the shear (P_γ) or its Fourier transform, the two-point correlation function (2PCF or $\xi_\pm$). The shear is a two-component quantity; similar to the decomposition of a vector field into a curl-free and divergence-free component, the shear can be decomposed into E- and B-modes. Whereas the first originate from lensing only (therefore contain the cosmological information), the second indicate remaining systematics in the data reduction/analysis. The 2PCF cannot separate E- and B-modes. In contrast, the so-called aperture mass dispersion ($\langle M_{\rm ap}^2 \rangle$), which measures the amount of density fluctuations within an aperture, is most commonly used to decompose the E- from the B-mode signal. In practice, $\langle M_{\rm ap}^2 \rangle$ cannot be measured from the cosmic shear data directly due to

gaps and holes in the data field caused by the masking of stars and satellite tracks. However, one can calculate $\langle M_{ap}^2 \rangle$ from the 2PCF, the latter being immune to these masking effects. The 2PCF can be obtained easily from the measured ellipticities of the observed galaxies and is considered the basic second-order cosmic shear measure.

Chapter 4: Parameter Estimation

In this chapter I introduce the basic concepts of cosmic shear parameter estimation which are intensively used in the following chapters. I explain the statistical framework using the 2PCF as a basic quantity and begin the chapter with introducing the 2PCF's estimators and covariances. I continue the chapter with a description of Baysian methods to infer parameters and introduce the concept of a Fisher matrix. The chapter concludes with an analysis of how the data's binning influences the parameter constraints and which numerical difficulties can occur when choosing a too narrow binning.

Chapter 5: Improvement of cosmic shear data vectors

In this chapter I show that, although 2PCF and $\langle M_{ap}^2 \rangle$ are both filtered versions of the power spectrum, a data vector of the first (ξ) contains more information on P_γ than a data vector of the second ($\langle M_{ap}^2 \rangle$). The reason for this is that ξ samples the power spectrum over a much broader range and also collects information on scales that are larger than the size of the survey. The data vector $\langle M_{ap}^2 \rangle$ lacks this large-scale information, but yields highly localized information on P_γ. Nevertheless, $\langle M_{ap}^2 \rangle$ has other advantages. First, the data points are much less correlated compared to the 2PCF data points. This leads to a mainly diagonal covariance matrix, which is numerically stabler during the inversion process in a likelihood analysis. Second, when considering higher-order statistics, $\langle M_{ap}^3 \rangle$ is much easier to handle than the three-point correlation function, and third, the aperture mass dispersion separates E- and B-modes. Based on these considerations I develop the combined data vector $\mathcal{N}$, which preserves the advantages of $\langle M_{ap}^2 \rangle$ and additionally provides large-scale information on P_γ through one data point of the 2PCF ($\hat{\xi}_+(\theta_0)$). Furthermore, the new data vector can be optimized with respect to the angular scale (θ_0). I compared the three data vectors in a detailed likelihood analysis and find that the combined data vector is a strong improvement over $\langle M_{ap}^2 \rangle$ in information content. The combined data vector $\mathcal{N}$ also maintains the other advantages of the aperture mass dispersion. Its covariance matrix is almost diagonal, and even the cross terms $C(\mathcal{M}(\theta_k), \hat{\xi}_+(\theta_0))$ are much smaller compared to the off-diagonal terms of $\mathbf{C}_\xi$. If B-modes are present, $\mathcal{N}$ is clearly less affected by the contamination compared to ξ. For the latter the parameter constraints are significantly weakened and, even worse, biased. These results are published in Eifler et al. (2008a).

Chapter 6: Ring statistics

As a second project related to data vectors, I examine the ring statistics, the most recently developed cosmic shear measure (Schneider & Kilbinger 2007). The ring statistics improves on deficits of commonly used measures to decompose E- and B-modes (e.g. $\langle M_{ap}^2 \rangle$). To calculate $\langle M_{ap}^2 \rangle$ properly from the 2PCF, the latter must be measured down to arbitrarily small scales which is not feasible in practice. As a consequence $\langle M_{ap}^2 \rangle$ suffers from mixing of E- and B-modes on

small angular scales (Kilbinger et al. 2006). In contrast, the ring statistics separates E- and B-modes properly using the measured 2PCFs, even on small scales. However, calculating the ring statistics from the 2PCF involves a filter function ($Z_\pm$), which has many free parameters and dependencies, and therefore is difficult to obtain. In the original paper the authors consider a special case of $Z_\pm$-functions which, once calculated, are held fixed independent of the considered scale of the ring statistics. In my thesis I relax this condition and choose a scale dependent filter function which improves the ring statistics' signal. A comparison of both methods shows, that the variable filter function significantly improves the signal strength, in particular on large scales. Furthermore, I develop a formula to calculate the ring statistics' covariance from the 2PCF covariance and use this formula to compare the correlation of data points and the information content of ring statistics and aperture mass dispersion. I find that the ring statistics is less correlated and therefore has a higher information content than $\langle M_{\mathrm{ap}}^2 \rangle$. In addition, the ring statistics can be calculated on smaller angular scales than $\langle M_{\mathrm{ap}}^2 \rangle$ which suffers from E/B-mode mixing on these scales. Finally, I employ the ring statistics to measure a shear signal from CFHTLS data. Note that this marks the first measured shear signal using the ring statistics. With the obtained shear signal I constrain cosmological parameters, in particular σ_8 in combination with Ω_{m}. As a result we obtain $\sigma_8 (\Omega_{\mathrm{m}}/0.25)^{0.49} = 0.82^{+0.02}_{-0.04}$. This is comparable to the results of foregoing cosmic shear analyses which employed the aperture mass dispersion or the E/B-mode correlation function, e.g. Benjamin et al. (2007) or Fu et al. (2008).

Chapter 7: Improved Likelihood analysis for cosmic shear data

In addition to data vectors, I intensively worked on cosmic shear covariances $\mathbf{C}$ and their impact on parameter constraints. In cosmic shear, there are several methods to derive covariances. First, one can calculate $\mathbf{C}$ analytically assuming a Gaussian shear field. This assumption breaks down on small angular scales (< 10 arcmin), where non-linearities of the matter density field start to become important. Second, covariances can be estimated from ray-tracing simulations. Although computationally more expensive, this method automatically accounts for the non-Gaussianity of the shear field. In both approaches the covariance is calculated assuming a specific cosmology, which must be chosen by the data analyst. Past cosmic shear data analyses approximate the covariance to be constant in parameter space, therefore assume that the chosen cosmology does not influence the result of a likelihood analysis significantly.

In my thesis I show that the covariance matrix depends non-negligibly on its underlying cosmology and that this "CDC-effect" significantly influences the likelihood contours of parameter constraints. To prove this, I calculate 2500 Gaussian covariance matrices for various parameters of $\Omega_{\mathrm{m}} \in [0.2; 0.4]$ and $\sigma_8 \in [0.6; 1.0]$; all other cosmological parameters are held fixed. Even a change of Ω_{m} and σ_8 within the WMAP5 68% confidence levels has a non-negligible impact on the covariances and the obtained likelihood contours. Furthermore, I show that the impact of the CDC-effect depends on survey parameters and that a proper treatment will be important in the future, for large and deep surveys. To account for the cosmology dependence of covariances I develop a likelihood analysis with an adaptive covariance matrix. Here, $\mathbf{C}$ is calculated from the cosmological model for which the likelihood is evaluated. For small surveys this method introduces a bias to the best-fit parameter set, which vanishes when going to larger survey sizes. A

disadvantage of this approach is its computational costs, as the covariance must be recalculated at every sampled point in parameter space. Using the analytic expression for Gaussian covariances is already time-consuming, using ray-tracing covariances to include the non-Gaussianity is not feasible with today's computer power. For the Gaussian case, I derive a scaling relation for covariances, which reduces the computational costs significantly. An additional effect is that the adaptive covariance approach enhances the constraints on cosmology. These results are published in Eifler et al. (2008b).

Chapter 8: Impact of non-Gaussianity and shot noise on likelihood contours
The last chapter of my thesis addresses the impact of non-Gaussian covariances on dark energy constraints with cosmic shear. I find that even when restricting the data vector to scales above $20'$ non-Gaussianity influences the parameter constraints and cannot be neglected. In addition, this chapter contains an analysis of the impact of shot noise on parameter constraints. As expected, the impact of shot noise is large if small angular scales ($\leq 5'$) are included into the data vector. Considering only 2PCFs above this angular separation reduces the impact of shot noise; however, even when only considering scales above $20'$ shot noise is non-negligible and contributes to the error bars significantly.

Chapter 10

Outlook

The research projects of this thesis can be continued and extended in the future. In particular, the work of chapter 6 provides a good starting point for further research, similar to the work of chapter 7 in combination with chapter 8. I will briefly outline my future plans in this chapter.

Ring statistics:
The ring statistics is a special case of a general E- and B-mode decomposition. I plan to investigate this general decomposition in detail, in particular, I will focus on optimizing the corresponding filter function with respect to the signal strength and the information content. Furthermore, I want to apply the ring statistics and the general E/B-mode statistics to weak lensing data from space-based observations (e.g. the COSMOS survey), which is a logical extension of my work with the CFHTLS data. Here, the ring statistics is especially useful for the following reasons. The E- and B-mode mixing of the aperture mass dispersion (Kilbinger et al. 2006) mainly occurs on small scales; the COSMOS analysis of Massey et al. (2007b) compensates for this issue by extrapolating the data using a theoretical model. The ring statistics is independent of any theoretical assumptions, in addition it provides tighter constraints on cosmological parameters. I consider the ring statistics, respectively the general E/B-mode statistics, a candidate for the best second-order cosmic shear measure.

Cosmic shear covariances:
Currently, covariances are derived from analytic formula/simulations or from the data itself. However, there exists another approach, called shrinkage estimation, which combines prior knowledge from theoretical models (contained in a so-called *target*) with information obtained from the data. This ansatz is well known in other science branches (Ledoit & Wolf 2004; Schäfer & Strimmer 2005). Recently Pope & Szapudi (2008) applied it to derive a power spectrum covariance for the matter density spectrum.
I want to apply the shrinkage approach to the 2PCF covariance of cosmic shear in order to derive an improved estimator. Second, I want to extend the work of Eifler et al. (2008b). As outlined there, an optimal likelihood analysis requires non-Gaussian covariances for every point in parameter space where the likelihood is evaluated (adaptive covariance). Deriving these covariances

via ray-tracing is computationally too expensive, such a brute-force approach can only be done for Gaussian covariances. My idea is to extend the concept of an adaptive covariance from Gaussian to non-Gaussian covariances using the shrinkage formalism. More precisely, I define the shrinkage target as a Gaussian covariance (which can be calculated for arbitrary cosmology) and incorporate the non-Gaussianity through the covariance estimated from the data. This provides a method to derive non-Gaussian covariances for arbitrary cosmology, in the next task I explain how to test the accuracy of this method.

In collaboration with AstroGrid[1] the lensing group at the Argelander-Institut für Astronomie (University Bonn) has access to ray-tracing simulations for 200 cosmologies, more precisely 200 different combinations of Ω_m and σ_8. I want to derive cosmic shear ray-tracing covariances from these simulations and compare them to the non-Gaussian covariance derived through the shrinkage approach.

The above two tasks aim to derive precise models for non-Gaussian covariances for different cosmologies. By comparing these model covariances with the one estimated from the data one can infer information on cosmological parameters. This method simply replaces the mean data vector of an "ordinary" likelihood analysis by a covariance. For the reason that non-Gaussian covariances are sensitive to higher-order information of the shear field, we can use this ansatz to incorporate this higher-order information into a likelihood analysis. The final goal is a joint likelihood formalism for mean data vector and estimated covariance. The main difficulty in this ansatz is that one has to derive a covariance for covariances. However, the mathematical formalism for this already exists (Basser & Pajevic 2002, 2007). It is hard to predict how much this ansatz can improve parameter constraints; Kilbinger & Schneider (2005) have shown that the inclusion of higher-order information can substantially enhance the parameter constraints, e.g. by a factor of two for the parameter combination Ω_m and σ_8.

Combining information from different cosmological probes:
Combining information from different cosmological probes in a joint likelihood analysis significantly improves the parameter constraints. In particular for future dark energy studies a combination of Planck data with cosmic shear data (e.g. JDEM, Euclid) is fruitful, as the CMB is sensitive to *early dark energy models* (Doran et al. 2001), whereas cosmic shear is very sensitive to dark energy in the late Universe.

Although the concept of a joint likelihood analysis is well-known in cosmology (e.g. Komatsu et al. 2008; Kilbinger et al. 2008), the currently used methods neglect cross-correlations between the individual cosmological probes, assuming them to be independent. This assumption can be true for some cases of cosmological probes; if so, the corresponding likelihood analyses can be performed independently, and later combined via Bayes theorem, where the result of the first likelihood analysis enters as prior information into the second likelihood analysis. For the case of cosmic shear and CMB data this assumption clearly fails, as the CMB is lensed by the large-scale structure. I want to develop a joint covariance matrix which also contains cross-correlations of CMB and cosmic shear data. These cross-terms quantify the amount of CMB lensing between the redshift of the cosmic shear source galaxies and the observer. Estimators for CMB lensing are

---

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

intensively discussed in the literature (e.g. Hu 2002; Hirata & Seljak 2003a,b); they are mainly based on the CMB observables T, E, B and therefore account for CMB lensing from $z = 1100$ until today. Recently, Marian & Bernstein (2007) constructed an estimator for the CMB lensing potential using the weak lensing signal of foreground galaxies. This estimator, originally designed to delens the CMB polarization map, can be employed to calculate the CMB lensing in the required redshift range yielding the desired cross-terms. After incorporating the full covariance matrix into a likelihood analysis, I want to test this method on simulated and, if possible, on real data.

Furthermore, the field of CMB lensing itself offers many research opportunities (see Lewis & Challinor 2006,for a review). Its sensitivity to the matter density field at intermediate redshifts, which are inaccessible to e.g. cosmic shear surveys, indicate the scientific potential. In addition, the CMB is unaffected from non-linear structure growth and has a precisely known redshift. Recently, there has been strong evidence for a detection of CMB lensing (Smith et al. 2007; Hirata et al. 2008). With upcoming Planck data (higher resolution, lower noise) CMB lensing can progress to a "competitive" cosmological probe. Equally interesting is the question how to remove lensing contaminations from CMB data properly in order to constrain primordial B-modes in the CMB polarization map. Their detection would strongly support the theory of inflation; from the amplitude of the B-modes one can even infer information on inflation's energy scale.

Appendix A

Comparison of two measures

We compare the information content of two arbitrary data vectors referring to them as primary data vector p and secondary data vector s. We further assume that s can be calculated from p by a *transfer matrix* $\mathbf{A}$ (dimension $n \times m$), with arbitrary n and m

$$p = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_m \end{pmatrix} \quad \text{and} \quad s = \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix} \quad \text{with} \quad s = \mathbf{A}\,p\,. \tag{A.1}$$

We define the covariance matrices of these data vectors as

$$\mathbf{C}_{\mathrm{p}} = \left\langle (p - \hat{p})(p - \hat{p})^{\mathrm{t}} \right\rangle, \tag{A.2}$$

$$\mathbf{C}_{\mathrm{s}} = \left\langle (s - \hat{s})(s - \hat{s})^{\mathrm{t}} \right\rangle, \tag{A.3}$$

where $\hat{p}$ ($\hat{s}$) denotes the estimated and p (s) the true values of primary (secondary) measure. Using (A.1) we can relate both covariances through

$$\mathbf{C}_{\mathrm{s}} = \mathbf{A}\,\mathbf{C}_{\mathrm{p}}\,\mathbf{A}^{\mathrm{t}}\,. \tag{A.4}$$

The transformation matrix $\mathbf{A}$ has to be of rank $\mathbf{A} = n$, otherwise the covariance matrix of the secondary data vector $\mathbf{C}_{\mathrm{s}} = (\mathbf{A}\,\mathbf{C}_{\mathrm{p}}\,\mathbf{A}^{\mathrm{t}})$ is singular and not invertible. Furthermore, as $\mathbf{A}$ is of dimension $(n \times m)$, rank $\mathbf{A} \leq m$ implying $n \leq m$. We take the χ^2-functions a measure for the information content

$$\chi_{\mathrm{p}}^2 = \Delta_{\mathrm{p}}^{\mathrm{t}}\,\mathbf{C}_{\mathrm{p}}^{-1}\,\Delta_{\mathrm{p}} \qquad \text{and} \qquad \chi_{\mathrm{s}}^2 = \Delta_{\mathrm{s}}^{\mathrm{t}}\,\mathbf{C}_{\mathrm{s}}^{-1}\,\Delta_{\mathrm{s}}\,, \tag{A.5}$$

where in our case $\Delta_{\mathrm{p}} = p^{\mathrm{f}} - p_{\pi}$ ($\Delta_{\mathrm{s}} = s^{\mathrm{f}} - s_{\pi}$) denotes the difference between the fiducial data vector p^{f} (s^{f}) and the data vector p_{π} (s_{π}) depending on the parameter vector π. If χ^2 is minimal, the posterior likelihood of the corresponding π being the correct parameter vector is maximized. The difference between χ_{p}^2 and χ_{s}^2 characterizes which probability function has a larger curvature, i.e. which data vector gives tighter constraints in parameter space. Therefore the information content of primary and secondary data vectors can be compared by calculating

$$\chi_{\mathrm{p}}^2 - \chi_{\mathrm{s}}^2 = \Delta_{\mathrm{p}}^{\mathrm{t}}\,\mathbf{C}_{\mathrm{p}}^{-1}\,\Delta_{\mathrm{p}} - \Delta_{\mathrm{p}}^{\mathrm{t}}\,\mathbf{A}^{\mathrm{t}}\left(\mathbf{A}\,\mathbf{C}_{\mathrm{p}}\,\mathbf{A}^{\mathrm{t}}\right)^{-1}\mathbf{A}\,\Delta_{\mathrm{p}}\,, \tag{A.6}$$

for arbitrary Δ_{p}. In case this difference is always positive, we can conclude that the primary data vector gives tighter constraints on parameters. We can always find transformation matrices $\mathbf{V}$ (dimension $m \times m$) and $\mathbf{U}$ (dimension $n \times n$) to rewrite the transfer matrix $\mathbf{A}$ as an $n \times m$ matrix

$$\left(\mathbf{E}_n \,\middle|\, \mathbf{0} \right) = \mathbf{S} = \mathbf{U} \mathbf{A} \mathbf{V}^{-1} \quad \longleftrightarrow \quad \mathbf{A} = \mathbf{U}^{-1} \mathbf{S} \mathbf{V}. \tag{A.7}$$

We can directly calculate these transformation matrices as a multiplication of elementary matrices (Fischer 1997a). Inserting (A.7) into (A.6), and after some lengthy but straightforward calculation we derive,

$$\chi_{\mathrm{p}}^2 - \chi_{\mathrm{s}}^2 = \Delta_{\mathrm{p}}'^{\mathrm{t}} \mathbf{C}'^{-1} \Delta_{\mathrm{p}}' - \Delta_{\mathrm{p}}'^{\mathrm{t}} \mathbf{S}^{\mathrm{t}} \left(\mathbf{S} \mathbf{C}' \mathbf{S}^{\mathrm{t}} \right)^{-1} \mathbf{S} \Delta_{\mathrm{p}}' \tag{A.8}$$

with

$$\mathbf{C}' = \mathbf{V} \mathbf{C}_{\mathrm{p}} \mathbf{V}^{\mathrm{t}} \quad \text{and} \quad \Delta_{\mathrm{p}}' = \mathbf{V} \Delta_{\mathrm{p}}. \tag{A.9}$$

For simpler notation we discard all " $'$ " later on. We define

$$\mathbf{C}^{-1} = \left(\begin{array}{c|c} \mathbf{C}_1 & \mathbf{C}_2 \\ \hline \mathbf{C}_2^{\mathrm{t}} & \mathbf{C}_3 \end{array} \right)^{-1} = \left(\begin{array}{c|c} \mathbf{D}_1 & \mathbf{D}_2 \\ \hline \mathbf{D}_2^{\mathrm{t}} & \mathbf{D}_3 \end{array} \right), \tag{A.10}$$

with $\mathbf{C}_1$ being an $n \times n$ matrix and calculate

$$\mathbf{S}^{\mathrm{t}} \left(\mathbf{S} \mathbf{C} \mathbf{S}^{\mathrm{t}} \right)^{-1} \mathbf{S} = \left(\begin{array}{c|c} \mathbf{C}_1^{-1} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array} \right). \tag{A.11}$$

Using (A.10) and (A.11) we can rewrite (A.8) as

$$\chi_{\mathrm{p}}^2 - \chi_{\mathrm{s}}^2 = \Delta_{\mathrm{p}}^{\mathrm{t}} \left(\begin{array}{c|c} \mathbf{D}_1 - \mathbf{C}_1^{-1} & \mathbf{D}_2 \\ \hline \mathbf{D}_2^{\mathrm{t}} & \mathbf{D}_3 \end{array} \right) \Delta_{\mathrm{p}}. \tag{A.12}$$

From $\mathbf{C} \mathbf{D} = \mathbf{E}_m$ we deduce

$$\mathbf{C}_1 \mathbf{D}_1 + \mathbf{C}_2 \mathbf{D}_2^{\mathrm{t}} = \mathbf{E}_n \longrightarrow \mathbf{D}_1 - \mathbf{C}_1^{-1} = -\mathbf{C}_1^{-1} \mathbf{C}_2 \mathbf{D}_2^{\mathrm{t}} \tag{A.13}$$

and

$$\mathbf{C}_1 \mathbf{D}_2 + \mathbf{C}_2 \mathbf{D}_3 = 0 \longrightarrow \mathbf{C}_2 = -\mathbf{C}_1 \mathbf{D}_2 \mathbf{D}_3^{-1}. \tag{A.14}$$

Inserting (A.14) into (A.13) we can rewrite (A.12) as

$$\chi_{\mathrm{p}}^2 - \chi_{\mathrm{s}}^2 = \Delta_{\mathrm{p}}^{\mathrm{t}} \left(\begin{array}{c|c} \mathbf{D}_2 \mathbf{D}_3^{-1} \mathbf{D}_2^{\mathrm{t}} & \mathbf{D}_2 \\ \hline \mathbf{D}_2^{\mathrm{t}} & \mathbf{D}_3 \end{array} \right) \Delta_{\mathrm{p}}. \tag{A.15}$$

The matrix $\mathbf{C}$ is positive definite and symmetric, therefore $\mathbf{D}_3$ as a submatrix is positive definite and symmetric, and the inverse $\mathbf{D}_3^{-1}$ also has these favorable properties (Anderson 2003). Hence, we can decompose $\mathbf{D}_3 = \mathbf{L}\mathbf{L}^t$ and finish our calculation as follows

$$
\chi_p^2 - \chi_s^2 = \Delta_p^t \begin{pmatrix} \mathbf{D}_2(\mathbf{L}^t)^{-1} \\ \mathbf{L} \end{pmatrix} \underbrace{\left(\mathbf{L}^{-1}\,\mathbf{D}_2^t\,\mathbf{L} \right)}_{\mathbf{T}} \Delta_p \tag{A.16}
$$

$$
\begin{aligned}
&= \Delta_p^t \mathbf{T}^t\,\mathbf{T}\Delta_p \\
&= \|\mathbf{T}\Delta_p\|^2 \\
&\geq 0.
\end{aligned} \tag{A.17}
$$

We now examine the case where $\chi_p^2 - \chi_s^2 = 0$. The information content of primary and secondary measures is considered to be equal if and only if this equality holds for **all** data vectors Δ_p. If there is only one Δ_p for which $\chi_p^2 - \chi_s^2 > 0$, the primary measure contains more information. The difference between the two χ^2-values is given by (A.6). In case it is zero for all Δ_p,

$$
\mathbf{C}_p^{-1} = \mathbf{A}^t \left(\mathbf{A}\,\mathbf{C}_p\,\mathbf{A}^t \right)^{-1} \mathbf{A} \tag{A.18}
$$

must hold (Fischer 1997b). The matrix $\mathbf{C}_p$ is of rank m, hence the lefthandside of (A.18) must also have rank m. Then $\mathbf{A}$ must have rank m and is therefore a quadratic $m \times m$ matrix, which is of course invertible. This result is intuitively clear, if one is able to calculate Δ_s from Δ_p, and vice versa the information content should be the same. We can summarize the results of the above calculation in two statements:

1. If a secondary measure can be calculated from a primary by a matrix $\mathbf{A}$ as described in (A.1), the secondary measure has less or equal information.

2. The amount of information is equal in case the rank of $\mathbf{A}$ equals the dimension of the primary data vector (m) implying that $\mathbf{A}$ is invertible.

Appendix B

Fit-coefficients for the linear term of the 2PCF covariance

Table B.1: *The fit-coefficients for the linear term of the* $\mathbf{C}_{++}$ *covariance. For each bin the upper value corresponds to α, the lower to β. The maximum standard error of the fit is smaller than 0.04 % for β, smaller than 0.02% for α. In most cases the standard error is ≈ 0.02 % for β and ≈ 0.01 % for α.*

$\vartheta_{i,j}$ [arcmin]	1.0	4.9	10.3	15.7	33.0	45.4	69.3	85.7	106.0	131.0	162.0	180.0
1.0	1.1892	1.3888	1.4163	1.4212	1.4485	1.4726	1.5186	1.5485	1.5841	1.6262	1.6768	1.7063
	2.689	2.4905	2.1437	2.0102	1.9220	1.9198	1.9322	1.9418	1.9522	1.9627	1.9734	1.9781
4.9	1.3888	1.2955	1.4126	1.4214	1.4483	1.4723	1.5183	1.5483	1.5838	1.6260	1.6767	1.7061
	2.4905	2.5137	2.1993	2.0283	1.9232	1.9201	1.9322	1.9417	1.9520	1.9628	1.9731	1.9782
10.3	1.4163	1.4126	1.3243	1.4190	1.4476	1.4712	1.5173	1.5474	1.5830	1.6253	1.6761	1.7055
	2.1437	2.1993	2.3725	2.1168	1.9283	1.9210	1.9319	1.9414	1.9518	1.9625	1.9729	1.9781
15.7	1.4212	1.4214	1.4190	1.3395	1.4465	1.4695	1.5156	1.5458	1.5816	1.6240	1.6750	1.7045
	2.0102	2.0283	2.1168	2.3110	1.9401	1.9230	1.9314	1.9408	1.9513	1.9621	1.9728	1.9778
33.0	1.4493	1.4483	1.4476	1.4465	1.3677	1.4607	1.5055	1.5365	1.5733	1.6166	1.6684	1.6982
	1.9222	1.9232	1.9283	1.9401	2.2460	1.9614	1.9301	1.9377	1.9484	1.9598	1.9710	1.9765
45.4	1.4726	1.4723	1.4712	1.4695	1.4607	1.3802	1.4944	1.5259	1.5637	1.6082	1.6609	1.6911
	1.9198	1.9201	1.9210	1.9230	1.9614	2.2323	1.9340	1.9351	1.9451	1.9571	1.9691	1.9749
69.3	1.5186	1.5183	1.5173	1.5156	1.5055	1.4944	1.396	1.4976	1.5370	1.5845	1.6399	1.6712
	1.9322	1.9322	1.9319	1.9314	1.9301	1.9334	2.2234	1.9491	1.9381	1.9496	1.9634	1.9701
85.7	1.5485	1.5483	1.5474	1.5458	1.5365	1.5259	1.4976	1.4023	1.5135	1.5629	1.6208	1.6532
	1.9418	1.9417	1.9414	1.9408	1.9377	1.9351	1.9491	2.2219	1.9433	1.943	1.9580	1.9655
106.6	1.5841	1.5838	1.5830	1.5816	1.5733	1.5637	1.5370	1.5135	1.4078	1.5308	1.5922	1.6264
	1.9522	1.9520	1.9518	1.9513	1.9484	1.9451	1.9381	1.9433	2.2218	1.94082	1.9497	1.9583
131.0	1.6262	1.6260	1.6253	1.6240	1.6166	1.6082	1.5845	1.5629	1.5308	1.4120	1.5496	1.5864
	1.9627	1.9628	1.9625	1.9621	1.9598	1.9571	1.9496	1.9433	1.9408	2.2226	1.9409	1.9473
162.0	1.6768	1.6767	1.6761	1.6750	1.6684	1.6609	1.6399	1.6208	1.5922	1.5496	1.4148	1.5279
	1.9734	1.9731	1.9729	1.9728	1.9710	1.9691	1.9634	1.9580	1.9497	1.9409	2.2242	1.9495
180.0	1.7063	1.7061	1.7055	1.7045	1.6982	1.6911	1.6712	1.6532	1.6264	1.5864	1.5279	1.4157
	1.9781	1.9782	1.9781	1.9778	1.9765	1.9749	1.9701	1.9655	1.9583	1.9473	1.9495	2.225

Table B.2: *The fit-coefficients for the linear term of the* **C**$_{--}$ *covariance. For each bin the upper value corresponds to* α, *the lower to* β. *The maximum standard error of the fit is smaller than 0.04 % for* β, *smaller than 0.02% for* α. *In most cases the standard error is* ≈ *0.02 % for* β *and* ≈ *0.01 % for* α.

$\vartheta_{i,j}$ [arcmin]	1.0	4.9	10.3	15.7	33.0	45.4	69.3	85.7	106.0	131.0	162.0	180.0
1.0	0.8146	1.1532	1.331	1.4036	0.6502	0.7446	0.6435	1.4427	0.8483	1.115	0.900	0.8539
	2.3102	2.9769	3.1407	3.0122	1.4087	1.5601	0.7488	4.3083	2.0481	5.0329	1.8159	2.6019
4.9	1.1532	0.9987	1.3058	1.3935	1.3981	1.3473	0.6796	1.1596	1.1655	1.1080	1.1260	1.063
	2.9769	2.5973	3.1406	2.9523	2.2487	2.0178	3.0619	2.5425	3.0956	3.9419	3.3497	3.2596
10.3	1.3307	1.3058	1.0946	1.3548	1.4015	1.375	1.3997	1.4481	1.3982	0.6967	1.2972	1.1362
	3.1407	3.1406	2.777	3.0472	2.280	2.053	1.8821	1.708	1.9703	6.7254	3.6954	7.0468
15.7	1.4036	1.3935	1.3548	1.1431	1.4052	1.3816	1.3711	1.3803	1.3884	1.3681	1.4430	1.4899
	3.0122	2.9523	3.0472	2.8242	2.3463	2.0730	1.9219	1.8869	1.8507	2.1066	1.8902	0.8581
33.0	0.6519	1.3981	1.4015	1.4052	1.2103	1.3983	1.3711	1.3728	1.3832	1.4006	1.4284	1.4404
	1.4122	2.2487	2.280	2.3463	2.7801	2.2915	1.9407	1.8973	1.8868	1.898	1.8862	1.9033
45.4	0.7447	1.3473	1.3750	1.3816	1.3983	1.2321	1.3761	1.3724	1.3813	1.40	1.4244	1.4405
	1.5601	2.0178	2.0530	2.0730	2.2915	2.7192	1.9933	1.910	1.8879	1.8913	1.9095	1.9213
69.3	0.64346	0.67964	1.3997	1.3711	1.3711	1.3761	1.2562	1.3808	1.3775	1.3931	1.4193	1.4355
	0.7488	3.0619	1.8821	1.9219	1.9407	1.9933	2.6167	2.0248	1.9049	1.8889	1.8982	1.9059
85.7	1.4427	1.1596	1.4481	1.3803	1.3728	1.3724	1.3808	1.2668	1.3793	1.3878	1.4133	1.4298
	4.3084	2.5425	1.7080	1.8869	1.8973	1.910	2.0248	2.5626	1.961	1.8931	1.8964	1.9045
106.6	0.8482	1.1655	1.3982	1.3884	1.3832	1.3813	1.3775	1.3793	1.2770	1.3833	1.4045	1.4211
	2.0481	3.0956	1.9703	1.8507	1.8868	1.8879	1.9050	1.961	2.5101	1.9236	1.8944	1.9005
131.0	1.1150	1.1080	0.69672	1.3681	1.4006	1.3995	1.3931	1.3878	1.3833	1.2871	1.3931	1.4081
	5.0329	3.9419	6.7254	2.1067	1.898	1.8913	1.8889	1.8930	1.9237	2.4610	1.9060	1.8975
162.0	0.90	1.1260	1.2972	1.4430	1.4284	1.4244	1.4193	1.4133	1.4045	1.3931	1.2973	1.3944
	1.8159	3.3497	3.6954	1.8902	1.8862	1.9095	1.8982	1.8964	1.8944	1.9060	2.4166	1.9428
180.0	0.8539	1.063	1.1362	1.4899	1.4404	1.4405	1.4354	1.4298	1.4211	1.4081	1.3944	1.3025
	2.6019	3.2596	7.0468	0.8581	1.9033	1.9213	1.906	1.9045	1.9005	1.8975	1.943	2.3965

Table B.3: *The fit-coefficients for the linear term of the* **C**$_{+-}$ *covariance. For each bin the upper value corresponds to α, the lower to β. In contrast to* **C**$_{++}$ *and* **C**$_{--}$ *the standard error here is higher and the fit is less accurate. For most of the coefficients the standard error is still below 0.1% for both α and β, however there are outliers (marked in red) where the standard error exceeds 1%.*

$\vartheta_{i,j}$ [arcmin]	1.0	4.9	10.3	15.7	33.0	45.4	69.3	85.7	106.0	131.0	162.0	180.0
1.0	0.9198	1.0776	1.2039	1.2690	1.3444	1.3588	1.3680	1.3720	1.3776	1.3857	1.3968	1.4036
	2.1759	2.7594	2.9587	2.9510	2.6699	2.4892	2.2673	2.1796	2.1105	2.0591	2.0232	2.0098
4.9	0.7385	1.8641	1.2002	1.2616	1.3433	1.3585	1.368	1.3721	1.3776	1.3857	1.3968	1.4036
	2.1639	4.6092	2.8784	2.9351	2.6728	2.4916	2.2684	2.1802	2.1109	2.0594	2.0232	2.0099
10.3	1.1172	1.0939	0.3245	1.2212	1.3388	1.3572	1.3679	1.3729	1.3777	1.3857	1.3968	1.4036
	3.3921	3.2594	1.5224	3.0572	2.6790	2.4998	2.2725	2.1827	2.1123	2.0601	2.0236	2.0102
15.7	1.0298	1.2851	1.2017	0.7253	1.3241	1.3544	1.3678	1.3721	1.3777	1.3857	1.3968	1.4036
	3.9545	3.7686	3.5351	2.2997	2.6384	2.5113	2.2797	2.1871	2.1147	2.0614	2.0243	2.0106
33.0	0.7624	1.5316	1.4728	1.4719	0.9712	1.3427	1.3660	1.3730	1.3785	1.3861	1.3968	1.4035
	1.5130	3.5641	3.7076	3.7696	2.8402	2.7344	2.321	2.2166	2.1312	2.0699	2.0284	2.0135
45.4	0.7845	1.0214	1.3619	1.4555	1.4584	1.0303	1.3650	1.3824	1.3810	1.3871	1.3970	1.4034
	1.5519	2.2561	3.0393	3.1508	3.6129	2.9527	2.4350	2.2705	2.1564	2.0818	2.0340	2.0173
69.3	0.5863	0.9360	-5.4805	1.3951	1.3855	1.4222	1.0895	1.3595	1.3635	1.4203	1.4007	1.4049
	0.6329	2.6217	5.1207	2.3331	2.4026	2.6896	3.0243	2.4258	2.1906	2.1918	2.0579	2.0318
85.7	1.1916	1.0523	0.8196	1.4040	1.3075	1.3291	1.4146	1.1132	1.3654	1.3648	1.4387	1.4129
	3.7856	2.9128	3.4400	1.9722	2.0450	2.1320	2.7362	3.0278	2.3185	2.1055	2.1568	2.060
106.6	0.8716	1.0998	1.4396	1.3243	1.2660	1.2710	1.3063	1.3641	1.1337	1.3707	1.3699	1.3430
	2.0595	3.5045	2.8142	0.8366	1.8996	1.9017	2.040	2.3663	3.0103	2.2253	2.0447	1.9414
131.0	1.1017	1.1333	1.1835	1.1471	1.2451	1.2547	1.2624	1.2748	1.3179	1.1518	1.3768	1.3791
	4.6080	3.7127	3.7228	3.1046	1.9377	1.8477	1.8426	1.8956	2.0997	2.9732	2.1490	2.0619
162.0	0.8894	1.1091	1.2446	1.4713	1.2692	1.2635	1.2663	1.2652	1.2674	1.2897	1.1679	1.3794
	1.8539	3.4068	4.2135	1.9034	1.6319	1.9126	1.8088	1.8213	1.8395	1.9450	2.9192	2.2337
180.0	1.0537	1.0673	1.3307	1.2411	1.2369	1.2882	1.2774	1.2751	1.2713	1.2731	1.3207	1.1755
	2.5585	3.1909	5.2090	5.6088	1.7140	1.9389	1.8113	1.8261	1.8255	1.8484	2.1093	2.8872

Bibliography

Adelman-McCarthy, J. K., Agüeros, M. A., Allam, S. S., et al. 2008, ApJS, 175, 297

Albrecht, A., Bernstein, G., Cahn, R., et al. 2006, astro-ph/0609591

Allen, S. W., Rapetti, D. A., Schmidt, R. W., et al. 2008, MNRAS, 383, 879

Anderson, T. W. 2003, An Introduction to Multivariate Statistical Analysis (Wiley-Interscience), 623/624

Astier, P., Guy, J., Regnault, N., et al. 2006, A&A, 447, 31

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

Bacon, D. J., Massey, R. J., Refregier, A. R., & Ellis, R. S. 2003, MNRAS, 344, 673

Bartelmann, M. & Schneider, P. 2001, Phys. Rep., 340, 291

Basser, P. J. & Pajevic, S. 2002, in ISBI, 927–930

Basser, P. J. & Pajevic, S. 2007, Signal Process., 87, 220

Bekenstein, J. D. 2004, astro-ph/0412652

Benjamin, J., Heymans, C., Semboloni, E., et al. 2007, MNRAS, 381, 702

Bennett, C. L., Halpern, M., Hinshaw, G., et al. 2003, ApJS, 148, 1

Bernardeau, F. 1998, A&A, 338, 375

Bernardeau, F., Colombi, S., Gaztañaga, E., & Scoccimarro, R. 2002, Phys. Rep., 367, 1

Bridle, S., Shawe-Taylor, J., Amara, A., et al. 2008, arXiv0802.1214B

Brown, M. L., Taylor, A. N., Bacon, D. J., et al. 2003, MNRAS, 341, 100

Buchert, T. 2008, General Relativity and Gravitation, 40, 467

Burke, W. L. 1981, ApJ, 244, L1+

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

Chang, T.-C., Refregier, A., & Helfand, D. J. 2004, ApJ, 617, 794

Clarkson, C., Bassett, B., & Lu, T. H.-C. 2008, Physical Review Letters, 101, 011301

Clowe, D., Bradač, M., Gonzalez, A. H., et al. 2006, ApJ, 648, L109

Cole, S., Percival, W. J., Peacock, J. A., et al. 2005, MNRAS, 362, 505

Cooray, A. & Sheth, R. 2002, Phys. Rep., 372, 1

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

Crittenden, R. G., Natarajan, P., Pen, U.-L., & Theuns, T. 2002, ApJ, 568, 20

Dodelson, S. 2003, Modern Cosmology (Elsevier), 180 ff

Doran, M., Lilley, M., Schwindt, J., & Wetterich, C. 2001, ApJ, 559, 501

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

Eifler, T., Kilbinger, M., & Schneider, P. 2008a, A&A, 482, 9

Eifler, T., Schneider, P., & Hartlap, J. 2008b, arXiv0810.4254E

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

Eisenstein, D. J., Zehavi, I., Hogg, D. W., et al. 2005, ApJ, 633, 560

Evrard, A. E., MacFarland, T. J., Couchman, H. M. P., et al. 2002, ApJ, 573, 7

Felten, J. E. & Isaacman, R. 1986, Reviews of Modern Physics, 58, 689

Fischer, G. 1997a, Lineare Algebra (Vieweg), 161

Fischer, G. 1997b, Lineare Algebra (Vieweg), 276

Freedman, W. L., Madore, B. F., Gibson, B. K., et al. 2001, ApJ, 553, 47

Fu, L., Semboloni, E., Hoekstra, H., et al. 2008, A&A, 479, 9

Gaztanaga, E., Cabre, A., Castander, F., Crocce, M., & Fosalba, P. 2008a, arXiv0807.2448G

Gaztanaga, E., Cabre, A., & Hui, L. 2008b, arXiv0807.3551

Gelman, A., Carlin, J., Stern, H., & Rubin, D. 2004, Bayesian Data Analysis (Chapman & Hall/CRC)

Guth, A. H. 1981, Phys. Rev. D, 23, 347

Hamana, T., Miyazaki, S., Shimasaku, K., et al. 2003, ApJ, 597, 98

Hamilton, A. J. S., Kumar, P., Lu, E., & Matthews, A. 1991, ApJ, 374, L1

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

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

Henry, J. P. 2004, ApJ, 609, 603

Hetterscheidt, M., Simon, P., Schirmer, M., et al. 2007, A&A, 468, 859

Heymans, C., Brown, M., Heavens, A., et al. 2004, MNRAS, 347, 895

Heymans, C., Brown, M. L., Barden, M., et al. 2005, MNRAS, 361, 160

Heymans, C., Van Waerbeke, L., Bacon, D., et al. 2006, MNRAS, 368, 1323

Hilbert, S., Hartlap, J., White, S. D. M., & Schneider, P. 2008, arXiv0809.5035H

Hinshaw, G., Spergel, D. N., Verde, L., et al. 2003, ApJS, 148, 135

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

Hirata, C. M. & Seljak, U. 2003a, Phys. Rev. D, 67, 043001

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

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

Hoekstra, H., Mellier, Y., van Waerbeke, L., et al. 2006, ApJ, 647, 116

Hoekstra, H., van Waerbeke, L., Gladders, M. D., Mellier, Y., & Yee, H. K. C. 2002a, ApJ, 577, 604

Hoekstra, H., Yee, H. K. C., & Gladders, M. D. 2002b, ApJ, 577, 595

Hu, W. 2002, Phys. Rev. D, 65, 023003

Hu, W. & Dodelson, S. 2002, ARA&A, 40, 171

Hubble, E. 1929, Proceedings of the National Academy of Science, 15, 168

Huetsi, G. 2005, astro-ph/0507678

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

Jarvis, M., Bernstein, G. M., Fischer, P., et al. 2003, AJ, 125, 1014

Jarvis, M., Jain, B., Bernstein, G., & Dolney, D. 2006, ApJ, 644, 71

Jenkins, A., Frenk, C. S., White, S. D. M., et al. 2001a, MNRAS, 321, 372

Jenkins, A., Frenk, C. S., White, S. D. M., et al. 2001b, MNRAS, 321, 372

Jing, Y. P. 2002, MNRAS, 335, L89

Joachimi, B. & Schneider, P. 2008, A&A, 488, 829

Joachimi, B., Schneider, P., & Eifler, T. 2008, A&A, 477, 43

Kaiser, N. 1992, ApJ, 388, 272

Kaiser, N. 1994, in Clusters of Galaxies, ed. F. Durret, A. Mazure, & J. Tran Thanh van, 269–+

Kaiser, N. 1998, ApJ, 498, 26

Kaiser, N., Wilson, G., & Luppino, G. A. 2000, astro-ph0003338K

Kendall, M. & Stuart, A. 1979, The Advanced Theory of Statistics (Charles Griffin & Company Limited)

Kilbinger, M., Benabed, K., Guy, J., et al. 2008, arXiv0810.5129K

Kilbinger, M. & Munshi, D. 2006, MNRAS, 366, 983

Kilbinger, M. & Schneider, P. 2004, A&A, 413, 465

Kilbinger, M. & Schneider, P. 2005, A&A, 442, 69

Kilbinger, M., Schneider, P., & Eifler, T. 2006, A&A, 457, 15

King, L. J. & Schneider, P. 2003, A&A, 398, 23

Knop, R. A., Aldering, G., Amanullah, R., et al. 2003, ApJ, 598, 102

Kochanek, C. S. & Schechter, P. L. 2004, in Measuring and Modeling the Universe, ed. W. L. Freedman, 117–+

Komatsu, E., Dunkley, J., Nolta, M. R., et al. 2008, arXiv0803.0547K

Kravtsov, A. V., Vikhlinin, A., & Nagai, D. 2006, ApJ, 650, 128

Kuznetsova, N., Barbary, K., Connolly, B., et al. 2008, ApJ, 673, 981

Ledoit, O. & Wolf, M. 2004, Journal of Multivariate Analysis, 88, 365

Lewis, A. & Challinor, A. 2006, Phys. Rep., 429, 1

Linde, A. D. 1982, Physics Letters B, 108, 389

Linder, E. V. 2003, Physical Review Letters, 90, 091301

Loredo, T. 1990, From Laplace to Supernovae SN 1987A: Bayesian inference in Astrophysics (Kluiver Academic Publishers), 81–142

Mandelbaum, R., Hirata, C. M., Ishak, M., Seljak, U., & Brinkmann, J. 2006, MNRAS, 367, 611

Mantz, A., Allen, S. W., Ebeling, H., & Rapetti, D. 2008, MNRAS, 387, 1179

Maoli, R., Van Waerbeke, L., Mellier, Y., et al. 2001, A&A, 368, 766

Marian, L. & Bernstein, G. M. 2007, Phys. Rev. D, 76, 123009

Massey, R., Heymans, C., Bergé, J., et al. 2007a, MNRAS, 376, 13

Massey, R., Refregier, A., Bacon, D. J., Ellis, R., & Brown, M. L. 2005, MNRAS, 359, 1277

Massey, R., Rhodes, J., Leauthaud, A., et al. 2007b, ApJS, 172, 239

Ménard, B., Hamana, T., Bartelmann, M., & Yoshida, N. 2003, A&A, 403, 817

Pan, J., Coles, P., & Szapudi, I. 2007, MNRAS, 382, 1460

Peacock, J. 1999, Cosmological Physics (Cambridge University Press)

Peacock, J. A. & Dodds, S. J. 1994, MNRAS, 267, 1020

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

Peacock, J. A., Schneider, P., Efstathiou, G., et al. 2006, ESA-ESO Working Group, Report Nr. 3 "Fundamental Cosmology", (ESA)

Penzias, A. A. & Wilson, R. W. 1965, ApJ, 142, 419

Percival, W. J., Cole, S., Eisenstein, D. J., et al. 2007, MNRAS, 381, 1053

Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, ApJ, 517, 565

Pope, A. C. & Szapudi, I. 2008, MNRAS, 389, 766

Press, W., Teukolsky, S., Vetterling, W., & Flannery, B. 1992, Numerical Recipes in C (Cambridge University Press)

Press, W. H. & Schechter, P. 1974, ApJ, 187, 425

Refregier, A., Rhodes, J., & Groth, E. J. 2002, ApJ, 572, L131

Rhodes, J., Refregier, A., & Groth, E. J. 2001, ApJ, 552, L85

Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, AJ, 116, 1009

Riess, A. G., Li, W., Stetson, P. B., et al. 2005, ApJ, 627, 579

Riess, A. G., Strolger, L.-G., Casertano, S., et al. 2007, ApJ, 659, 98

Riess, A. G., Strolger, L.-G., Tonry, J., et al. 2004, ApJ, 607, 665

Sachs, R. K. & Wolfe, A. M. 1967, ApJ, 147, 73

Sánchez, A. G. & Cole, S. 2008, MNRAS, 385, 830

Schäfer, J. & Strimmer, K. 2005, Statistical Applications in Genetics and Molecular Biology, 4, Article 32

Schneider, P. 1996, MNRAS, 283, 837

Schneider, P. 2006, Einführung in die extragalaktische Astronomie und Kosmologie (Springer Science+Business Media)

Schneider, P. 2008, Cosmology lecture notes: University of Bonn

Schneider, P. & Kilbinger, M. 2007, A&A, 462, 841

Schneider, P., Kilbinger, M., & Lombardi, M. 2005, A&A, 431, 9

Schneider, P., Kochanek, C. S., & Wambsganss, J. 2006, Gravitational Lensing: Strong, Weak and Micro (Springer-Verlag Berlin)

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

Schneider, P., van Waerbeke, L., Kilbinger, M., & Mellier, Y. 2002a, A&A, 396, 1

Schneider, P., van Waerbeke, L., & Mellier, Y. 2002b, A&A, 389, 729

Schrabback, T., Erben, T., Simon, P., et al. 2007, A&A, 468, 823

Schramm, T. & Kayser, R. 1995, A&A, 299, 1

Scoccimarro, R. & Couchman, H. M. P. 2001, MNRAS, 325, 1312

Scoccimarro, R. & Frieman, J. A. 1999, ApJ, 520, 35

Seitz, C. & Schneider, P. 1995, A&A, 297, 287

Seitz, C. & Schneider, P. 1997, A&A, 318, 687

Semboloni, E., Mellier, Y., van Waerbeke, L., et al. 2006, A&A, 452, 51

Semboloni, E., van Waerbeke, L., Heymans, C., et al. 2007, MNRAS, 375, L6

Sheth, R. K. & Tormen, G. 1999, MNRAS, 308, 119

Simon, P., Hetterscheidt, M., Schirmer, M., et al. 2007, A&A, 461, 861

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

Smith, R. E., Peacock, J. A., Jenkins, A., et al. 2003, MNRAS, 341, 1311

Smoot, G. F., Bennett, C. L., Kogut, A., et al. 1991, ApJ, 371, L1

Spergel, D. N., Bean, R., Doré, O., et al. 2007, ApJS, 170, 377

Springel, V., White, S. D. M., Jenkins, A., et al. 2005, Nature, 435, 629

Sugiyama, N. 1995, ApJS, 100, 281

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

Trefethen, L. N. & Bau, D. 1997, Numerical Linear Algebra (Society for Industrial and Applied Mathematics, Philadelphia)

van Waerbeke, L. & Mellier, Y. 2003, astro-ph/0305089

van Waerbeke, L., Mellier, Y., Erben, T., et al. 2000, A&A, 358, 30

van Waerbeke, L., Mellier, Y., & Hoekstra, H. 2005, A&A, 429, 75

Van Waerbeke, L., Mellier, Y., Pelló, R., et al. 2002, A&A, 393, 369

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

Wiltshire, D. L. 2007a, New Journal of Physics, 9, 377

Wiltshire, D. L. 2007b, Physical Review Letters, 99, 251101

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

Wood-Vasey, W. M., Friedman, A. S., Bloom, J. S., et al. 2007, arXiv0711.2068

Yadav, A. P. S. & Wandelt, B. D. 2008, Physical Review Letters, 100, 181301